



HAL
open science

Microbial risk assessment of *Nocardia cyriacigeorgica* in polluted environments, case of urban rainfall water

Florian Vautrin, Petar Pujic, Christian Paquet, Emmanuelle Bergeron, Delphine Mouni e, Thierry Marchal, H el ene Salord, Jeanne-Marie Bonnet-Garin, Benoit Cournoyer, Thierry Winiarski, et al.

► To cite this version:

Florian Vautrin, Petar Pujic, Christian Paquet, Emmanuelle Bergeron, Delphine Mouni e, et al.. Microbial risk assessment of *Nocardia cyriacigeorgica* in polluted environments, case of urban rainfall water. Computational and Structural Biotechnology Journal, 2021, 19, pp.384-400. 10.1016/j.csbj.2020.12.017 . hal-03157072

HAL Id: hal-03157072

<https://hal.science/hal-03157072>

Submitted on 3 Feb 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin e au d p t et   la diffusion de documents scientifiques de niveau recherche, publi s ou non,  manant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv s.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 Microbial risk assessment of *Nocardia* 2 *cyriacigeorgica* in polluted environments, 3 case of urban rainfall water

4 Florian Vautrin^{a,d,e}, Petar Pujic^a, Christian Paquet^b, Emmanuelle Bergeron^{a,d}, Delphine
5 Mouni e^a, Thierry Marchal^c, H el ene Salord^d, Jeanne-Marie Bonnet^b, Benoit Cournoyer^a,
6 Thierry Winiarski^e, Vanessa Louzier^{b*}, Veronica Rodriguez-Nava^{a,d*}

7
8 ^a UMR Ecologie Microbienne, CNRS 5557, INRA 1418, Research team on « bacterial
9 opportunistic pathogens and environment », VetAgro Sup and University Lyon 1, F-69363
10 Lyon, France; ^b APCSe, Agressions Pulmonaires et Cardiovasculaires dans le Sepsis,
11 VetAgro Sup – Campus v et erinaire de Lyon, 69280 Marcy l’Etoile, France; ^c Laboratoire
12 v et erinaire d’histopathologie, VetAgro Sup – Campus v et erinaire de Lyon, 69280 Marcy
13 l’Etoile, France; ^d Observatoire Fran ais des Nocardioses, Institut des Agents Infectieux (IAI),
14 Centre de Biologie et Pathologie Nord, H opital de la Croix-Rousse, 69317 Lyon, France; ^e
15 UMR LEHNA, CNRS 5023, ENTPE, University Lyon 1, F-69622 Villeurbanne, France. *

16 These authors contributed equally to this work.

17 **e-mail addresses of the authors:**

18 florianvautrin@gmail.com; petar.pujic@univ-lyon1.fr; christian.paquet@vetagro-sup.fr;
19 emmanuelle.bergeron@univ-lyon1.fr; delphine.mouni e@univ-lyon1.fr;
20 thierry.marchal@vetagro-sup.fr; helene.salord@chu-lyon.fr; jeanne-marie.bonnet@vetagro-
21 sup.fr; benoit.cournoyer@vetagro-sup.fr; thierry.winiarski@entpe.fr;
22 vanessa.louzier@vetagro-sup.fr; veronica.rodriguez-nava@univ-lyon1.fr

23 **Address correspondence to:** Veronica Rodriguez-Nava, UMR Ecologie Microbienne, CNRS
24 5557, INRA 1418, 8 avenue Rockefeller, 69373 Lyon cedex, France. E-mail:

25 veronica.rodriguez-nava@univ-lyon1.fr_ (+33)478777276 **and** Vanessa Louzier, APCSe,
26 Agressions Pulmonaires et Cardiovasculaires dans le Sepsis, VetAgro Sup – Campus
27 vétérinaire de Lyon, 69280 Marcy l’Etoile, France. E-mail: vanessa.louzier@vetagro-sup.fr
28 (+33)478872760

29 **Running title:** Environmental pollution and assessment of microbial risk

30 **Declarations of interest:** none

31 **Author contributions:**

32 **Florian Vautrin:** writing, data analysis, methodology, investigation. **Petar Pujic:**
33 investigation, data analysis. **Christian Paquet:** investigation. **Emmanuelle Bergeron:**
34 investigation, **Delphine Mouniée:** investigation, **Thierry Marchal:** investigation. **Hélène**
35 **Salord:** investigation. **Jeanne-Marie Bonnet:** validation. **Benoit Cournoyer:** data analysis,
36 validation, reviewing. **Thierry Winiarski:** methodology, reviewing. **Vanessa Louzier:**
37 supervision, data analysis, reviewing. **Veronica Rodriguez-Nava:** supervision, data analysis,
38 writing-reviewing.

39

40

41 **Highlights**

42 Urban infiltration basins are a reservoir of a high diversity of *Nocardia* encompassing both
43 pathogenic and not-pathogenic species.

44 Relative abundance of pathogenic *Nocardia* species presents a positive correlation with metal
45 trace elements.

46 High infraspecific variability within *N. cyriacigeorgica*, forming three phylogroups.

47 Immunosuppressive window may be enough to infect immunocompetent mice.

48 To better understand the ecology of *Nocardia* to elucidate sources of clinical infections.

49 First isolation of *N. cyriacigeorgica* in Europe in a polluted urban environment.

50 Environmental *N. cyriacigeorgica* strains may be as virulent as clinical GUH-2 strain.

51 *hsp65* marker can be used by metabarcoding approach for assessment of environmental
52 *Nocardia* biodiversity.

53

54 **Abstract**

55 Urban Infiltration Basins (UIBs) are used to manage urban runoff transfers and feed aquifers.

56 These UIBs can accumulate urban pollutants and favor the growth of potentially pathogenic
57 biological agents as *Nocardia*.

58 **Objectives:** To assess the spatio-temporal dynamics of pathogenic *Nocardia* in UIBs and to
59 establish phylogenetic relationships between clinical and UIB *N. cyriacigeorgica* strains. To

60 assess pathogenicity associated with environmental *N. cyriacigeorgica* using an animal
61 model, and to identify genetic elements that may be associated to its virulence.

62 **Methods:** A well-characterized UIB in terms of chemical pollutants from Lyon area was used
63 in this study during a whole year. Cultural and Next-Generation-Sequencing methods were
64 used for *Nocardia* detection and typing. Clinical and environmental isolates phylogenetic
65 relationships and virulences were compared with Multilocus-Sequence-Analysis study
66 together with a murine model.

67 **Results:** In autumn, *N. cyriacigeorgica* and *N. nova* were the pathogenic most prevalent
68 species in the UIB. The complex *N. abscessus/asiatica* was also detected together with some
69 other non-pathogenic species. The presence of pathogenic *Nocardia* was positively correlated
70 to metallic trace elements. Up to 1.0×10^3 CFU/g sediment of *N. cyriacigeorgica* and 6 OTUs
71 splitted in two different phylogroups were retrieved and were close to clinical strains. The
72 EML446 tested UIB isolate showed significant infectivity in mice with pulmonary damages
73 similar to clinical clone (GUH-2).

74 **Conclusion:** Hsp65 marker-based metabarcoding approach allowed detecting *N.*
75 *cyriacigeorgica* as the most abundant *Nocardia* pathogenic species in a UIB. Metal trace
76 elements-polluted environments can be reservoirs of pathogenic *Nocardia* which may have a
77 similar virulence to clinical strains.

