



## Upstream land use with microbial downstream consequences: Iron and humic substances link to *Legionella* spp.

Jon Ahlinder<sup>a,b,1</sup>, Karolina Ida Anna Eriksson<sup>c,d,1,\*</sup>, Maria Hellmér<sup>g</sup>, Emelie Salomonsson<sup>a</sup>, Malin Granberg<sup>a</sup>, Ingrid Dacklin<sup>a</sup>, Josefine Elving<sup>e</sup>, Björn Brindefalk<sup>a,f</sup>

<sup>a</sup> Division of CBRN Defence and Security, Swedish Defence Research Agency (FOI), Sweden

<sup>b</sup> Department of Tree Breeding, Skogforsk, Sävar, SE-91821, Sweden

<sup>c</sup> Department of Ecology and Environmental Sciences, Faculty of Science and Technology, Umeå University, Sweden

<sup>d</sup> Umeå Marine Sciences Centre, Umeå University, Hörnefors, 905 71, Sweden

<sup>e</sup> Department of Chemistry, Environment and Feed Hygiene, Swedish Veterinary Agency, Sweden

<sup>f</sup> Department of Molecular Biosciences, The Wenner-Gren Institute, Stockholm University, Sweden

<sup>g</sup> Department of Biology, Science Division, Swedish Food Agency, Sweden

### ARTICLE INFO

#### Keywords:

Land-use effects on microbiota and water quality  
Water safety  
Water resources/management  
Anthropogenic effects  
bacterial pathogens

### ABSTRACT

Intensified land use can disturb water quality, potentially increasing the abundance of bacterial pathogens, threatening public access to clean water. This threat involves both direct contamination of faecal bacteria as well as indirect factors, such as disturbed water chemistry and microbiota, which can lead to contamination. While direct contamination has been well described, the impact of indirect factors is less explored, despite the potential of severe downstream consequences on water supply. To assess direct and indirect downstream effects of buildings, farms, pastures and fields on potential water sources, we studied five Swedish lakes and their inflows. We analysed a total of 160 samples in a gradient of anthropogenic activity spanning four time points, including faecal and water-quality indicators. Through species distribution modelling, Random Forest and network analysis using 16S rRNA amplicon sequencing data, our findings highlight that land use indirectly impacts lakes via inflows. Land use impacted approximately one third of inflow microbiota taxa, in turn impacting ~20–50 % of lake taxa. Indirect effects via inflows were also suggested by causal links between e.g. water colour and lake bacterial taxa, where this influenced the abundance of several freshwater bacteria, such as *Polynucleobacter* and *Limnohabitans*. However, it was not possible to identify direct effects on the lakes based on analysis of physicochemical- or microbial parameters. To avoid potential downstream consequences on water supply, it is thus important to consider possible indirect effects from upstream land use and inflows, even when no direct effects can be observed on lakes. *Legionella* (a genus containing bacterial pathogens) illustrated potential consequences, since the genus was particularly abundant in inflows and was shown to increase by the presence of pastures, fields, and farms. The approach presented here could be used to assess the suitability of lakes as alternative raw water sources or help to mitigate contaminations in important water catchments. Continued broad investigations of stressors on the microbial network can identify indirect effects, avoid enrichment of pathogens, and help secure water accessibility.

### 1. Introduction

Land use and climate change can induce imbalances in the microbial community that threaten reliable access to clean water, since many

bacterial pathogens are opportunistic and naturally occurring. The availability of clean water is crucial for all societies and intrinsically linked to human health. Access to clean water has its own sustainable development goal (UN General Assembly, 2015) and is a central global

**Abbreviations:** ASV, amplicon sequence variant; CODMN, chemical oxygen demand; DAG, directed acyclic graphs; GLMM, generalized linear mixed models; RF, random forest.

\* Corresponding author.

1 Authors contributed equally.

E-mail address: [karolina.eriksson@umu.se](mailto:karolina.eriksson@umu.se) (K.I.A. Eriksson).

<https://doi.org/10.1016/j.watres.2024.121579>

Received 31 August 2023; Received in revised form 28 March 2024; Accepted 5 April 2024

Available online 9 April 2024

0043-1354/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

societal risk factor (World Economic Forum, 2023). Increasing population growth and climate change are pressuring ecosystems that have good water quality (Vörösmarty et al., 2000). Stressors from land use (e.g., farms, fields, buildings, and pasture) and climate change (e.g., humidification caused by increased precipitation) can lead to increasing nutrient loads, especially in boreal lake systems used as water sources for many communities in temperate regions. Importantly, bacteria not only cause problems but are also essential to functioning ecosystems, where a balanced network of microbes provide ecosystem services such as degradation of organic matter. Stressors can lead to ecosystem imbalances and consequences for water quality due to changes in the microbial community. Humans and animals can then be exposed to the microbes when the water is used recreationally or after treatment for drinking water. Water quality reduction is often associated with increases in turbidity, colour, or faecal bacteria. In Sweden, risks to reliable clean water access are mainly associated with disease caused by faecal bacteria like *Campylobacter*, pathogenic *E. coli* and *Shigella* (Public Health Agency of Sweden, 2015), but also include environmental pathogens that can cause disease such as diarrhoea (e.g. *Vibrio*) or pulmonary infection when inhaled such as several species of e.g. *Legionella* (World Health Organization, 2011).

Waterborne pathogens can be part of the natural microbial community and are often environmental opportunists (Anttila et al., 2015). *Legionella* is such an opportunistic bacteria that has its natural life cycle in water and soil as part of biofilms and intracellular parasites of protozoans, sporadically causing problems in water supplies (Hamilton et al., 2018; Leoni et al., 2018; Holsinger et al., 2022). *Legionella* is a public-health-relevant genus, with over 20 of 71 known species being potential human pathogens and cases of disease have been coupled to local watershed hydrology (Ng et al., 2008). Although *Legionella* abundance has been shown to increase near agricultural sites and to vary depending on land use type in Canada (Peabody et al., 2017), there is still a knowledge gap regarding what selects for these species in environmental water (Eriksson et al., 2022) and how they colonize human-made water systems.

Studies of threats to water quality often focus on separate factors, such as the composition of microorganisms in raw water (Comté et al. 2017; Brindefalk et al. 2022), riverine inflows (Wu et al. 2022), land use effects (Schelker et al. 2016), and nutrient regimes (Soranno et al. 2015; Yang and Kim, 2019). For example, Brindefalk et al. (2022) investigated water quality of six important raw water sources in Sweden across a gradient of land-use factors as proportion of farmland and urban area at the respective water catchments, and found associations of bacterial metacommunities with water quality indicators and degree of urbanisation in the catchment area. Soranno et al. (2015) investigated the effects of spatial extent for measuring land use, hydrologic connectivity, and the regional differences on the amount of nutrients in 346 lakes in the USA, where the strongest association was found to be due to lake hydrologic connectivity. However, studying these factors in isolation hinders identification of the complex systems governing water quality, since these parameters may display effects only when analysed jointly, not individually. This includes factors influencing the lake inflows as well as factors influencing the lakes themselves. Stressors on the microbial network are fundamental, so a broader approach could help in securing water accessibility (Fig. 1a).

In this study of five Swedish lakes, together with their inflows and surrounding land use, we addressed these questions (Fig. 1b-c): I) Does land use influence lake inflows? II) Do upstream inflows influence the lakes? III) Does the disturbance due to land use/anthropogenic action have a measurable effect on the lakes? IV) Does land use affect the occurrence of the genus *Legionella* (exemplifying how upstream land use could affect downstream water quality)?

## 2. Materials and methods

### 2.1. Parameters of study waters

Five lakes (Supplementary Fig. 1, coded as By, Ho, Ok, Sk, and Sv) with potential for surface raw water use were sampled in southern Sweden, in an area with mixed, but generally low, levels of human disturbance, totalling 160 samples. The samples were taken in October 2018, November 2018, May 2019, and August 2019, capturing the full yearly variability. On each sampling occasion, water-samples from inflows to the lake (referred to as “inflow” samples), shore lake water and a mid lake sampling point were collected (both referred to as “lake” samples). Inflows relevant to sampling based on land use were identified using available maps (e.g., WaterInformationSystem Sweden [WISS]) and aerial imagery, selecting streams assumed to contribute to lake contamination, i.e., passing through or near agricultural land or buildings. To identify anthropogenic influences, all lake inflows were screened for the proximity of farms, fields, buildings, and pasture (Table 1, visualized for one lake in Fig. 1b).

In addition to the condensed materials and methods below, in depth details concerning indicator analysis, sample preparation methods for DNA extraction, sequencing, qPCR, and the GLMM of bacterial ASVs and *Legionella* specific ASVs can be found in the Supplementary Material (Supplementary Fig. 1–13).

### 2.2. Sample collection

Lake samples were collected from the shore by lowering the sample bottles under the water surface and by boat at depths of either 5 or 10 m using a Ruttner water sampler (Hydro-X 1.7 L; Swedaq, Höör, Sweden), after which water was distributed into sampling bottles. Samples from the inflows to the lakes were collected from land, downstream from potential contamination points near the lakeside, by lowering the sample bottles under the water surface (Fig. 1a, Supplementary Fig. 1). Disposable gloves were used during all sampling to avoid sample contamination.