78 **Keywords:** *Nocardia*; Opportunistic pathogen; Environment; Murine model of transient
79 immunoparalysis; Urban pollution; *hsp65* metabarcoding

80

81 **1. Introduction**

82 *Nocardia* are Gram-positive facultative intracellular bacteria responsible for
83 nocardiosis. Manifestations of disease range from cutaneous infection caused by traumatic
84 inoculation of the organism in a normal host to severe pulmonary or central nervous system
85 (CNS) disease in an immunocompromised host (1). Pulmonary infection is the most frequent
86 clinical form which is similar to pneumonia in 80% of cases and that can be fatal in patients
87 who are immunocompromised or affected by chronic pulmonary diseases (2, 3). *Nocardia*
88 cells are ubiquitous in the environment but distribution biases per species are still poorly
89 documented. Nocardiosis are caused by inhalation of these bacteria from aerosolized soils.
90 *Nocardia* cells are metabolically versatile, and some species such as *N. cyriacigeorgica* can
91 harbor genes involved in the degradation of petrol-derivatives (4–6). This property likely
92 explains the tropism of *Nocardia* cells for polluted environments (7). *N. cyriacigeorgica* is
93 also known for its ability to propagate in alveolar macrophages, inducing pulmonary damage.
94 Some more virulent strains are able to disseminate and reach the brain (8). Prevalence of *N.*
95 *cyriacigeorgica* in nocardiosis was estimated to be around 20% in the USA (9), 25% in Spain
96 (10) and 13% in France with a global nocardiosis incidence between 0.33 and 0.87 / 100 000
97 inhabitants (11). The environmental occurrence, persistence and enrichment of *N.*
98 *cyriacigeorgica* remain to be defined. Here, the hypothesis of a tropism of *Nocardia* cells, and
99 *N. cyriacigeorgica*, for biotopes found in a city, was tested because of their ability at using
100 petrol-derivatives as carbon sources. Furthermore, we have tested the hypothesis of an on-
101 going evolution in the virulence traits of these urban *Nocardia* cells. The hypothesis was that
102 city strains should have a reduced but significant virulence in comparison with clinical
103 isolates, when tested on mice as an alternative host system.

104

105 *Nocardia* cells together with other microorganisms such as *Pseudomonas aeruginosa*,
106 *Aeromonas caviae* and fecal indicators, have recently been shown to be recurrent
107 contaminants of the urban deposits of a detention basin of a Stormwater Infiltration System
108 (SIS) located in the Lyon area (France). Presence of *Nocardia* cells at the outflow of this
109 basin redirecting to an Urban Infiltration Basin (UIB) were also observed, but the transferred
110 species from the detention to the UIB remained to be defined (12). To decrease the
111 environmental impacts of runoff flooding, SISs have been constructed all over the world to
112 manage runoff transfers and favor the recharging of local aquifers. Today, more than 5000
113 SISs are monitored around the world (13). Runoff waters getting into SISs are loaded with
114 organic and mineral particles such as PAHs (Polycyclic Aromatic Hydrocarbons), PCBs
115 (Polychlorinated Biphenyls), heavy metals and microorganisms (14), which accumulate on
116 the surface of the infiltration basins to generate the so-called “urban sediments” (15).

117 The presence of pathogenic microorganisms in these sediments constitutes a public
118 health risk because they can contribute in multiple ways at disseminating hazardous biological
119 agents either through (i) a transfer into natural water systems such as an aquifer which can be
120 used for gardening (16), (ii) a contamination of animals feeding in these systems that can
121 come into contact with humans (dogs, cats, rats, birds), or (iii) through an aerosolization
122 towards environments with environmental characteristics (moisture, pollutants, etc.) that may
123 also favor their growth (moist urban zones, gas-stations, major road axes, petrochemical
124 factories, etc.) and the consequent inhalation of these microbial cells by the local populations.
125 It has to be noted that aerosolized bacterial cells can migrate over large distances as observed
126 for *P. aeruginosa*, *Escherichia coli*, and *Klebsiella pneumoniae* (17).

127 Moreover, the interest of studying UIBs is that, due to their heterogeneity (gradient of
128 pollutants and moisture), they can mimic various environments that can be found in many

129 other different locations, such as urban water areas (blue-zones), green areas or puddles after
130 a rainfall event.

131 The aims of this study were thus to determine the spatiotemporal distribution biases of
132 *Nocardia* cells and pathogenic species such as *N. cyriacigeorgica* in an urban SIS, and to
133 evaluate their hazards for local populations. Epidemiological molecular investigations were
134 performed to define the phylogenetic relations between SIS *N. cyriacigeorgica* isolates and
135 clinical strains. The virulence of these isolates was then compared using a double-hit murine
136 model of transient immunoparalysis, but also through an analysis of virulence gene content or
137 virulome.

138 **2. Materials and Methods**

139 *2.1. Stormwater Infiltration System*

140 The studied SIS (named Django-Reinhardt) is part of a long term monitoring site of
141 OTHU (Field Observatory for Urban Water Management; <http://www.graie.org/othu/>) (18). It
142 is located in Chassieu, France (eastern part of Lyon). This system has been operational for
143 approximately 40 years and consists of a detention basin receiving runoff water from the
144 stormwater network, and discharging its waters into an Urban Infiltration Basin (UIB),
145 hereafter named DRIB (Django-Reinhardt Infiltration Basin) (Figure 1). The DRIB has a 1 ha
146 surface and a volume of 61,000 m³. The drained surfaces are in a stabilized industrial area,
147 and the main pollutants found in the accumulated sediments are heavy metals, cyanides, inks,
148 fats, hydrocarbons and solvents (19).

149 The DRIB has been extensively studied. Some of its sediments properties were
150 characterized according to normalized procedures such as ISO10390 for pH and ISO13320
151 for granulometry. Soil moisture was determined by comparative weighing before and after 24

152 h at 105°C. Additional DRIB data were also extracted from the Gessol report, and indicated a
153 mean of 2-3.5 mg for the sum of the 16 well-defined PAHs /kg dry sediment (19). These
154 values were found equivalent to those of industrial soils, *e.g.* 3.5 mg/kg according to Muntean
155 *et al.* (20), while PAHs levels in unpolluted soils were 4-12 µg/kg (21).

156 2.2. Sampling

157 Urban sediments of the DRIB were sampled during three different sampling
158 campaigns in, one per season (autumn (November), spring (April) and summer (July) 2015-
159 2016) in three contrasted areas: near the detention basin discharging pipe (inflow zone), in
160 the middle of the basin (bottom zone), and at the southern end of the basin (upper zone, five
161 samples per area). These positions match different concentration of pollutants, distinct
162 hydrological behaviors, soil moisture and vegetation (Figure 1).

163 **Table 1.** PCR primers and DNA amplification conditions for *rrs*, *hsp65*, *sodA*, *secA1* genes and *Nocardia* specific PCR.

164

Target	Length (bp)	Forward primer (5'-3')		Reverse primer (5'-3')		PCR cycling conditions	Reference
		Primer name	Sequence	Primer name	Sequence		
<i>rrs</i>	569	Noc1	GCTTAACACATGCAAG TCG	Noc2	GAATTCCAGTCTCCCC TG	5 min 94°C; 40x 1 min 94°C, 1 min 58°C, 1 min 72°C; 10 min 72°C	Rodriguez-Nava et al. 2006 (22)
<i>hsp65</i>	401	TB11	ACCAACGATGGTGTGT CCAT	TB12	CTTGTCGAACCGCATA CCCT	5 min 94°C; 35x 1 min 94°C, 1 min 55°C, 1 min 72°C; 10 min 72°C	Telenti et al. 1993 (23)
<i>sodA</i>	406	SodV1	CACCAYWSCAAGCAC CA	SodV2	CCTTGACGTTCTGGTA CTG	5 min 94°C; 35x 1 min 94°C, 1 min 52°C, 1 min 72°C; 10 min 72°C	Sánchez-Herrera et al. 2017 (24)
<i>secA1</i>	469	SecA1	GTAAAACGACGGCCA GGACAGYGAGTGGAT GGGYCGSGTGCACCG	SecA2	CAGGAAACAGCTATG ACGCGGACGATGTAG TCCTTGTC	5 min 95°C; 35x 1 min 95°C, 1 min 60°C, 1 min 72°C; 10 min 72°C	Conville et al. 2006 (25)
NG*	590	NG1	ACCGACAAGGGGG	NG2	GGTTGTAAACCTCTTT CGA	11 min 94°C; 30x 1 min 94°C, 20 sec 55°C, 1 min 72°C; 10 min 72°C	Laurent et al. 1999 (26)