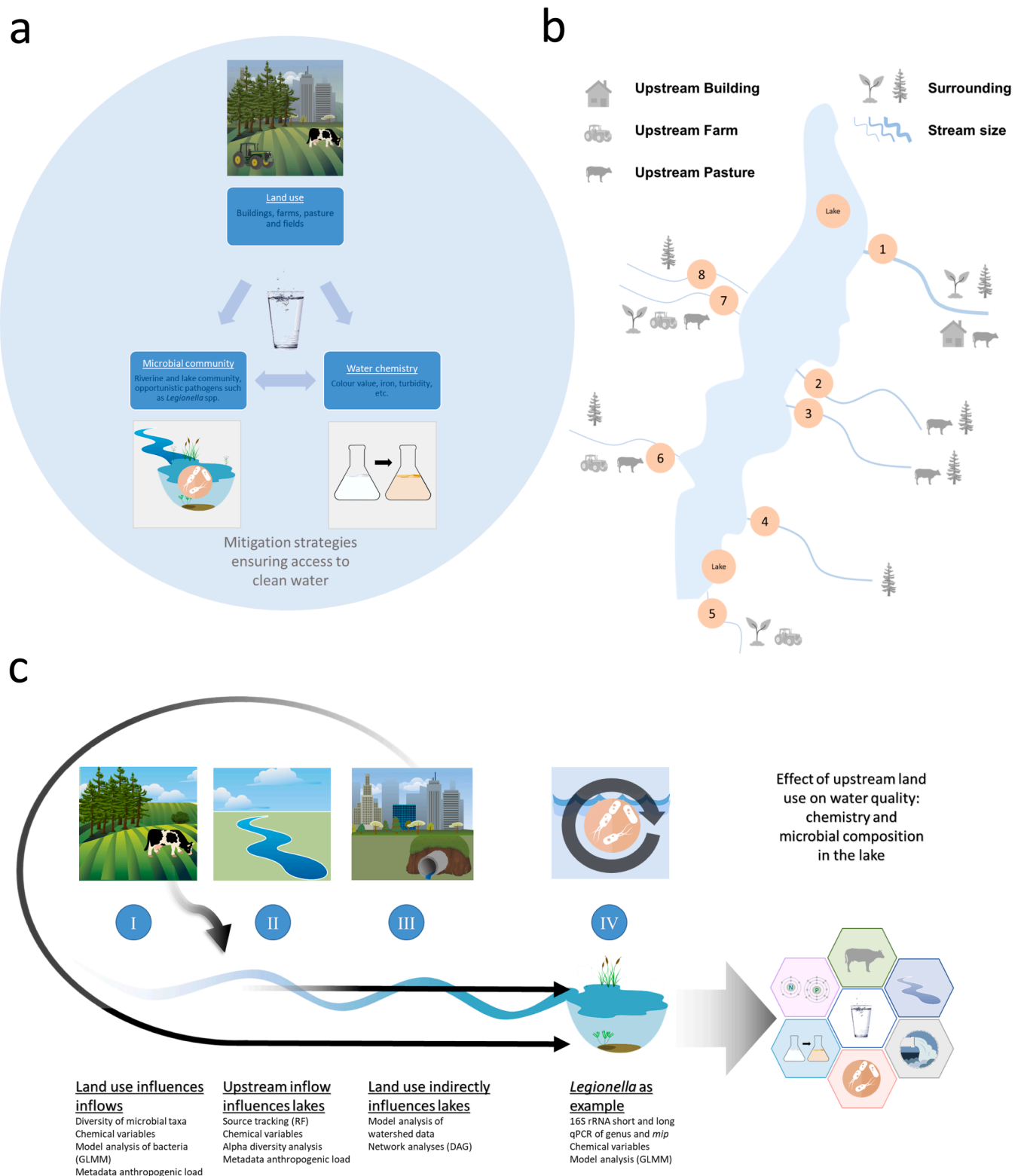
### 2.3. Chemical analysis

We investigated a wide range of water property indicators covering effects of disturbed inflowing water with respect to land-use effects. Chemical analysis of the collected samples was performed by Eurofins. All samples, except those collected at the potential point for withdrawal of raw water in November 2018 and August 2019, were analysed using the PSL2V package for drinking water chemical analysis (supplied by Eurofins Environment Testing Sweden AB). Samples collected at the potential point for withdrawal of raw water in November 2018 and August 2019 were subjected to an extended analysis using the PSL3Z package that, in addition to physical chemical parameters and elements, also includes pesticides.

### 2.4. DNA sequencing and classification, molecular analyses

For the full set of samples Illumina sequencing was performed on the V3-V4 16S rRNA gene region using primers Bact\_341F and Bact\_785R. For a subset of samples full-length sequencing was done on the full 16S rRNA gene using Oxford Nanopore technology using primers Bact\_27F and Bact\_1492R. Resulting sequences were subjected to standard quality control and trimming and classified against the SILVA no. 99 v. 138 dataset for V3-V4 sequences and a database based on GTDB version Release 08-RS214 for the full length sequences.

For phylogenetic placement of sequences assigned to the genus *Legionella* a reference dataset consisting of 17 *Aquicella* sequences (as outgroup) and 393 *Legionella* sequences was assembled and phylogenetic relationships were inferred using a maximum likelihood approach (IQ-TREE v. 2.0.7, Nguyen et al., 2015). For the detection of L.



**Fig. 1.** Experimental design. Panel a: Land use can affect the microbial community and water chemistry. Also, the microbial community and water chemistry can affect each other. This study therefore focuses on a holistic approach examining the potential effects of upstream land use on water quality. Panel b: Sampling design and land use categories for one of the five lakes. For each of the four sampling occasions, stream inflows (numbered 1–8) and lake water including shore water (both referred to as “Lake”) was sampled. Panel c: Hypotheses and analyses for the study of the effect of upstream land use on the water quality. I: Land use does have an influence on the water properties; II: The upstream inflows do have an effect on the overall water properties and the composition of the microbiota in the lakes; III: The degree of disturbance of water quality due to land use/anthropogenic action does affect the inflows and has a measurable effect on the lakes; IV: Land use can affect the occurrence of genus *Legionella*, a potentially pathogenic organism that could cause problems in raw water, demonstrating how upstream land use could impact downstream water quality.

**Table 1**

Sampling points included in the statistical analyses, using representatives from each lake and chemical data (Supplementary Fig. 4) for the four sampling occasions. These points also represent mixed levels of land use (Fields, Buildings, Farms and Pastures), including “undisturbed” inflows (all 0).

SamplePoint	Type	Surroundings	UpstreamBuildings	UpstreamFarms	UpstreamPastures	Size(stream)
By1	Inflow	Field	1	0	1	Middle
By3	Inflow	Field	0	0	0	Middle
ByMidlake	Lake	NA	NA	NA	NA	NA
ByShore	Lake	NA	NA	NA	NA	NA
Ho1	Inflow	Forest	1	0	0	Small
Ho2	Inflow	Field/Forest	1	0	1	Small
Ho3	Inflow	Forest	0	0	0	Small
Ho4	Inflow	Forest	1	1	0	Small
Ho5	Inflow	Field	1	1	0	Middle
Ho8	Inflow	Forest	1	0	0	Middle
HoMidlake	Lake	NA	NA	NA	NA	NA
HoShore	Lake	NA	NA	NA	NA	NA
Ok1	Inflow	Field/Forest	1	0	1	Large
Ok5	Inflow	Field	0	1	0	Small
OkMidlake	Lake	NA	NA	NA	NA	NA
OkShore	Lake	NA	NA	NA	NA	NA
Sk3	Inflow	Forest	0	0	0	Small
Sk4	Inflow	Field/Forest	1	0	0	Small
SkMidlake	Lake	NA	NA	NA	NA	NA
SkShore	Lake	NA	NA	NA	NA	NA
Sv1	Inflow	Forest	0	0	0	Small
Sv2	Inflow	Field	0	1	0	Middle
Sv4	Inflow	Forest	0	0	0	Middle
SvMidlake	Lake	NA	NA	NA	NA	NA
SvShore	Lake	NA	NA	NA	NA	NA

*pneumophila* by qPCR, the specific *mip* gene target was used (Nazarian et al., 2008).

### 2.5. Generalized linear mixed effect modelling

For questions I: land use effects on lake inflows, III: disturbance of land use on the lakes, and IV: land use effects on the occurrence of the genus *Legionella*, the following generalized linear mixed model (GLMMs) were used to estimate land-use effects on the water quality of the inflows and lakes:

$$g(m_{ij}) = \alpha_i + \beta_{1jk}x_{1i} + \beta_{2jl}x_{2i} + \beta_{3jn}x_{3i} + \beta_{4lm}x_{4i} + \beta_{5jp}x_{5i} + \beta_{6jq}x_{6i} + \beta_{7jr}x_{7i} + \beta_{8js}x_{8i} + \beta_{9jt}x_{9i}, \tag{1}$$

where  $g()$  is the log link function defining the mean of the linear function of predictors,  $m_{ij}$  is the  $j$ th ASV abundance in sample  $i$  ( $i = 1, \dots, n_{\text{tot}}$ ) with  $n_{\text{tot}}$  being the total number of samples included ( $n_{\text{tot}} = 115$  for inflow analysis, and  $n_{\text{tot}} = 63$  for physiochemical analysis),  $\alpha_{0i}$  is the effect of the total sample sequence abundance of sample  $i$ ,  $\beta_{1ik}$  is the effect of time point  $k$  when collecting sample  $i$  on ASV  $j$  ( $k = 1, \dots, 4$ ),  $\beta_{2il}$  is the effect of sampling year  $l$  on ASV  $j$  ( $l = 0, 1$ , corresponding to 2017 or 2018),  $\beta_{3jn}$  is the effect of the lake system (level 3 in the model hierarchy) for sample  $i$  on ASV  $j$  ( $n = 1, \dots, 5$ ),  $\beta_{4jno}$  is the effect of sampling point  $o$  in lake system  $n$  (level 2 in the model hierarchy) for sample  $i$  on ASV  $j$  ( $n = 1, \dots, 5$ , with  $m$  differing depending on the lake),  $\beta_{5jp}$  is the effect of pasture near the sampling location  $i$  on the abundance of ASV  $j$  ( $p = 0, 1$ ),  $\beta_{6jq}$  is the stream size ( $q = 1, 2, \text{ or } 3$  corresponding to small, medium, or large),  $\beta_{7jr}$  is the effect of farm presence on ASV  $j$  ( $r = 0, 1$ ),  $\beta_{8js}$  is the effect of building presence on ASV  $j$  ( $s = 0, 1$ ), and  $\beta_{9jt}$  is the effect of agricultural field presence in the surroundings on ASV  $j$  ( $t = 0, 1$ ).

The regression models were implemented and analysed using the RStan library (Carpenter et al., 2017) within the R statistical software (R Core Team, 2021). Eight parallel chains were used to collect posterior draws with a total of 2000 iterations per chain with the first 1000

iterations removed (as burn-in), providing 8000 iterations in total. Results were visualized using R package ggplot2 (Wickham, 2016). R code and data for reproducing the results are available at [www.github.com/jonhar97/LandUseEffects](http://www.github.com/jonhar97/LandUseEffects).

### 2.6. Random forest: source tracking

To find the degree of inflow source contribution to the lake samples (i.e., question II), the random forest (RF) classification method implemented in the RandomForest R package was used (Liaw and Wiener, 2002). The source environments (i.e., training data) were defined as:

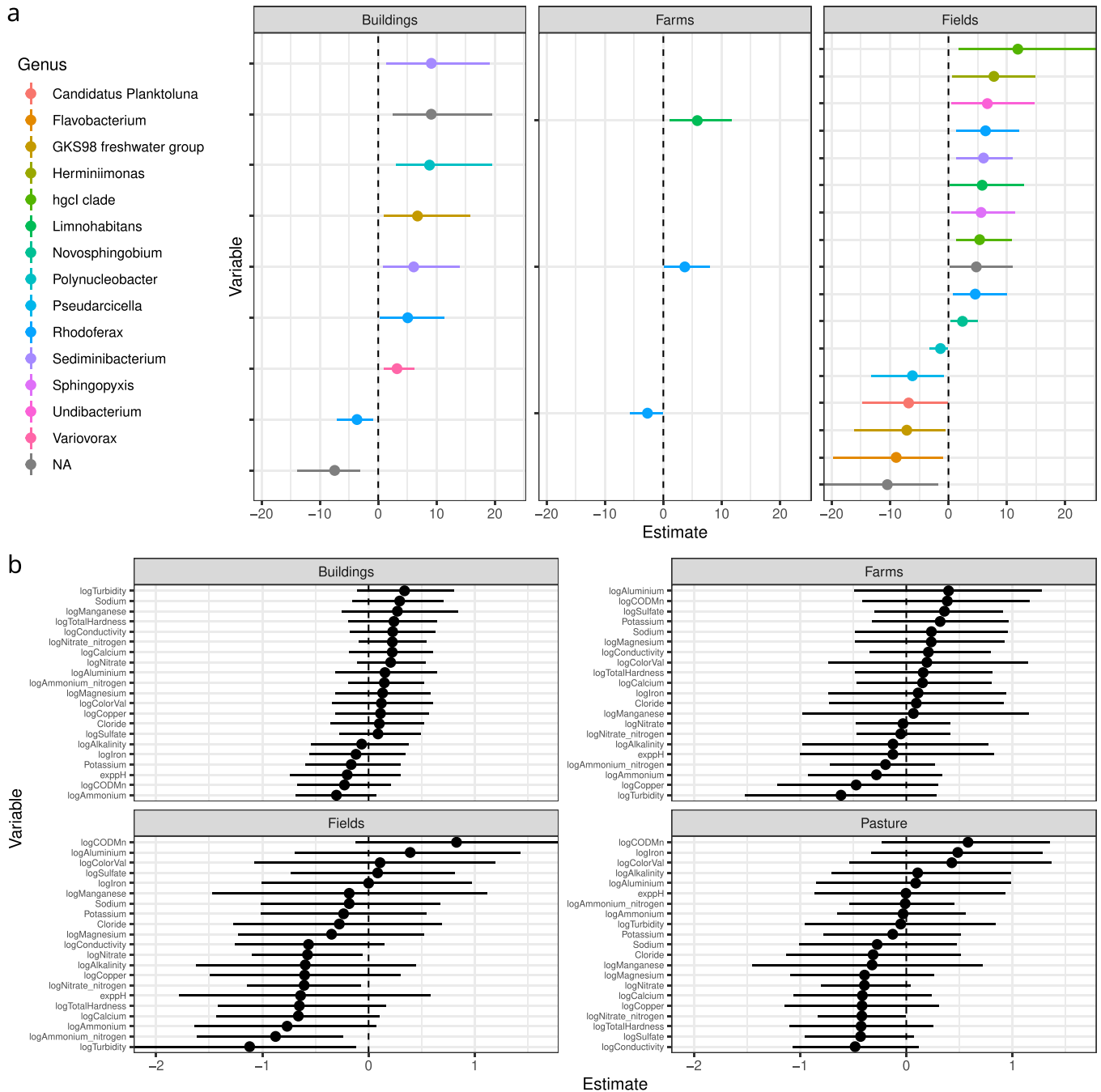
inflows ( $n = 115$  and 39 samples for ASV and physiochemical signatures, respectively) and lakes ( $n = 32$  and 24, respectively). The bacterial signatures consisted of the abundance data for the same top 100 ASVs as in the GLMM modelling. The 21 variables forming the physiochemical signature were first log transformed, apart from chloride, sodium, and potassium, which already displayed good distributional properties. Both sets of signatures were then scaled to zero mean and unit variance. RF models were first trained on the training data via a grid search of parameters to minimize the out-of-bag prediction error:  $n_{\text{tree}}$  (i.e., number of trees in the forest) = 250, 500, 1000, and 1500;  $m_{\text{try}}$  (i.e., number of candidates drawn to feed the algorithm) = 2, 3, 4, and 5; and minimum node size (i.e., depth of trees) = 1, 2, 10, and 20. After model optimization, the source label was removed for one of the lake samples (i.e., leave-one-out cross-validation) and the mixing proportions of the source environment were predicted for all lake samples.

### 2.7. Directed acyclic graphs (DAG)

To investigate if land use affect the lake water quality (i.e. question III), directed acyclic graphs (DAGs) were used to predict causal links between bacteria amplicon sequence variants (ASVs) affected by land-

use factors and water property variables, to capture possible indirect effects. Dynamic Bayesian Networks were used to model dependencies between land use-associated bacterial taxa and physiochemical variables. The inferred topology (i.e. relationships between nodes: here, ASVs and physiochemical variables) can be interpreted as chain effects of many dependent variables which might not be detected if analyzed separately. To infer causal relationships between nodes and not confound the results with design effects, we first corrected the bacteria abundance and physiochemical variables for design factors (i.e., removing effects of lake system, sampling year, and season), by using the GLMM outlined above. In this way, we could improve on earlier applications of DAG to microbial data.

First, the data were normalized such that the counts of all samples were equal in the data (to compensate for differences in the total number of reads between samples). The network was created by collapsing the V3-V4 16S rRNA ASV sequence data to the highest taxonomic resolution possible (genus). Selected environmental variables were included to infer dependencies (edges) between bacteria and environmental factors (nodes), and thus potential drivers of community structure. To avoid including highly correlated variables as predictors in the model, a pre-selection step was performed based on the PCA score plot of the environmental variables (i.e., only one of a group of tightly clustered variables was selected). Then, all features (i.e., bacteria and environmental variables) were normalized such that the mean value was centred at zero



**Fig. 2.** Effects of land-use on bacteria. Panel a: significant GLMM predictions of the estimated effects of land-use predictors on ASV abundance. Non-zero inclusion in estimated credible intervals were defined as significant. The bars highlight 2.5 % and 97.5 % of the posterior density, while the posterior mean is indicated with a dot. Colours highlight taxonomic assignment at the genus level. Panel b: effect of land use on physiochemical variables in the inflows for 39 samples for which extended chemical data were available. The bars show 5 % and 95 % of the posterior density, while the posterior mean is shown with a dot.

and the variation was transformed to one to enable comparison across datasets.

The network was inferred using BANJO version 2.2.0 (<http://www.cs.duke.edu/~amink/software/banjo/>, Smith et al., 2002) using the following analysis conditions: discretization = 2 intervals; max parents = 3; min lag = 0; max lag = 0; and the Simulated Annealing search algorithm. These parameter choices were based on preliminary runs minimizing the negative log-likelihood score (i.e., resulting in best goodness of fit to the data). For other settings, default parameters were used.