165

166 Note: Y=C or T, W=A or T and S=C or G, * NG = *Nocardia* Genus

167

168 2.3. *Physical-chemical parameters statistical analyses*

169 All datasets were analyzed with the R software (V.3.1.3) (27). The distribution of the
170 physical-chemical parameters (PAHs, trace metal elements (Cd, Cu, Hg, Pb and Zn),
171 granulometry, water content) and relative abundance of *Nocardia* (pathogenic vs non-
172 pathogenic) was represented by a between-class analysis (BCA) allowing a longitudinal
173 analysis. Packages ade4 (28), mixOmics and RVAideMemoire (29) were used. Trace
174 elements and metals were log transformed because they are not normally distributed. All the
175 other parameters were normally distributed according to a Shapiro test. The diversity within
176 each individual sample was estimated using the non-parametric Shannon index. Statistical
177 analyses were performed using ANOVA2 and normality of the residues was tested in order to
178 establish the significance of the groupings. Only p-values lower to 0.05 were considered as
179 statistically significant. The correlogram and p-values were obtained on R software using the
180 ade4, corrplot and Hmisc packages (30).

181 2.4 *Identification approaches*

182 2.4.1. *Culture-independent approach*

183 A metabarcoding approach by using *hsp65* gene was used. Genomic DNA from
184 environmental samples of the DRIB was extracted using the FastDNA SPIN Kit for Soil (MP
185 Biomedicals, France) according to the manufacturer's instructions. Amplifications were
186 performed on the *hsp65* gene (Table 1) using PuReTaqTM Ready-To-Go PCR Beads (GE
187 Healthcare) in a final volume of 50 μ L and sequenced (together with purification and quality
188 control) by Biofidal (<https://www.biofidal.com/>) using high-throughput Illumina MiSeq with
189 2x250 bp, paired-end chemistry to obtain 20,000 paired reads per sample. Bioinformatics
190 analysis were performed using the MOTHUR pipeline (31), and according to the protocol

191 previously defined by Marti et al., (16). The number of sequences was normalized between
192 the samples and was set to 7,128 sequences. *Hsp65*-OTUs at a 99% cutoff for species level
193 identification were used.

194

195 2.4.2. Culture-dependent approach

196 Diluted suspensions of the urban sediments from the DRIB were cultured on Bennett
197 and Middlebrook 7H10 agar (Thermo Fisher ®) semi-selective medium, and colonies with
198 morphological features typical of *Nocardia* (presenting a white and powdery aspect and
199 embedded in the agar) were purified, and then identified at the species level by sequencing
200 and analysis of the 16S rRNA gene, according to Rodriguez-Nava *et al.* (32) and following
201 the CLSI guidelines of similarity percentages greater than or equal to 99.6% (27). When
202 facing ambiguity or when identification is not possible with this gene we decided to use a
203 more discriminative one according to Sanchez-Herrera *et al.* (24).
204 *N. cyriacigeorgica* isolates were obtained from the 2015-2016 sampling campaign. An
205 additional *N. cyriacigeorgica* isolate (EML446), obtained from a bottom area sediment
206 sample in a previous campaign in 2013, was added to our work. It is the first strain of the
207 pathogenic *N. cyriacigeorgica* species that has been isolated in an anthropized environment in
208 Europe. Sampling conditions for this last isolate were identical, just the culture media
209 changed as the Actinomytes isolation Agar (Difco™) culture media was used in that case.
210 The list of isolates is presented in Table 2.

211
212

Table 2. List of SIS and clinical strains used in this study, and their origin, main features and MLSA-phylogroups.

Patient record	Sample date	Sample	Host Immunosuppressed	Tropism of nocardiosis or origin of isolate	Location	MLSA-Phylogroup
OFN.4	03/2015	Pus from cutaneous abscess	Yes	Cutaneous	Neuilly sur Seine	PI
OFN.5	10/2017	Bronchial aspirate	Yes	Lung	Lyon	PI
OFN.6	06/2010	Pus from cerebral abscess	Yes	Brain	Bron	PI
OFN.7	06/2015	Bronchial aspirate	No	Brain ^d	Bron	PI
OFN.13	02/2015	Pus from skin abscess	Yes	Lung ^d	Montpellier	PI
OFN.14	03/2015	Cervical biopsy	Yes	Brain	Montpellier	PI
DjRm.12	11/2015	UIB <i>hsp65</i> metabarcoding	-	Bottom	Chassieu	PI
DjRm.14	11/2015	UIB <i>hsp65</i> metabarcoding	-	Upper	Chassieu	PI
OFN.8	02/2016	Cervical biopsy	NA	Brain	Metz	PII
OFN.9	11/2015	Bronchial aspirate	Yes	Lung	Montpellier	PII
OFN.10	11/2015	Bronchial aspirate	Yes	Lung	Aulney sous Bois	PII
OFN.11	07/2011	Pus from cutaneous abscess	No	Cutaneous	Belley	PII
OFN.12	04/2014	Blood culture	Yes	Brain ^d	Réunion island	PII
OFN.1	02/2013	Blood culture	No	Lung ^d	La Roche sur Yon	PIII
OFN.2	03/2016	Pleural puncture	Yes	Lung	Agen	PIII
OFN.3	01/2016	Lung biopsy	Yes	Lung	Tours	PIII
EML446	04/2013	UIB polluted sediments	-	Bottom	Chassieu	PIII
EML1456	11/2015	UIB polluted sediments	-	Bottom	Chassieu	PIII
DjR.1	11/2015	UIB polluted sediments	-	Bottom	Chassieu	PIII
DjR.2	11/2015	UIB polluted sediments	-	Bottom	Chassieu	PIII
DjR.3	07/2016	UIB polluted sediments	-	Upper	Chassieu	PIII
DjR.4	11/2015	UIB polluted sediments	-	Bottom	Chassieu	PIII
DjR.5	07/2016	UIB polluted sediments	-	Upper	Chassieu	PIII
DjR.6	11/2015	UIB polluted sediments	-	Bottom	Chassieu	PIII
DjR.7	07/2016	UIB polluted sediments	-	Upper	Chassieu	PIII
DjR.8	07/2016	UIB polluted sediments	-	Bottom	Chassieu	PIII
DjR.9	07/2016	UIB polluted sediments	-	Bottom	Chassieu	PIII
DjR.10	07/2016	UIB polluted sediments	-	Bottom	Chassieu	PIII

DjR.11	04/2016	UIB polluted sediments	-	Upper	Chassieu	PIII
DjRm.13	07.2016	UIB <i>hsp65</i> metabarcoding	-	Bottom	Chassieu	PIII

213

214 Note: ^d indicates a disseminated case of nocardiosis, NA: data not available, MLSA: multilocus sequence analysis, SIS: stormwater infiltration
215 system

216

217 2.5. Phylogenetic analysis

218 DNA sequences (*rrs-hsp65-sodA-secA1*) were generated for the DRIB *N.*
219 *cyriacigeorgica* isolates (n=13). These sequences were compared with those of clinical *N.*
220 *cyriacigeorgica* (n=14), and of other *Nocardia* reference strains (n=10). The following strains
221 were considered: *N. abscessus* DSM 44432^T, *N. anaemiae* DSM 44821^T, *N. asteroides* ATCC
222 19247^T, *N. brasiliensis* ATCC 19296^T, *N. cyriacigeorgica* DSM 44484^T, *N. farcinica* IFM
223 10152^T, *N. nova* DSM 44481^T, *N. otitidiscaviarum* ATCC 14629^T, *N. vinacea* JCM 10988^T
224 and *N. cyriacigeorgica* GUH-2, the reference pathogenic strain isolated from a fatal case of
225 nocardiosis after a renal transplant in the 1970s (33). These reference strains chosen in this
226 study are phylogenetically closely related to *N. cyriacigeorgica* according to Yassin *et al.*
227 (34). Clinical *N. cyriacigeorgica* were obtained from OFN (French Observatory of
228 Nocardiosis, <http://ofn.univ-lyon1.fr/>), and originated from French patients affected by
229 nocardiosis (cutaneous, pulmonary, and cerebral infections) (Table 2). By selecting these
230 strains, we obtained a good representation of the main clinical forms of this disease that
231 encompassed the environmental sampling period of this study (2015-2016).