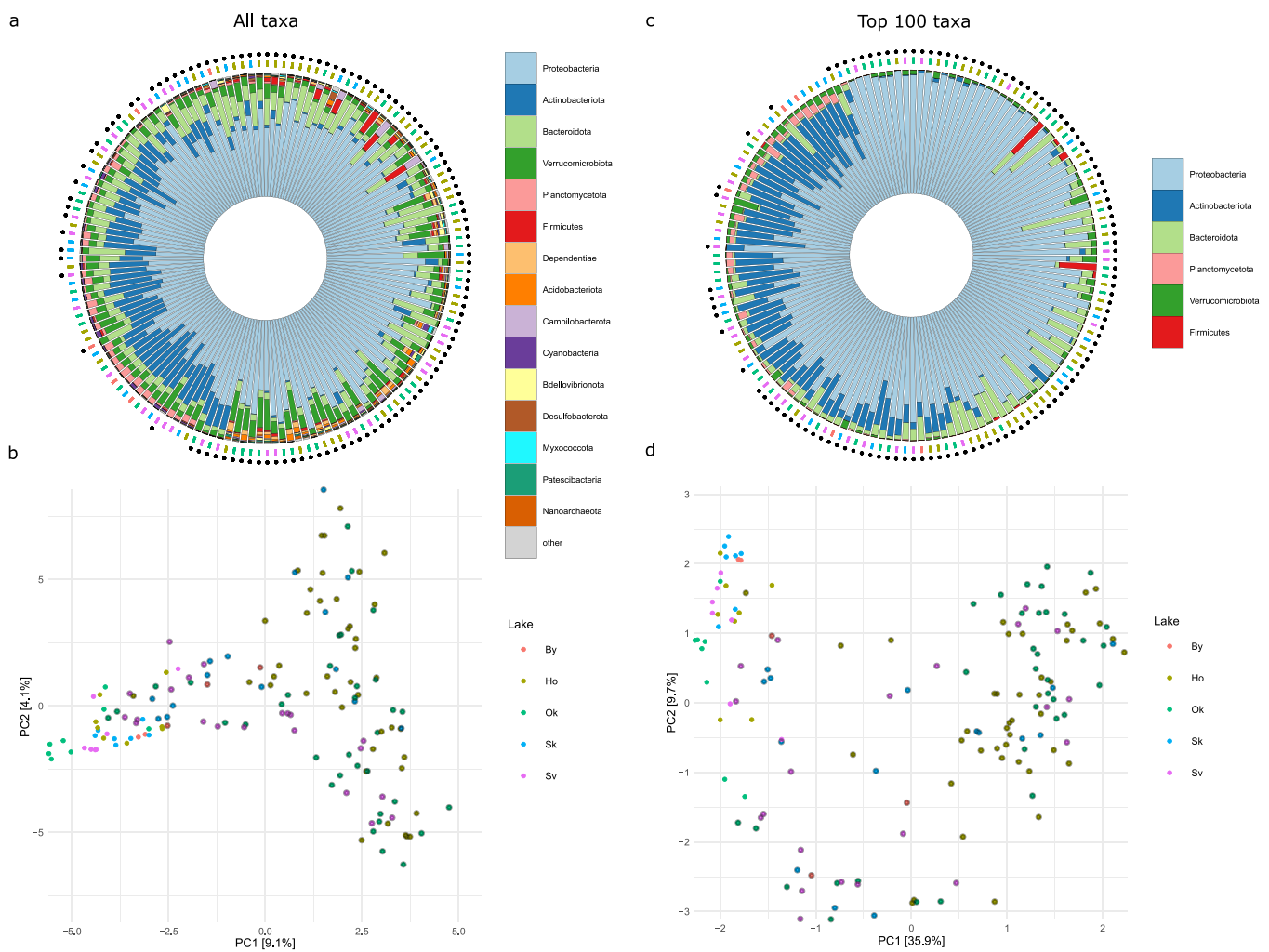
### 3. Results

#### 3.1. The bacterial composition of lake inflows was influenced by land use

The bacterial composition of the lake inflows was influenced by the presence of buildings and agricultural fields, where many common ASVs (of the top 100) increased in relative abundance due to anthropogenic action (Fig. 2a). To identify anthropogenic influences on bacterial communities and physiochemical signatures, all lake inflows were screened for anthropogenic factors (by presence/absence) near the sampling point: farms (small scale crop production and animal husbandry), fields (under active cultivation), buildings (single-family homes and farm outbuildings), and pasture (horses, goats, cows)

(Fig. 1b, Table 1). To infer the associations between ASVs and anthropogenic factors, a GLMM was fitted to the abundance data of the 115 inflow samples. Of the top 100 most abundant ASVs, 27 were associated with one or more of the land-use categories (Fig. 2a, Table 1), with both buildings ( $n_{\text{taxa}} = 9$ ) and agricultural fields ( $n_{\text{taxa}} = 17$ ) having the highest number of associations, while farms ( $n_{\text{taxa}} = 3$ ) and pasture ( $n_{\text{taxa}} = 0$ ) were less influential. For these 27 ASVs, most associations (20/29) were positive, indicating higher frequency (in relative composition), compared to undisturbed environments, under the influence of anthropogenic factors. The most common bacterial order among the associated set of ASVs was Burkholderiales, represented by the genera *Limnochabellites*, *Variovorax*, *Undibacterium*, *Polynucleobacter*, *Rhodiferax*, and *Hermimonas* (Supplementary Table 2). In addition, we investigated the impact of land use on bacterial alpha diversity: the presence of land use factors resulted in a predicted increase in the species diversity in most cases (Supplementary Fig. 3). The model predicts that farm presence would increase the Fisher alpha diversity by approximately 35 % in the inflows, all else being equal. Both bacterial composition and diversity thus displayed similar patterns, with the presence of farms, fields, and buildings producing a positive response. None of the classic faecal indicators (coliforms, *Escherichia coli*, enterococci and *Clostridium perfringens*) were positively influenced by land use factors.

A smaller proportion (<20 %) of water physiochemical variables were affected by anthropogenic activity (Fig. 2b: GLMM of 21 physiochemical

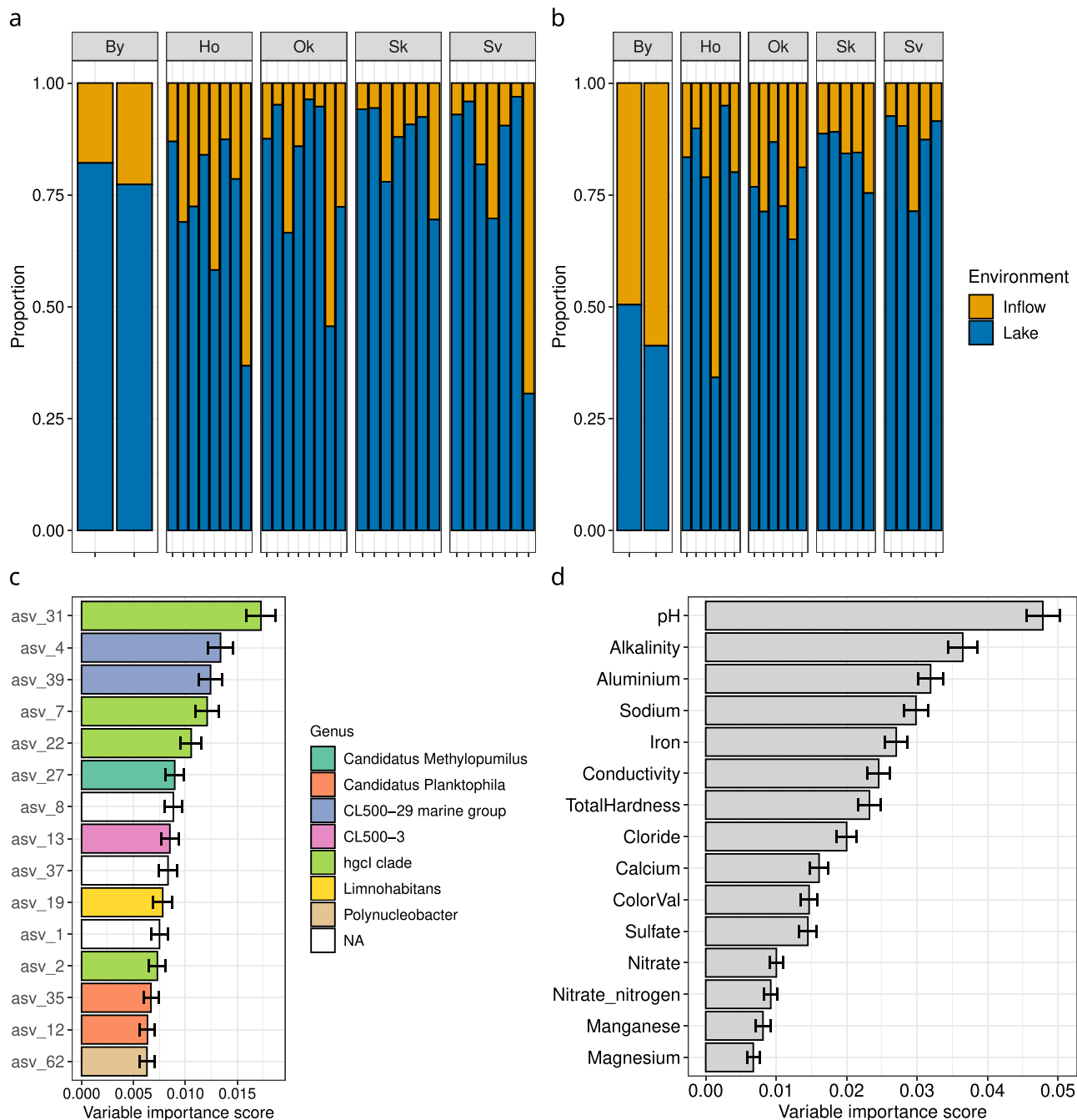


**Fig. 3.** Bacterial community composition. Panel a Iris plot showing the taxonomic composition of the samples on the phylum level for all taxa. Black dot: inflow samples. Panel b Centre-log-transformed PCA ordination of samples and all taxa. No outline: lake samples, Black outline: inflow samples. Panel c Iris plot showing just the top 100 taxa. Panel d Centre-log-transformed PCA ordination showing just the top 100 taxa. No outline: lake samples, Black outline: inflow samples.

variables and 39 samples), compared to the proportion of bacterial ASVs (27 %). Physiochemical measurements tended to increase in the presence of buildings (15 out of 21 estimates) and farms (13/21), while the measurements tended to decrease for fields (16/21) and pastures (14/21). Fields were the most influential of the anthropogenic factors: turbidity ( $\beta_{9,turb} = -1.11$ , 90 % CI  $[-2.21, -0.11]$ ), nitrate nitrogen ( $\beta_{9,nit} = -0.61$ ,  $[-1.14, -0.07]$ ), nitrate ( $\beta_{9,nitr} = -0.57$ ,  $[-1.10, -0.06]$ ), and ammonium nitrate ( $\beta_{9,amm} = -0.88$ ,  $[-1.62, -0.24]$ ). Nitrate was negatively associated with pastures ( $\beta_{5,nitr} = -0.42$ ,  $[-0.83, -0.01]$ ). To summarize, anthropogenic activity affected bacterial community composition and water chemistry in these inflows.