232 Bacterial DNAs of these strains were extracted by the boiling method using
233 achromopeptidase (10 U.μL⁻¹, Sigma-Aldrich). Amplifications of the following genes were
234 performed: 16S rRNA (*rrs*) (positions 64 to 663, *N. cyriacigeorgica* DSM 44484^T), *hsp65*
235 (positions 398 to 836, *N. cyriacigeorgica* DSM 44484^T), *sodA* (positions 99 to 505, *N.*
236 *cyriacigeorgica* DSM 44484^T), *secA1* (positions 396 to 864, *N. cyriacigeorgica* DSM
237 44484^T). PCR were performed using PuReTaqTM Ready-To-Go PCR Beads (GE Healthcare)
238 in a final volume of 25 μL with 200 ng of DNA. Primers and PCR conditions are listed in
239 Table 1. PCR products were sequenced by Biofidal (Vaulx-en-Velin, France). Multiple
240 alignments were generated by ClustalW using Seaview version 4.4.2 (35). Only for the

241 phylogenetic analysis of *N. cyriacigeorgica* species based on *hsp65*-gene, we added to our
242 sequences dataset i) some sequences from the historically known *hsp65*-based genotypes (9,
243 36–38) that are available on Genbank, and ii) some sequences coming from our
244 metabarcoding analysis of strains isolated from the DRIB. For MLSA (Multilocus Sequence
245 Analysis), the *rrs-hsp65-sodA-secA1* sequences were concatenated (1845 bp). Phylogenetic
246 relationships were resolved using the maximum-likelihood method through the MEGA
247 software, version 7.0.16 (39). A cutoff of 99.5% was used for MLSA-phylogroups
248 differentiation. Bootstrapping using 1,000 replicates was performed for each analysis (single
249 locus or multiple ones).

250 2.6. Comparative analysis of virulence genes among a SIS *N. cyriacigeorgica* isolate

251 The EML446 strain was selected to represent the SIS *N. cyriacigeorgica* isolates. This
252 strain was chosen as it has been extensively studied in our laboratory. It was grown on BHI
253 agar medium (Difco BD), and its genomic DNA was extracted with the standard phenol-
254 chloroform-isoamyl alcohol method (40). Genome sequencing was performed on a
255 HiSeq2000 Illumina system by GATC (Mulhouse, France). Assemblage and annotation were
256 performed on MicroScope (41). Contigs are available from the BioProject PRJNA542857
257 (42). Comparative analyses were performed against *N. farcinica* IFM 10152^T, *Mycobacterium*
258 *tuberculosis* H37Rv (both already available on NCBI) and the GUH-2 genome reported by
259 Zoropogui *et al.* (43). This strain has been chosen as representative of clinical strains because
260 it has been commonly used as model strain in many physiopathology studies (8). Virulence
261 genes targeted were those identified by Zoropogui *et al.* in the genomic study of GUH-2 strain
262 (43).

263 2.7. Virulence tests with a double-hit murine model of transient immunoparalysis

264 2.7.1. General issues

265 All experiments presented below were approved by the Institutional Animal Care and
266 Use Committee at VetAgro Sup (proposal 1403) in accordance with the European Convention
267 for the Protection of Vertebrate Animals used for Experimental and other Scientific Purposes.

268 Male mice C57BL/6J (Charles River, L'Arbresle, France) of 7-9 weeks of age (20-25
269 g) were housed for one week before beginning the experiments at the Veterinary school
270 (VetAgro Sup, Marcy l'Etoile, France). A 12 hours' dark/light cycle was applied over all
271 experiments. Immunoparalysis of some mice was induced via a moderate cecal ligation and
272 puncture (CLP 30%), and *Nocardia* cells (the SIS EML446 or GUH-2 clinical isolate) were
273 then instilled in the pulmonary airways to test their virulence properties. The CLP procedure
274 (externalization, puncture, antibiotic treatment and pain control) was performed as described
275 in Restagno *et al.* (45). Sham (control-operated mice) underwent laparotomy with exposition
276 of the cecum but without CLP.

277 The starting mouse population was made of 144 individuals. Animals were randomly
278 split into six groups: 1) Sham-NaCl mice which received a saline solution; 2) Sham-GUH-2
279 mice instilled with GUH-2; 3) Sham-EML446 mice instilled with EML446; 4) CLP-NaCl
280 mice which received a saline solution; 5) CLP-GUH-2 mice instilled with GUH-2; and 6)
281 CLP-EML446 mice instilled with EML446. Due to the unpredictable exact mortality rate
282 induced by CLP, more mice were CLP-treated than Sham-treated to reach the right number of
283 CLP subjects. The aim of the “*Nocardia*-free” groups (*i.e.*, Sham-NaCl and CLP-NaCl) was
284 to ensure that mortality was not CLP-dependent. The final experimental design is shown in
285 Figures 2A & B.

286 Prior studies were performed with three different bacterial concentrations (data not
287 shown) to determine the sublethal dose. Strains GUH-2 and EML446 were cultivated in BHI
288 medium and adjusted to 2.0×10^7 CFU/mL to provide an instillation of 1.0×10^6 bacteria in 50
289 μ L of physiological saline.

290 2.7.2. *Mouse survival monitorings*

291 After instillation, the mice were monitored and weighed every day until death or at the
292 end of the experiment, *i.e.* 41 days. In agreement with the Remick laboratory report (46), mice
293 were systematically euthanized when they reached the cutoff point, *i.e.* when they were found
294 in a moribund state identified by the inability to maintain an upright position associated or not
295 with labored breathing and cyanosis. Classical signs of distress, such as anorexia and weight
296 loss (> 20%), hunching, prostration, impaired motility, labored breathing, ruffled haircoat,
297 and dehydration, were assessed. Mice exhibiting at least four of these criteria were euthanized
298 via isoflurane (5%) anesthesia followed by cervical dislocation. Mice exhibiting less than four
299 of these criteria were re-inspected each 8 hours. Then, if the conditions of the mice worsened,
300 they were euthanized. Surviving mice were used for detection of *Nocardia* cells at days 33
301 and 41. Survival curves (Kaplan–Meier plots) were compared by log rank test and performed
302 on Prism8 software. P-values < 0.05 were considered statistically significant.

303 2.7.3. *Detection and visualization of Nocardia cells among mouse organs*

304 Organ histologic examinations were performed on the dead or euthanized mice. The
305 inflammatory response and tissue damages due to *Nocardia* were evaluated at 4, 10, 33 and
306 41 days according to animal mortality curve due to animal ethical reasons linked to the
307 concept of reduction (Figure 2B). Small pieces of kidneys, spleen and liver were removed and
308 fixed in formalin. For brain analysis and lung, the organs were removed and fixed by
309 intratracheal infusion of paraformaldehyde (4%). They were kept in 4% paraformaldehyde for

310 at least 36 hours, dehydrated in successive baths with 30, 50 and 70% ethanol, embedded in
311 paraffin, cut into 8 µm sections and stained with hematoxylin and eosin (47).

312 Crushed organs (lung, kidneys, brain, spleen and liver) diluted in 4.5 mL of a saline
313 physiological solution, and serially diluted (10^{-1} to 10^{-4}) were used to estimate *Nocardia* plate
314 count numbers. These plate counts were performed on BHI agar medium after a validation of
315 the bacterial colonies using a *Nocardia*-specific PCR (NG1/NG2 primers (Table 1)). DNA
316 extracts were produced from 200 µL of the above crushed organs using the NucleoSpin®
317 Tissue kit (Macherey-Nagel, France). The *Nocardia*-specific PCR was then applied on these
318 extracts to verify the presence of *Nocardia* cells in these organs.

319 **3. Results**

320 *3.1. Spatio-temporal diversity of Nocardia cells in soils from an infiltration basin*

321 Regarding physical-chemical parameters, soil water content of the SIS was quite
322 variable between the three sampled areas and campaigns, varying from 9% in the upper zone
323 in summer to saturation (155%) in the inflow zone in spring (Figure 3). General trend was a
324 higher moisture, at least 115%, in the inflow zone, and a lower water content, not higher than
325 56%, in the upper zone. Granulometry did not exhibit much variability between samples, with
326 a relative mean sand content of 55-60% and a clay content of 40-45% (data not shown).
327 Regarding metal trace elements, they have been shown to be constant over the time: Zn
328 concentration remained around 0.5 mg/L and it was the highest concentration detected when
329 comparing to other elements that respect the following relationship: Cd < Pb < Cu < Zn (data
330 not shown). These pollutants were more abundant in the bottom zone than in the inflow and
331 upper zones (p-value = 0.0195) (**Figure 4 A**). The PAHs, taken individually, harbored two
332 clusterization patterns for 12 out of 16. So they were clustered according to the zone

333 (fluoranthene, pyrene, phenanthrene, benzo(a)anthracene, benzo(a)pyrene, hereafter “5
334 PAHs”) or the period of sampling (naphthalene, acenaphthene, fluorene,
335 benzo(b)fluoranthene, dibenzo(a,h)anthracene, benzo(ghi)perylene, indeno(1,2,3-cd)pyrene,
336 hereafter “7 PAHs”). The 4 remaining PAHs (acenaphthylene, anthracene, chrysene and
337 benzo(k)fluoranthene) were not considered in this study because they don’t exhibit any
338 variability. The 5 PAHs were significantly more abundant in the inflow zone (p-value =
339 0.0157), while the 7 PAHs were significantly more abundant during the summer (p-value =
340 6.97×10^{-6}) (**Figure 4 B & C**).

341 DNA sequencing of the *hsp65* PCR products yielded good quality reads for each
342 sediment sample (Table 3). Throughout the three sampling sites, the *hsp65* metabarcoding
343 analytical scheme revealed a high diversity through the actinobacterial community according
344 to Shannon index higher than 5 in all the samples except of SP2 in autumn (Table 3) and the
345 Shannoneven index at 0.8 indicates that microbial populations are on evenness proportions.
346 On a total of 320,760 reads allocated to the *Actinobacteria* class, 125,378 could not be
347 allocated to a defined order. The *Mycobacterium* genus was the most abundant with 132,379
348 reads representing about 41.3% of the identified actinobacteria, but 92,757 reads from this
349 genus could not be allocated to a specific species within the *Mycobacterium* genus. Part of the
350 most abundant *Mycobacterium* species were not pathogen such as *M. neglectum* (31.8%) or
351 *M. sediminis* (19.4%) but some others were opportunistic nontuberculous Mycobacteria (*M.*
352 *doricum*: 22.7%, *M. fluoranthenvivorans*: 2.1% or *M. manitobense*: 1.2%). The *Streptomyces*
353 genus was another important group representing 2.7% of the identified *Actinobacteria* but no
354 pathogenic species were recovered. *Gordonia* and *Nocardia* genus represented each about
355 0,4% and 0.2% of identified *Actinobacteria*, respectively.

356 The *hsp65* metabarcoding approach gave a more general view of *Nocardia* diversity

357 in SIS allowing resolution to the species level (Figure 3). Regarding the *Nocardia* community
358 detected in the DRIB (Figure 3), on the 507 *Nocardia* affiliated reads, only 203 reads could
359 not be identified at the species level, indicating part of diversity that still needs to be resolved.
360 To confirm the accuracy of the Wang text-based Bayesian taxonomic classifications
361 performed with MOTHUR (48), representative sequences of the *Nocardia hsp65*-OTUs were
362 analyzed by BLASTn searches using the GenBank database. These searches confirmed all
363 taxonomic inferences and showed that the unclassified sequences did not share enough
364 identities to be clearly allocated to a particular species taking into account the chosen species
365 identification criterion. However, even in these cases, closest results in terms of sequence
366 similarity belonged to *Nocardia* genus. Hereafter, only reads allocated to well-defined species
367 were analyzed and compared. A total of fourteen *Nocardia* species (relative abundance >1%)
368 could be tracked using the *hsp65* metabarcoding approach (Figure 3). Most of these are
369 opportunistic pathogens which also belong to the species the most clinically relevant in
370 France (the classification of most/intermediate/less frequent pathogens was based on Lebeaux
371 *et al.* (43)). In autumn, the most prevalent *Nocardia* species in the SIS (inflow, bottom, upper
372 zones) were *N. cyriacigeorgica* and *N. nova* (clinically-relevant) together with *N. asteroides*
373 and *N. cunmideleus*. A segregation between zones was observed (Figure 3). *N.*
374 *cyriacigeorgica hsp65* reads were mainly observed in bottom zone (humidity rate = 72%) of
375 the DRIB, while *N. nova* reads were mainly recovered from the inflow zone (water saturated
376 samples). In spring, *N. asteroides hsp65* reads were observed in the three zones, and this
377 species appeared to be the most prevalent in this environment. Reads from nonpathogenic *N.*
378 *globerula* and the complex *N. abscessus/asiatica* (human pathogens) were also obtained in
379 high numbers in the inflow and upper zones. In summer, most of the detected species were
380 non-pathogenic ones: *N. globerula* showed the highest number of reads over the DRIB, but *N.*
381 *salmonicida* (a fish pathogen) reads were higher in the inflow zone (see Figure 3 for a

382 summary of the taxonomic allocations).