### 3.2. The variable and diverse bacteria in the inflows influenced the composition in the lakes

To examine how much influence the bacterial composition from the inflow had on the lake water composition, we first performed alpha- and beta-diversity analyses for the taxa variation. The inflows displayed higher diversity than did the lakes (Figs. 3b, 3d). The variability in the inflows was higher for both the bacterial composition and most physiochemical signatures, such as iron, colour value, and turbidity, throughout the study period. Furthermore, the alpha diversity (Supplementary Fig. 5) was higher in the inflows than in the lakes (Welch



**Fig. 4.** Predicted source environment for physiochemistry and bacteria. Random forest (RF) analysis of the predicted source of each feature (physiochemical signatures and V3-V4 16S rRNA ASVs). Panel a: inferred source proportion of the ASV profile data for each lake. Panel b: inferred source proportion of the physiochemical profile data for each lake. Panel c: variable importance score (i.e. mean decrease in accuracy if the variable is left out) for each of the top 15 bacterial ASVs highlighted at the genus level for discriminating the inflow and lake sources. Panel d: variable importance score for discriminating lake and inflow sources using the top 15 physiochemical variables.

Two Sample *t*-test,  $t = 4.61$ ,  $p$ -value =  $1.76e-05$ ). Most ASVs found in the lakes were also found in the inflows, while the inflows had a large observed proportion of unique ASVs (Supplementary Fig. 6). The lakes had higher relative abundances of Actinobacteria and Planctomycetota (Figs. 2a and 2c), while the inflows had higher relative abundances of Proteobacteria and Bacteroidota. The inflow communities contained a higher number of taxa, which were more variable (PERMANOVA;  $f = 8.84$ ,  $p$ -value = 0.001, PERMDISP;  $f = 60.2$ ,  $p$ -value =  $1.42e-12$ ) and diverse than in the lake communities.

To better understand how much the inflow influenced the lake water with respect to the bacterial and physiochemical compositions, we used the random forest (RF) classification method to model whether the bacteria (115 samples) and water property (39 samples) signatures in the lakes originated from the inflows. In general, most of the data were assigned to the lake environment (Fig. 4a–b), with some variation: the mean proportion of inflow in the lakes was 0.26 (0.17) for ASVs and 0.23 (0.17) for the physiochemistry, with standard deviations within parentheses. In some cases, lake samples consisted of more than 50 % inflow signature: one sample each for ASVs in lakes Ho and Sv, and one sample each for physiochemistry in By and Ho. Some bacterial taxa and physiochemical variables were especially influential on the water composition of the lakes (Fig. 4c–d). For example, the occurrence of ASV31 from the hgcl clade was almost four times higher in relative abundance in lake communities (mean frequency:  $\mu_{\text{Lake,asv31}} = 0.031$  (0.023),  $\mu_{\text{Inflow,asv31}} = 0.008$  (0.066)). One ASV from the hgcl clade was preferentially found in the inflow environment, as was the CL500–29 marine clade. In the model, pH and alkalinity had the highest discriminating power (Fig. 4d). Both pH and alkalinity levels were elevated in the inflows:  $\mu_{\text{Lake,pH}} = 7.08$  (0.13),  $\mu_{\text{Inflow,pH}} = 6.35$  (0.78), and  $\mu_{\text{Lake,Alk}} = 8.33$  (2.25),  $\mu_{\text{Inflow,Alk}} = 7.53$  (11.00). To summarize, some ASVs and physiochemical variables that were associated with the inflows considerably influenced the water properties in the lakes.

### 3.3. Downstream indirect influence of upstream characteristics on composition of microbiota

To investigate the land-use effects in the watersheds, we fitted a GLMM to the lake bacterial communities ( $n = 35$ ). For the analysis, the top 100 ASVs were selected according to their overall frequency in the lake assemblages. No significant effects (the predicted distributions overlap zero) on the lake ASVs were obtained for the proportion of fields or forest in the watershed, the lake waterline length, or the water body volume (a proxy for the sensitivity to perturbations) (Supplementary Table 5) on the lake ASVs were obtained (Supplementary Fig. 7).

For the holistic understanding of possible indirect downstream effects, DAG analysis was used to predict causal links between land use affected ASVs and water property variables. Since land use can affect the microbial community and water chemistry, and the microbial community and water chemistry can affect each other, indirect effects of land use could still be possible (Fig. 1a), even though significant effects were not identified by the GLMM on the physiochemical data in inflow samples. The land use-associated bacterial taxa (Fig. 2a) and physiochemical variables (Supplementary Fig. 4, 9) were merged to construct a DAG network (Fig. 5), resulting in physiochemical variables situated at the top of the network, linked to each other. Of these, aluminium, colour value, and turbidity were central in the topology: the upper position in the network topology suggested that these variables influenced the ASVs below. Aluminium and colour value had several links with negative influence scores, while the overall network contained many positive links (Fig. 5, Supplementary Table 4). A limited number of ASVs occupied central parts of the network, mainly Proteobacteria comprising a large number of gram-negative bacterial genera. Of these, *Poly-nucleobacter*, *Herminiimonas*, *Rhodoferrax*, *Variovorax*, and *Limnhabitans* were central in the network. In the phylum Bacteroidota, *Sed-iminibacterium* had several nodes, while most Actinobacteria, such as *Planktoluna* and the hgcl clade, had fewer nodes and were found further

down the topology. To summarize, the DAG analysis provides an understanding of the effects of various factors on the microbiota; although a singular factor may not capture the complete effects of land use disturbance on water quality, indirect effects of e.g. changes in colour value may be revealed when whole systems are analysed, due to the complex interaction of variables such changes can cause ripple effects down in the network.

### 3.4. Occurrence of legionella

Of the bacterial genera identified as transmitting waterborne disease on the World Health Organization (2011) list, only the genus *Legionella* displayed high presence and diversity in the studied lake systems (584 ASVs: Fig. 6, 7, Supplementary Fig. 10), enabling a focused case study of this bacteria that has possible public health implications. *Legionella* had more ASVs than any of the 10 most abundant genera in the dataset (Supplementary Fig. 11), e.g. compared to 36 ASVs of *Mycobacterium*. The alpha diversity and relative abundance of *Legionella* were significantly higher in the watershed inflows ( $n = 115$ ) than the lakes ( $n = 35$ ) (Wilcoxon rank sum test  $W = 1440$ ,  $p$ -value = 0.0355) (Fig. 6d). To better understand the ecology of this diverse genus, we performed a taxonomic assessment and quantification as well as a deeper analysis to find links between the occurrence of *Legionella* and land use and/or environmental factors.

A general qPCR for *Legionella* spp. was positive in all four of the tested samples, selected to represent a mix of low and high abundance of the taxa. The highest concentration was found in the Ho inflow with ~4668 bacteria/L (Fig. 6a). The presence of ASVs classified as *L. pneumophila* were detected with full-length 16S analysis (Figs. 6b–c, Supplementary Table 3). However, results for the presence of the *mip* gene in the four samples were negative (data not shown), suggesting that the detected species does not belong to the virulent *L. pneumophila* taxa.

Three *Legionella* ASVs were present in more than eight samples in total and were included in GLMM analyses of anthropogenic factors (in inflow samples,  $n = 115$ ) and in samples with available physiochemical data ( $n = 63$ ), resulting in a strong inferred association with the presence of farms, pasture, and fields (Supplementary Fig. 12). In addition, the ASVs were positively associated with colour value ( $\beta_{\text{ASV5659}} = 2.90$ , 95 % CI [0.39, 5.89]) and aluminium content in the samples ( $\beta_{\text{ASV4077}} = 6.85$ , [1.68, 13.63]).

### 3.5. Prediction of Legionella in response to changed water chemistry

In future scenarios of climate change and anthropogenic activities, it is expected that human activities and their impacts on water quality will increase. Therefore, we performed simulations using the inferred model above: three *Legionella* ASVs presence were simulated in the case of a 50 % increase in colour values and aluminium levels, but also with the same levels as measured (i.e. the baseline). The average detection frequency of ASV4077 more than doubled across samples with a 50 % increase in aluminium at 0.193 (0.248) compared to the baseline at 0.092 (0.199) (Supplementary Fig. 13). A 50 % increase in colour value resulted in an average detection frequency of ASV5659 of 0.150 (0.193) compared to the background of 0.094 (0.158). The last ASV2280 increased in detection frequency as well, but not as pronounced: 0.093 (0.122) and 0.080 (0.111) for 50 % increase in aluminium and colour value respectively as compared to the background of 0.066 (0.101). It is thus likely that variants of *Legionella* will increase in presence with intensified land use.