383 **Table 3.** Diversity and richness indices of the *hsp65* metabarcoding analysis

384

Zone	Sample	Shannon	Shannon-	Number	Shannon	Shannon-	Number	Shannon	Shannon-	Number
		Autumn			Spring			Summer		
		even	even	of reads	even	even	of reads	even	even	of reads
Inflow	SP1	6.769454	0.850362	14,852	6.417540	0.809400	20,781	6.223242	0.789785	18,477
	SP2	4.599154	0.618105	20,279	5.138384	0.668644	21,826	6.114367	0.767870	20,255
	SP3	6.372891	0.806140	16,725	6.049869	0.776262	19,069	6.129255	0.782352	18,223
	SP4	6.451997	0.807211	17,855	6.105381	0.777861	19,702	5.597741	0.729251	16,189
	SP5	6.147986	0.787445	16,296	5.602082	0.741520	14,079	5.602386	0.734353	21,506
Bottom	SP6	5.862727	0.763773	15,311	6.324151	0.800348	15,733	6.416230	0.808466	20,265
	SP7	5.402151	0.710725	21,924	5.654639	0.731502	21,151	6.048270	0.769293	21,551
	SP8	6.374864	0.802821	19,138	6.418229	0.808028	17,795	6.207508	0.782555	22,832
	SP9	6.621081	0.835147	17,992	5.838846	0.752964	16,046	6.156486	0.778947	23,940
	SP10	6.360877	0.810332	19,628	6.180622	0.787997	15,732	6.580971	0.829486	16,981
Upper	SP11	6.597122	0.830997	19,100	6.246701	0.788356	18,921	7.194320	0.873336	25,995
	SP12	6.013042	0.764439	19,922	6.188532	0.779951	19,689	7.179103	0.877277	22,632
	SP13	6.879646	0.858737	17,783	6.350625	0.800308	19,069	7.185295	0.878364	24,621
	SP14	6.405939	0.807024	17,387	5.948913	0.754248	19,584	7.126191	0.873712	23,833
	SP15	6.636493	0.833514	17,664	6.278243	0.795136	17,258	6.496152	0.814367	24,676

385 Shannon is indicative of microbial community diversity whereas Shannoneven is
 386 indicative of the evenness proportion of different populations within the microbial
 387 community.

388

389 Regarding the distribution of *hsp65*-OTUs per species, some were repeatedly observed
 390 from one campaign to another. Some *hsp65*-OTUs were found in multiple areas of the DRIB.
 391 Regarding *N. cyriacigeorgica*, which represents one of the species of most health concern for
 392 this genus, a total of 39 sequences was identified, representing 6 *hsp65*-OTUs. Most of these
 393 sequences (n=36) were recovered from the bottom zone of the DRIB, but one sequence was
 394 recovered from the upper zone. Only one sequence of this species was recovered over two
 395 sampling campaigns.

396 Regarding the relative abundance of *Nocardia* species, they were clustered according
397 to their potential health hazard or not. The pathogenic species are significantly less abundant
398 in spring and in summer than in autumn and different between these two first seasons (p-
399 values = 0.04066 for spring vs autumn and 0.02953 for summer vs autumn). The non-
400 pathogenic species, *i.e.* not recognized as implicated in human nocardiosis, are most present
401 in the upper zone than in the two others (p-value = 0.0294 for the upper zone vs inflow and
402 bottom zones) (**Figure 4 D & E**).

403 When we compare both physical-chemical parameters and *Nocardia* relative
404 abundance, we could see that there exists a clusterization according to the sampling areas.
405 Humidity and 5 PAHs explained the clusterization in the inflow zone, while only non-
406 pathogenic *Nocardia* species explained the clusterization in the upper zone. In the bottom
407 zone, the explaining parameters were the pathogen *Nocardia* species, 7 PAHs, the soil
408 granulometry and the heavy metals content (**Figure 5 A&B**).

409 The correlogram highlighted the potential correlation between metal trace elements
410 and the relative abundance of non-pathogenic and pathogenic *Nocardia* species (**Figure 6**).
411 Indeed, a negative correlation was observed between the presence of zinc and mercury (p-
412 value < 0,03) and the non-pathogenic species while there was a positive correlation between
413 the presence of mercury and copper and pathogenic species. It could be also noted a negative
414 correlation between the 7 PAHs and the relative abundance of pathogenic species (p-value <
415 0,08).

416 To support the inferences made by the *hsp65* metabarcoding approach, attempts at
417 isolating *N. cyriacigeorgica* strains from the DRIB sediment samples were performed. An
418 averaged *N. cyriacigeorgica* plate count number of 1.0×10^3 CFU/g dry sediment was obtained
419 (Figure 1). From these platings, some isolates were purified, and twelve were confirmed to be

420 *N. cyriacigeorgica* strains (by *rrs* and *hsp65* gene-based typing) and named DjR1 to 12
421 (2015-2016 isolates) and EML1456, already sequenced in Vautrin *et al.* (42) (Table 2).
422 Several other species were also detected (*N. abscessus*, *N. nova*, *N. ignorata*, *N. asteroides*, *N.*
423 *salmonicida*, etc.) but out of scope in this study. The DjRm13-15 are sequences identified
424 according to the metabarcoding analysis on the sampled sediments of the DRIB (2015-2016).
425 Location of these SIS isolate community and sequences over the DRIB is indicated on Figure
426 1. Deeper taxonomic allocations of these strains by phylogenetic analysis are shown below.

427 3.2. Phylogenetic relatedness of SIS and clinical isolates

428 To go deeper into the evaluation of health hazards that can be associated with SIS *N.*
429 *cyriacigeorgica* isolates, we obtained the phylogenetic trees built from individual locus and
430 concatenated genetic loci (*rrs*, *hsp65*, *sodA*, *secA1*) (Figure 7). The phylogenetic tree based on
431 *rrs* gene grouped all the clinical and environmental *N. cyriacigeorgica* strains into a single
432 phylogroup (data not shown). The phylogenetic tree obtained from *hsp65* gene sequences of
433 the analyzed strains in this work showed a distribution into three significant phylogroups (PI,
434 PII and PIII) that matched previous groups defined by Schlaberg *et al.* (9). Thanks to some
435 sequences representatives from the metabarcoding analysis and integrated in *hsp65*-tree, we
436 could observe that different genotypes were splitted onto two different *hsp65*-phylogroups (I
437 and III) for *N. cyriacigeorgica* species (Figure 7A: DjRm sequences only). Phylogroup I (PI)
438 harbored the *N. cyriacigeorgica* type strain *hsp65* sequence, environmental sequences
439 obtained from the metabarcoding analysis (DjRm.12 and DjRm.14) and French clinical
440 strains, including one belonging to a patient from Lyon, which was a resident in the
441 geographic area of the DRIB. This phylogroup can be seen as a clonal complex as it
442 encompassed 5 different OTUs issued from the metabarcoding analysis. Phylogroup II (PII)
443 harbored clinical strains including GUH-2. Phylogroup III (PIII) also harbored environmental

444 sequences obtained from the metabarcoding analysis (DjRm.13, SIS *N. cyriacigeorgica*
445 isolates (EML446, DjR.1 and DjR9) and clinical strains. Only 1 OTU joined this phylogroup.
446 Bootstrap values were >80 only for PI (85) and PIII (98) (Figure 7a).

447 For *sodA* and *secA1*, the phylogenetic tree structures were in agreement with the one
448 derived from *hsp65* gene (data not shown). The MLSA phylogenetic tree of concatenated
449 gene sequences increased the reliability of the groupings (Figure 7B). This tree better resolved
450 the relatedness of the SIS isolates with the clinical strains. Low infra-specific divergences
451 were observed within PI (similarity mean = 99.7%, min-max = 99.4-100%) and PII (similarity
452 mean = 99.8%, min-max = 99.6-100%). In PIII, infra-specific divergences were higher (mean
453 identities = 99.5%, min-max = 99.2%-100%) (Supplementary Table 1). However, only PI and
454 PIII clustered clinical and environmental strains, instead of PII that clusters only clinical
455 strains. The MLSA tree showed a clear phylogroups distribution harboring SIS and also
456 clinical strains from the French Observatory of Nocardiosis. These phylogroups that matched
457 those of *hsp65* ones. So, in the same way as for *hsp65* tree, we can state that MLSA-
458 phylogroup I represent a clonal complex inside *N. cyriacigeorgica* species.