## 4. Discussion

Upstream land use can have downstream consequences for the water chemistry and microbial community in inflows and lakes. Traditional approaches often compartmentalise factors influencing water quality, such as microorganisms, inflows, land use effects, and water chemistry,

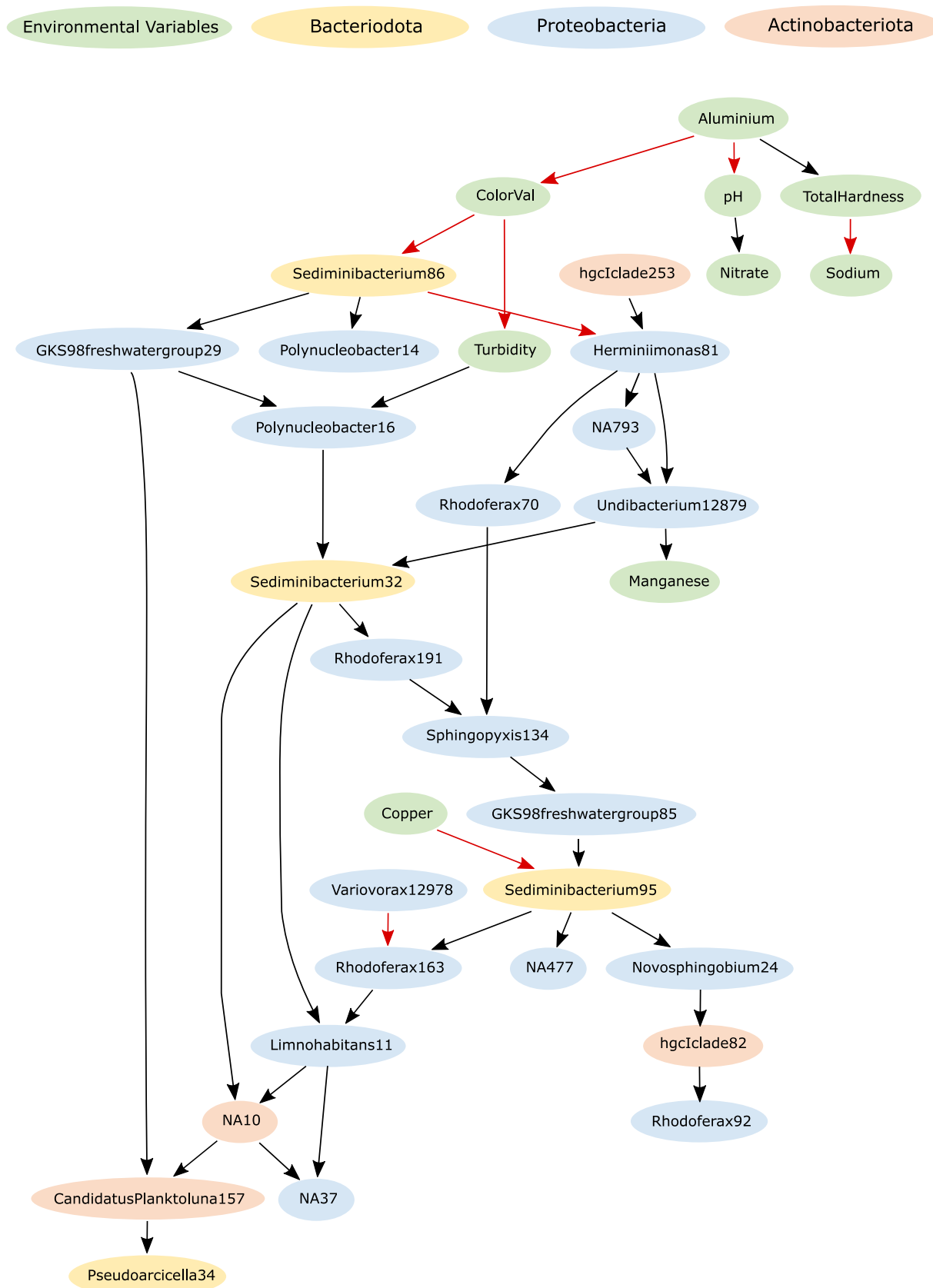
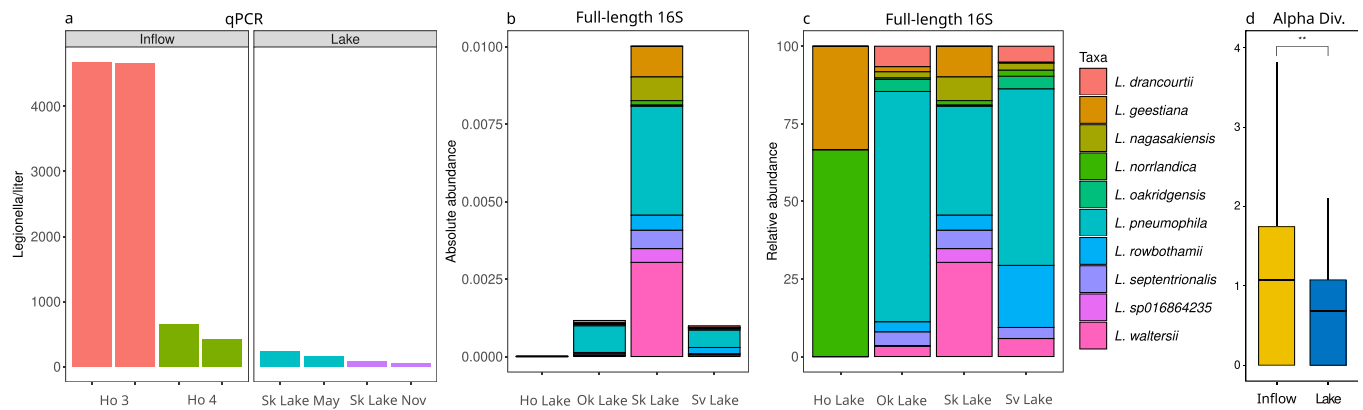
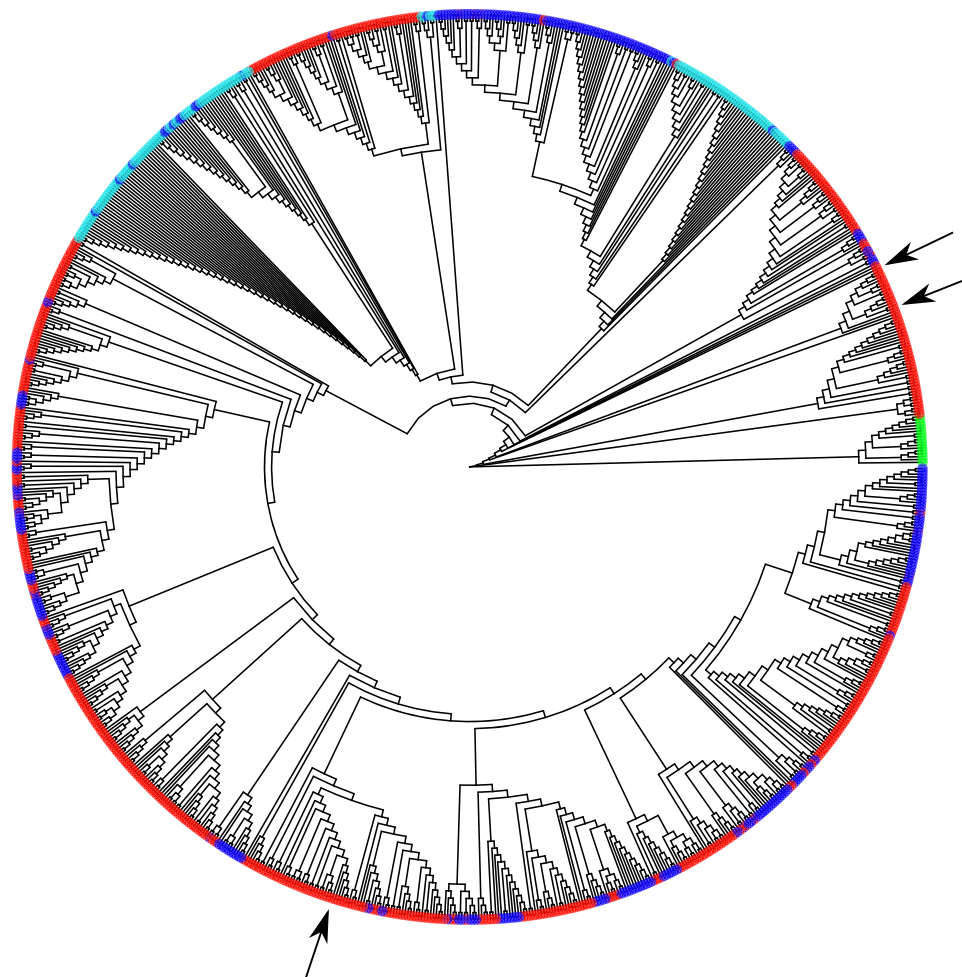


Fig. 5. Causal downstream effects from land use characteristics. Direct acyclic graph (DAG) analysis (using BANJO) of environmental data and V3-V4 16S rRNA ASVs (naming: GenusASVnumber). Colours denote environmental parameters or bacterial phyla. Red arrows denote negative links.



**Fig. 6.** *Legionella* abundances and species. Panel a Results of a real-time qPCR assay based on primers targeting a *Legionella*-specific part of the 16S rRNA gene. The resulting bacterial concentration was calculated based on cq values in the linear regression equation of a calibration curve specific to the assay:  $y(\log \text{ bacterial concentration}) = -0.2911 * x(\text{cq value}) + 10,361$ . Identical colours denote technical replicates. Panel b Abundances of the 10 most common full length 16S rDNA sequencing taxa identified as belonging to genus *Legionella* as proportions of total abundance. Panel c Relative proportions of the taxa shown in panel b. For b and c, taxa were classified by the Epi2Me pipeline using a custom database based on GTDB.



**Fig. 7.** Cladogram showing the diversity of *Legionella* taxa present in the dataset, with red dots names representing studied ASVs, green dots representing *Aquicella*/the outgroup, cyan dots representing *Legionella pneumophila* reference sequences, and blue sequence names representing other *Legionella* reference sequences. Arrows indicate the three most numerous taxa subjected to model analysis.

hindering the identification of intricate causal links. To address this limitation, our study comprehensively investigated five lakes, their inflows, and surrounding land use, emphasising that: I) land use influenced both bacteria and chemistry of lake inflows; II) upstream inflow impacted lake compositions; III) the lake microbiome can be indirectly

influenced by land use via e.g. changes in colour value and; IV) upstream land use can have downstream consequences relevant for water supply, illustrated by *Legionella* (a genus containing pathogens), that increased by the presence of pastures, fields, and farms.

Therefore, land use can either directly or indirectly threaten public

access to clean water, since many environmental and opportunistic bacterial pathogens could increase in occurrence. Threats to public access to clean water could include wide-scale alteration of the microbiota and the transmission of potential pathogens to surface raw water, sources providing water for consumption, agriculture and recreational lakes.

#### 4.1. Bacteria as sentinels of land use changes in lake inflows

We found that a significant portion (approximately one-third) of the inflow microbiota responded to land use, demonstrating the role of bacteria as sensitive indicators of environmental change. The observed shifts in bacterial communities, despite weak evidence for corresponding alterations in water chemistry, can be attributed to several key factors. First, bacteria exhibit remarkable sensitivity to changes in their surroundings, allowing them to serve as early indicators of ecological shifts (Faust et al., 2015). Additionally, the dynamic nature of microbial communities, influenced by factors such as functional redundancy and metabolic plasticity (Comte et al., 2013), may contribute to fluctuations in bacterial composition even when traditional chemical parameters remain relatively stable. Even though there was weak evidence for large alterations in water chemistry, it is important to consider the discrepancy in number of measurements between bacteria and chemistry (115 and 39, respectively). Since there were indications of land use effects on the water chemistry (Fig. 2b), we speculate that a more distinct impact could have emerged with a revised model design, including increased chemical measurements in the inflows. Still, the temporal dynamics of bacterial responses and their rapid adaptation to alterations in land use provide a more nuanced insight into environmental stressors (Hermans et al., 2020), offering a comprehensive perspective on ecosystem health that extends beyond conventional chemical assessments.

Conventional chemical assessments include measuring humic substances or the colour value of water, to understand the extent of organic matter and its potential influence on aquatic life. Humic substances leach from soils into waters (Susic, 2016), where they can be consumed by bacteria in an oxygen-demanding process (Farjalla et al., 2009). Runoff from fertilizers and manure can also increase the carbon content, contributing to brownification (Kritzberg et al., 2020). Iron and aluminium easily bind to humics (Zhou et al., 2015) and are transported with these substances (Herzog et al., 2020). Earlier studies have found that land use such as forest plantations can increase both the iron and humic substance contents in freshwater (Škerlep et al., 2022). In this study, the inflows often contained browner water compared to the lakes, but a connection to land use could not be statistically determined. However, changes in land use may introduce new organic substrates or alter nutrient availability, influencing the types of bacteria that thrive without necessarily altering overall water chemistry in terms of coloration of the water. Consequently, land use-induced alterations, such as changes in humic substances, can disrupt the microbial community, illustrating how disturbances initiated upstream can travel down the stream.

#### 4.2. Land use can indirectly affect the downstream bacterial community in lakes, via inflows

We found both that land use impacted approximately one third of the analyzed inflow microbial taxa, and that inflow associated taxa were detected in the lakes up to a total frequency of ~10–50 %. Land use could thus lead to indirect effects on a balanced microbiome due to the characteristics of the affected inflows (Fig. 4), which have an impact on the bacteria (Fig. 5). The bacterial compositions were more diverse in the inflows than in the lakes (Fig. 3), which could be explained by a more variable water chemistry in combination with the inflow of terrestrial taxa, as well as by land use effects. Earlier studies have shown that land use can affect bacterial communities (Kraemer et al., 2020). To understand how the upstream community can affect the downstream,

we performed a RF analysis. In this analysis, the communities in the inflows showed varying impacts on the lake communities, and on average about 20 % of the most abundant taxa were predicted to originate from inflows (Fig. 4a), an impact that we hypothesize could increase with increased precipitation. The same was found for the water chemistry, where e.g. colour value as well as iron and aluminium contents were generally higher in the inflows, impacting the lakes. These downstream propagations demonstrate the indirect ways land use can affect aquatic microbial ecosystems.

Indirect effects via inflows were also suggested by causal links between e.g. water colour and lake bacterial taxa, where this influenced the abundance of several freshwater bacteria, such as *Polynucleobacter* and *Limnohabitans*. In the network analysis, aluminium and colour value that were characteristics of the inflows had causal links with lake bacteria. In addition to involvement in phosphorous limitation, iron and aluminium can promote the formation of biofilms (Hu et al., 2019; Kappler et al., 2021). Changes in humic substances, iron, and aluminium can thus select for bacteria adapted to these conditions. Several bacteria in the DAG network (e.g., *Sediminibacterium*, *Herminiimonas*, *Rhodoferrax*, and *Polynucleobacter*) have been described as mainly providing ecosystem services, such as denitrification, phosphorous release, and organic matter degradation (Watanabe et al., 2012; Logue et al., 2016; Mandel et al., 2019). The hgcl clade and the CL500–29 marine clade were also identified as important for the inflow (Fig. 4c), and both these taxa have been shown to be involved in the carbon cycle and can use different forms of carbon-based compounds (Zhang et al., 2020). Although bacteria commonly break down organic matter, changes in the origin of this organic matter can lead to indirect effects on the community, since humic substances have varying nutrient contents and can be hard to degrade (Moran and Hodson, 1990). Thus, inflows contribute in shaping the microbial community in lakes, since their water properties can contribute in forming new niches, leading to the enrichment of other taxa, including potential pathogens.

#### 4.3. Land use can affect *Legionella* via iron, aluminium, and colour value

In the context of water supply significance, our investigation reveals a noteworthy relationship between upstream land use and *Legionella*. *Legionella* taxa were more abundant in watershed inflows than in lakes, in accordance with previous findings (Peabody et al., 2017), hinting at potential threats to water quality. For the first time, we report positive associations between *Legionella* occurrence and anthropogenic disturbances like farms, pasture, and fields. Our predictions revealed *Legionella* ASVs as positively linked to aluminium and colour value, suggesting opportunistic growth where other bacteria are limited (Fig. 5). *Legionella*'s propensity for biofilm living and intracellular lifestyles may contribute to its resilience in the face of e.g. nutrient limitations (Leon-Sicairos et al., 2015). Our findings align with previous suggestions of association between *Legionella* and humic substances (Eriksson et al., 2023), exhibiting a significant correlation with iron. The intricate interplay of iron, humic substances, and land use poses potential challenges to microbial communities, especially in humic-rich boreal lakes (Kritzberg et al., 2020), which according to our predictions can be favourable for *Legionella* occurrences. Our study highlights the need for future research, since the disentanglement of complex interplays require holistic approaches. This includes experiments introducing humic substances to environmental water, to understand the dynamics impacting this bacterial niche and its implications for water quality management.

## 5. Conclusion

To summarize:

- 1) 27 % of the most common bacterial taxa in the lake inflows were influenced by buildings, farms, and fields, increasing the bacterial diversity.

- II) The bacteria and water chemistry upstream can considerably affect the composition downstream in the lakes, where in some cases more than half of the features originated from the inflows.
- III) Land use can, through causal links, lead to indirect downstream effects on lake bacterial taxa, e.g. through changes in humic substances influencing common freshwater bacteria.
- IV) *Legionella* was particularly abundant in the lake inflows and increased in occurrence with the presence of pastures, fields, farms, aluminium, iron, and humic substances.

To ensure access to reliable clean water and suitable water treatment, it is therefore important to consider upstream land use and inflows even if no direct effects can be observed on water quality in terms of chemical measurements, since land use can influence the microbial communities in inflows where pathogens can naturally occur. There were no positive associations between land-use and faecal indicator bacteria, suggesting that these indicators are insufficient to detect land use effects. Considering climate change in Nordic lakes, with increasing humification, precipitation, flooding and freshwater inflow, bacteria utilizing these conditions, including potential pathogens, might become more common. It is our opinion that these findings are of interest to fundamental microbial ecology as well as for water management to identify robust water sources and to avoid enrichment of *Legionella*. Additional broader investigations of stressors on the microbial network and the ecology of this diverse genus would facilitate preventative measures, helping secure water accessibility.

## Funding

This project was financed by the Swedish Civil Contingencies Agency (project "Reservvatten - från bristfälligt resurs till resurs vid brist") and the marine research environment EcoChange (FORMAS).

## CRediT authorship contribution statement

**Jon Ahlinder:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Karolina Ida Anna Eriksson:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis. **Maria Hellmér:** Writing – review & editing, Resources, Investigation. **Emelie Salomonsson:** Writing – review & editing, Resources, Investigation. **Malin Granberg:** Writing – review & editing, Resources, Investigation. **Ingrid Dacklin:** Writing – review & editing, Resources, Investigation. **Josefine Elving:** Writing – review & editing, Supervision, Resources, Project administration, Investigation, Conceptualization. **Björn Brindefalk:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Formal analysis, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

Supplementary data related to this article can be accessed in the Supplementary Data1 file, and at the Short Read Archive (<https://www.ncbi.nlm.nih.gov/sra>), project accession: PRJNA, ID: 798233.

## Acknowledgments

We thank Stina Bäckman and Mattias Jonsson for the help with the qPCR, and Agneta Andersson, Mats Forsman, Johanna Thelaus, David

Sundell, Karin Jacobsson, and Rikard Dryselius for productive discussions and help with data collection. We acknowledge SNIC for facilitating the computational analysis.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at [doi:10.1016/j.watres.2024.121579](https://doi.org/10.1016/j.watres.2024.121579).