459 3.3. EML 446 and GUH-2 genome comparisons

460 The whole genome sequencing (WGS) of DRIB *N. cyriacigeorgica* EML446 (MLSA-
461 PIII) resulted in the obtaining of 41 contigs that could be assembled into a circular
462 chromosome of 6,530,670 bp with a G+C content of 68.21%. This genome encodes 51 tRNA,
463 3 rRNA and 6,230 CDSs (coding sequences). Analysis on the MicroScope platform with the
464 Virulome tool highlighted the presence of 130 CDSs (*i.e.* 2.09% of the CDSs) that can be
465 involved in virulence in EML446 (MLSA-PIII) while 108 CDSs involved in virulence were
466 found among GUH-2 (MLSA-II), the model strain to investigate virulence in *Nocardia*.
467 Among these CDSs, 96 were found to be in common. A Venn diagram was drawn to highlight

468 the number of shared CDSs between *N. cyriacigeorgica* EML446 (MLSA-PIII), GUH-2
469 (MLSA-PII) and DSM 44484^T (MLSA-PI), and *N. farcinica* IFM10152^T and *M. tuberculosis*
470 H37Rv (Figure 8a). The EML446 genome shared 4,392 CDSs with GUH-2 and 4,883 with
471 DSM 44484^T. GUH-2 and DSM 44484^T shared 4,408 CDSs, and the number of shared CDSs
472 between any *N. cyriacigeorgica* strain and *N. farcinica* IFM 10152^T or *M. tuberculosis*
473 H37Rv was lower. Comparison of 11 gene families between these genomes did not show
474 many differences except for polyketide synthase (27 for *M. tuberculosis* compared to 7 for the
475 three *Nocardia* genomes), lipoproteins (62 compared to 7-16) and PE_PGRS ((proline-
476 glutamic acid_polymorphic guanine-cytosine-rich sequence) 62 compared to 0-1). Only the
477 NRPS (nonribosomal peptides synthetases) CDSs were lower for *Mycobacterium*, having 3
478 compared to 14-17 for *Nocardia* (Table 4). According to these analyses, the EML446 strain
479 from the UIB polluted sediments has the genetic potential to be as virulent as the *N.*
480 *cyriacigeorgica* GUH-2 clinical isolates. On the other hand, the MAUVE analysis highlights a
481 decrease in CDSs content in the region of genomic plasticity (RGP) between the two genomes
482 *N. cyriacigeorgica* EML446 and GUH-2 (Figure 8b)

483 **Table 4.** Number of virulence CDSs (coding sequences) per functional categories shared between *N. cyriacigeorgica* GUH-2, EML446 and two
 484 other pathogenic species *N. farcinica* IFM10152^T and *Mycobacterium tuberculosis* H37Rv:
 485

Virulence CDSs	<i>N. cyriacigeorgica</i> EML446	<i>N. cyriacigeorgica</i> GUH-2	<i>N. farcinica</i> IFM10152 ^T	<i>M. tuberculosis</i> H37Rv
PKS (polyketide synthase)	7	7	7	27
NRPS (nonribosomal peptides synthetases)	17	17	14	3
Lipoproteins (Lpps)	16	17	7	62
Hemolysin	2	2	2	2
Esterases	18	17	24	15
PE_PGRS (proline-glutamic acid_polymorphic guanine-cytosine-rich sequence)	1	1	0	62
<i>sod</i> (superoxide dismutase)	2	2	2	2
<i>mce</i> (mammalian cell entry)	6	6	7	4
<i>cat</i> (catalase)	3	3	3	1
<i>narBGHIJK</i> (nitrate reductase)	6	6	5	6
<i>nirBD</i> (nitrite reductase)	2	2	2	2

486
 487 Note: number of genes were found by keyword search tool on the MicroScope platform (<http://www.genoscope.cns.fr/agc/microscope/>).
 488 Virulence genes were selected according to (43).

489

490 3.4. *Physiopathology of N. cyriacigeorgica in a murine model of transient immunoparalysis*

491 In the present study, 37% of the CLP-operated mice died prior to the instillation step
492 (37/101) in accordance with the model of septic immunoparalysis established by Restagno *et*
493 *al.* (45). Five days after the first hit, *i.e.* the CLP operation, Sham and CLP-operated mice
494 were randomized as described in Figure 2.

495 The survival results showed a survival rate of 100% after 41 days for Sham-NaCl and
496 CLP-NaCl groups. For mice intratracheally instilled by a load of *Nocardia* at 1.0×10^6
497 CFU/mouse, Sham-EML446 showed the same survival rate (Figure 9). In the Sham-GUH-2
498 group, only one mouse died at D6 (survival rate = 86% (6/7 mice)). For both the CLP-GUH-2
499 and CLP-EML446 groups, during the first 10 days following intratracheal bacterial
500 challenges, the survival rates were not statistically different (p-value = 0.36): CLP-GUH-2,
501 survival 67% (6/9 mice) and CLP-EML446, survival 64% (7/11 mice). A second episode of
502 mortality occurred at 30 days for CLP-GUH-2, decreasing the survival rate at 44% (4/9 mice).
503 As compared to respective sham groups survival decreased significantly for CLP-GUH2 (p-
504 value = 0.02) and CLP-EML446 (p-value = 0.04) groups.

505 The TCBD (time course bacterial detection) experiment showed that in the Sham-
506 operated group, soon after intratracheal instillation, the lung was the primary infection site of
507 *Nocardia*, but other organs were also affected. At D4, 4/5 Sham-GUH-2 mice presented
508 *Nocardia* in the lungs; in two of them, *Nocardia* was detected in all the studied organs (Table
509 5). Only in one mouse, *Nocardia* could not be detected. At D10 and D33, the number of
510 organs positive for *Nocardia* decreased except in the lungs (2/3 mice at D10 and 3/4 mice at
511 D33) and kidneys (1/3 at D10 and 3/4 at D33). For the Sham-EML446 mice, the occurrence
512 rate and dissemination were lower than for the Sham-GUH-2 mice. At D4, *Nocardia* was
513 found in the lungs of 3/5 mice, but the incidence in other organs was lower (1/5) and even

514 null in the brain and spleen. At D10, only 1/5 mice presented *Nocardia* in each organ, and
515 nothing was detected at D33. However, *Nocardia* was found in all organs of one mouse at
516 D41.

517 In the CLP-operated group, at D4, the presence of *Nocardia* was observed in all of the
518 inspected lungs for both strains (5/5 for CLP-GUH-2 and 6/6 for CLP-EML446), but it was
519 almost missing in the other organs (only in 3/5 CLP-GUH-2 and 2/7 CLP-EML446 in the
520 kidneys). At D10, 100% of the inspected lungs were still positive for *Nocardia*, and all the
521 other organs became progressively positive: 3/5 (CLP-GUH-2) and 4/4 (CLP-EML446)
522 positive in the kidneys, 4/5 (CLP-GUH-2) and 3/4 (CLP-EML446) positive in the brains and
523 spleens, 3/5 (CLP-GUH-2) and 3/4 (CLP-EML446) positive in the livers. At D33, the
524 presence of *Nocardia* in the lungs remained especially high for CLP-GUH-2 mice (7/8) but
525 relatively low for CLP-EML446 mice (2/5). Dissemination in other organs was almost similar
526 for both strains excepting for the liver. At D41, almost all the organs of CLP-EML446 mice
527 were positive at high rates (Table 5). As expected, no NaCl-operated mice (controls) showed
528 *Nocardia* cells in their organs.

529 **Table 5.** Heatmap representing rates of contaminated mice by *N. cyriacigeorgica* and
530 detected with specific *Nocardia*-genus PCR. Mice were sacrificed 4, 10 and 33 days after
531 instillation for the GUH-2 strain and until 41 days for the EML446 strain. Lungs, brain,
532 kidneys, spleen and liver were sampled for *Nocardia* detection. Clinical strain GUH-2 is
533 colored in blue, and environmental EML446 is colored in orange. The deeper is the color, the
534 more contaminated are the mice. The heatmap only reveals the presence or absence of
535 *Nocardia* but is not quantitative. ^H some organs were used for histological analysis and were
536 not available for bacterial detection since it is not technically possible to do both. ^{n=x}
537 corresponds to the number of mice reserved for histological analysis.