## References

- Anttila, J., Kaitala, V., Laakso, J., Ruokolainen, L., 2015. Environmental variation generates environmental opportunist pathogen outbreaks. *PLoS ONE* 10, e0145511. <https://doi.org/10.1371/journal.pone.0145511>.
- Brindefalk, B., Brolin, H., Sæve-Söderbergh, M., Karlsson, E., Sundell, D., Wikström, P., 2022. Bacterial composition in Swedish raw drinking water reveals three major interacting ubiquitous metacommunities. *Microbiologyopen* 11, e1320. <https://doi.org/10.1002/MBO3.1320>.
- Carpenter, B., Gelman, A., Hoffman, M.D., Lee, D., Goodrich, B., Betancourt, M., 2017. Stan : a Probabilistic Programming Language. *J. Stat. Softw.* 76 <https://doi.org/10.18637/jss.v076.i01>.
- Comte, J., Fauteux, L., del Giorgio, P.A., 2013. Links between metabolic plasticity and functional redundancy in freshwater bacterioplankton communities. *Front. Microbiol.* 4 <https://doi.org/10.3389/fmicb.2013.00112>.
- Comte, J., Berga, M., Severin, I., Logue, J.B., Lindström, E.S., 2017. Contribution of different bacterial dispersal sources to lakes: population and community effects in different seasons. *Environ. Microbiol.* 19, 2391–2404. <https://doi.org/10.1111/1462-2920.13749>.
- Eriksson, K.I.A., Ahlinder, J., Ramasamy, K.P., Andersson, A., Sundell, D., Karlsson, L., 2023. Association between *Legionella* species and humic substances during early summer in the northern Baltic Sea. *Front. Mar. Sci.* 9, 2835. <https://doi.org/10.3389/fmars.2022.1070341>.
- Eriksson, K.I.A., Thelaus, J., Andersson, A., Ahlinder, J., 2022. Microbial Interactions — Underexplored Links Between Public Health Relevant Bacteria and Protozoa in Coastal Environments. *Front. Microbiol.* 13, 2154. <https://doi.org/10.3389/fmicb.2022.877483>.
- Farjalla, V.F., Amado, A.M., Suhett, A.L., Meirelles-Pereira, F., 2009. DOC removal paradigms in highly humic aquatic ecosystems. *Environ. Sci. Pollut. Res.* 16, 531–538. <https://doi.org/10.1007/s11356-009-0165-x>.
- Hamilton, K.A., Prussin, A.J., Ahmed, W., Haas, C.N., 2018. Outbreaks of Legionnaires' Disease and Pontiac Fever 2006–2017. *Curr. Environ. Heal. Reports* 5, 263–271. <https://doi.org/10.1007/s40572-018-0201-4>.
- Herzog, S.D., Persson, P., Kvashnina, K., Kritzberg, E.S., 2020. Organic iron complexes enhance iron transport capacity along estuarine salinity gradients of Baltic estuaries. *Biogeosciences* 17, 331–344. <https://doi.org/10.5194/bg-17-331-2020>.
- Holsinger, H., Tucker, N., Regli, S., Studer, K., Roberts, V.A., Collier, S., 2022. Characterization of reported legionellosis outbreaks associated with buildings served by public drinking water systems: United States, 2001–2017. *J. Water Health* 20, 702–711. <https://doi.org/10.2166/WH.2022.002>.
- Hu, X., Yang, L., Lai, X., Yao, Q., Chen, K., 2019. Influence of Al(III) on biofilm and its extracellular polymeric substances in sequencing batch biofilm reactors. *Environ. Technol.* 40, 53–59. <https://doi.org/10.1080/09593330.2017.1378268>.
- Kappler, A., Bryce, C., Mansor, M., Lueder, U., Byrne, J.M., Swanner, E.D., 2021. An evolving view on biogeochemical cycling of iron. *Nat. Rev. Microbiol.* 19, 360–374. <https://doi.org/10.1038/s41579-020-00502-7>.
- Kraemer, S.A., Barbosa da Costa, N., Shapiro, B.J., Fradette, M., Huot, Y., Walsh, D.A., 2020. A large-scale assessment of lakes reveals a pervasive signal of land use on bacterial communities. *ISME J.* 14, 3011–3023. <https://doi.org/10.1038/s41396-020-0733-0>.
- Kritzberg, E.S., Hasselquist, E.M., Škerlep, M., Löfgren, S., Olsson, O., Stadmark, J., 2020. Browning of freshwaters: consequences to ecosystem services, underlying drivers, and potential mitigation measures. *Ambio* 49, 375–390. <https://doi.org/10.1007/s13280-019-01227-5>.
- Leon-Sicairos, N., Reyes-Cortes, R., Guadrón-Llanos, A.M., Madueña-Molina, J., Leon-Sicairos, C., Canizalez-Román, A., 2015. Strategies of Intracellular Pathogens for Obtaining Iron from the Environment. *Biomed Res. Int.* 2015, 1–17.
- Leoni, E., Catalani, F., Marini, S., Dallolio, L., 2018. Legionellosis Associated with Recreational Waters: a Systematic Review of Cases and Outbreaks in Swimming Pools, Spa Pools, and Similar Environments. *Int. J. Environ. Res. Public Heal.* 15, 1612. 2018, Vol. 15, Page 1612.
- Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. *R news* 2, 18–22. Available at: <http://www.stat.berkeley.edu/> [Accessed August 23, 2021].
- Logue, J.B., Stedmon, C.A., Kellerman, A.M., Nielsen, N.J., Andersson, A.F., Laudon, H., 2016. Experimental insights into the importance of aquatic bacterial community composition to the degradation of dissolved organic matter. *ISME J.* 10, 533–545.
- Mandel, A., Zekker, I., Jaagura, M., Tenno, T., 2019. Enhancement of anoxic phosphorus uptake of denitrifying phosphorus removal process by biomass adaptation. *Int. J. Environ. Sci. Technol.* 16, 5965–5978. <https://doi.org/10.1007/s13762-018-02194-2>.
- Moran, M.A., Hodson, R.E., 1990. Bacterial production on humic and nonhumic components of dissolved organic carbon. *Limnol. Oceanogr.* 35, 1744–1756.

- Ng, V., Tang, P., Jamieson, F., Drews, S.J., Brown, S., Low, D.E., 2008. Going with the flow: legionellosis risk in Toronto, Canada is strongly associated with local watershed hydrology. *Ecohealth* 5, 482–490.
- Nguyen, L.T., Schmidt, H.A., Von Haeseler, A., Minh, B.Q., 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32 (1), 268–274.
- Peabody, M.A., Caravas, J.A., Morrison, S.S., Mercante, J.W., Prystajecy, N.A., Raphael, B.H., 2017. Characterization of *Legionella* Species from Watersheds in British Columbia, Canada. *mSphere* 2. <https://doi.org/10.1128/mSphere.00246-17>.
- Public Health Agency of Sweden (2015). Outbreaks of disease caused by drinking water in Sweden 1992–2011. Available at: <https://www.folkhalsomyndigheten.se/publikationer-och-material/publikationsarkiv/s/sjukdomsutbrott-orsakade-av-dricksvattnen/> [Accessed December 15, 2022].
- Schelker, J., Sponseller, R., Ring, E., Högbom, L., Löfgren, S., Laudon, H., 2016. Nitrogen export from a boreal stream network following forest harvesting: seasonal nitrate removal and conservative export of organic forms. *Biogeosciences* 13, 1–12. <https://doi.org/10.5194/bg-13-1-2016>.
- Škerlep, M., Nehzati, S., Johansson, U., Kleja, D.B., Persson, P., Kritzberg, E.S., 2022. Spruce forest afforestation leading to increased Fe mobilization from soils. *Biogeochemistry* 157, 273–290. <https://doi.org/10.1007/s10533-021-00874-9>.
- Soranno, P.A., Cheruvilil, K.S., Wagner, T., Webster, K.E., Bremigan, M.T., 2015. Effects of Land Use on Lake Nutrients: the Importance of Scale, Hydrologic Connectivity, and Region. *PLoS ONE* 10, e0135454. <https://doi.org/10.1371/journal.pone.0135454>.
- Susic, M., 2016. Replenishing Humic Acids in Agricultural Soils. *Agronomy* 6, 45. <https://doi.org/10.3390/agronomy6040045>.
- UN General Assembly (2015). Transforming our world : the 2030 agenda for sustainable development. 21 Oct. 2015, A/RES/70/1. Available at: <https://www.refworld.org/docid/57b6e3e44.html> [Accessed February 7, 2023].
- Vörösmarty, C.J., Green, P., Salisbury, J., Lammers, R.B., 2000. Global water resources: vulnerability from climate change and population growth. *Science* (80) 289, 284–288. <https://doi.org/10.1126/SCIENCE.289.5477.284/ASSET/8277DF2D-D298-4C85-88F2-B3A4F9CE0C3B/ASSETS/GRAPHIC/SE2508647003.JPEG>.
- Watanabe, K., Komatsu, N., Kitamura, T., Ishii, Y., Park, H.D., Miyata, R., 2012. Ecological niche separation in the *Polynucleobacter* subclusters linked to quality of dissolved organic matter: a demonstration using a high sensitivity cultivation-based approach. *Environ. Microbiol.* 14, 2511–2525. <https://doi.org/10.1111/j.1462-2920.2012.02815.x>.
- Wickham, H., 2016. *Ggplot2*. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-24277-4>.
- World Economic Forum (2023). *The Global Risks Report 2023 18th Edition*. Available at: [www.weforum.org](http://www.weforum.org) [Accessed February 7, 2023].
- World Health Organization (2011). *Guidelines for drinking-water quality*. 4th ed. Available at: [https://apps.who.int/iris/bitstream/handle/10665/44584/9789241548151\\_eng.pdf](https://apps.who.int/iris/bitstream/handle/10665/44584/9789241548151_eng.pdf).
- Wu, K., Hu, M., Zhang, Y., Zhou, J., Wu, H., Wang, M., 2022. Long-term riverine nitrogen dynamics reveal the efficacy of water pollution control strategies. *J. Hydrol.* 607, 127582 <https://doi.org/10.1016/j.jhydrol.2022.127582>.
- Yang, T., Kim, H.J., 2019. Nutrient management regime affects water quality, crop growth, and nitrogen use efficiency of aquaponic systems. *Sci. Hortic. (Amsterdam)*. 256, 108619 <https://doi.org/10.1016/j.scienta.2019.108619>.
- Zhang, L., Zhong, M., Li, X., Lu, W., Li, J., 2020. River bacterial community structure and co-occurrence patterns under the influence of different domestic sewage types. *J. Environ. Manage.* 266, 110590 <https://doi.org/10.1016/j.jenvman.2020.110590>.
- Zhou, S., Chen, S., Yuan, Y., Lu, Q., 2015. Influence of humic acid complexation with metal ions on extracellular electron transfer activity. *Sci. Rep.* 5, 17067. <https://doi.org/10.1038/srep17067>.