		Lungs	Kidneys	Brain	Spleen	Liver
Sham-GUH-2	4 days	4/5	2/5	2/5	2/5	2/5
	10 days	2/3	1/3	0/3	0/3	0/3
	33 days	3/4	3/4	1/4	2/4	0/4
Sham EML446	4 days	3/5	1/5	0/5	0/5	1/5
	10 days	1/5	1/5	1/5	1/5	1/5
	33 days	0/2	0/2	0/2	0/2	0/2
	41 days	1/3 ^{H n=2}	1/5	1/5	1/5	1/5
CLP GUH-2	4 days	5/5	3/5	0/5	0/5	0/5
	10 days	5/5	3/5	4/5	4/5	3/5
	33 days	7/8	4/8	4/8	3/8	0/8
CLP EML446	4 days	6/6 ^{H n=1}	2/7	0/7	0/7	0/7
	10 days	4/4	4/4	3/4	3/4	3/4
	33 days	2/5	3/5	3/5	2/5	3/5
	41 days	4/4 ^{H n=3}	7/7	4/5 ^{H n=2}	7/7	6/7

539

GUH-2	EML446
0 – 19 %	0 – 19 %
20 – 39 %	20 – 39 %
40 – 59 %	40 – 59 %
60 – 79 %	60 – 79 %
80 – 100 %	80 – 100 %

540

541 In the lungs of CLP-EML446 mice (1/1), histological signs of pneumonia similar to
542 nocardiosis were clearly observed, as histologic pictures showed multiple cavity lesions at D4
543 (Figure 10A & B). A strong mononuclear infiltrate in the periphery of the microabscesses
544 among the collagen fibers was also observed (Figure 10A1, A2 & B3). The presence of
545 numerous filamentous bacteria in the caseous necrosis area suggests that these granulomatous
546 lesions were infectious and that the mice developed nocardiosis (Figure 10B). At D41, 1/3
547 CLP-EML446 mice exhibited mild pneumonia (Figure 10C). Nothing was observed at D41 in
548 Sham-EML446 mice (0/2) (Figure 10D). As expected, no histological lesions in any organ
549 were observed for Sham-NaCl (n=2 at day 41) or CLP-NaCl (n=2-3 at D4, D10, D33 and n=6
550 at D41) mice. No lesions were found in the brains of CLP-EML446 mice at D41 (n=2). No

551 specific nocardiosis lesions were found in other organs (kidney, spleen, liver) at D4, D10,
552 D33 of either Sham-EML446 (n=2/organ) or CLP-EML446 (n=2/organ) mice. At D41, no
553 specific nocardiosis lesions were found in kidney, spleen, liver of Sham-EML446 (n=5/organ)
554 mice, but small granulomas were observed in the livers of 5 CLP-EML446 (n=7/organ) mice
555 (Figure 10E).

556 **4. Discussion**

557 The dissemination of hazardous biological agents in cities, outside hospital settings, remain
558 largely under explored. Urban soils and waters can offer shelters for some pathogenic micro-
559 organisms such as the opportunistic ones. With their increasing contact with human
560 populations, these pathogens might be undergoing selective processes that will make them
561 better fit for a colonization of the human host. Furthermore, urban chemical pollutants seem
562 to generate a “dangerous liaison” with these micro-organisms as demonstrated by Cui *et al.*
563 (49) who found the presence of 16 bacterial genera harboring pathogenic species such as
564 *Aeromonas* and *Mycobacterium* in polluted lakes in an industrial area in China. Furthermore,
565 Chulan *et al.* (50) showed a relation between concentrations of chemical pollutants and
566 airborne pathogenic bacteria in air samples. Regarding *Nocardia*, also, some authors have
567 already described a potential relationship between organic pollutants and presence of
568 *Nocardia*. As explained by Arrache *et al.* (51) this relationship could explain the infective
569 source associated to a case of cerebral nocardiosis of an immunocompetent individual
570 exposed to the inhalation during a long period of time of dusts rich in hydrocarbons in a
571 refinery which probably hosted the *Nocardia cyriacigeorgica* species responsible of his
572 pathology.

573 *Nocardia* are known to be widely spread among outdoor environments but several
574 species represent a public health concern. This genus includes opportunistic pathogens that

575 primarily cause pulmonary infection following inhalation (52, 53). These species can cause
576 pulmonary nocardiosis in immunocompromised individuals associated with high-dose
577 corticosteroids treatments (53, 54). Non-immunocompromised patients like cigarette smokers,
578 or those affected by bronchiectasis and acute bronchitis, and other chronic pulmonary
579 diseases, are also at risk of pulmonary nocardiosis. These clinical pictures affect about 60
580 million people around the world, according to the WHO, and are often related to high
581 atmospheric pollution including high content in aerosolized dusts. These dusts can be
582 generated by several urban components such as motor engines, chemical industries, garbage
583 incinerators, stored garbage on sidewalks, plant and animal detritus, etc. They are
584 accumulating on urban surfaces and washed away with the runoff waters during rain events or
585 aerosolized. Polluted urban runoffs are nowadays transferred either to wastewater treatment
586 plants (WWTP), SIS, or natural waterways. These washed urban sediments can thus create
587 novel growth conditions for opportunistic human pathogens that are known to be well-adapted
588 for a growth on chemical pollutants. Results obtained in this study supported this hypothesis
589 as high numbers of *N. cyriacigeorgica* cells were observed among SIS sediments, and these
590 cells were allocated to a MLSA-phylogroup harboring clinical strains that had been involved
591 in lung infections. Other pathogenic *Nocardia* species such as *N. abscessus* (associated
592 mainly with cerebral and pulmonary infections), *N. nova* (related with pulmonary and
593 cutaneous cases), and *N. otitidiscaviarum* (causing mainly cerebral infections and multidrug
594 resistant) were also identified among SIS through the use of a novel metabarcoding approach
595 based on the *hsp65* gene target. These species can thus also be disseminated through SIS and
596 aerosolized deposits.

597 Sediments analyzed from the DRIB showed variable water contents, high hydrocarbon
598 pollution and variable plant cover. The main pollution recorded in this urban environment is
599 due to PAHs. The high amount of PAHs has a pyrogenic origin according to the

600 phenanthrene/anthracene <10 and fluoranthene/pyrene >1 ratio for all the samples in
601 accordance with Budzinski *et al.* (55) and Yunker *et al.* (56). This pyrogenic origin could be
602 explained by the industrial activity in this area and the engine gasoline combustion. No
603 petrogenic origin could be identified in this study, contrary to Marti *et al.* (16) that reported a
604 petrogenic origin for most of the PAHs in the detention basin upstream the infiltration basin,
605 indicating a good performance in removing oils from water and avoiding the plugging of the
606 infiltration basin. Nadudvari & Fabianska (57) reported a pyrogenic origin in sediments in a
607 river in Poland arising from runoff water and city waste combustion. The same phenomenon
608 was first described by Radke & Welte (58) in oil wells in Canada. Five of the detected PAHs
609 (phenanthrene, fluoranthene, pyrene, benzo(a)anthracene and benzo(a)pyrene) were already
610 reported by Sebastian *et al.* (59) in the same DRIB as being specially more abundant for the
611 inflow zone which is confirmed by our study. On the other side, we observed a group of 7
612 PAHs (naphthalene, acenaphthene, fluorene, benzo(b)fluoranthene, dibenzo(a,h)anthracene,
613 benzo(ghi)perylene, indeno(1,2,3-cd)pyrene) that are more abundant in summer regardless the
614 measured zone which agrees with the study of Belles *et al.*, (2016) also in the same studied
615 system. Regarding the correlation between pollutants and relative abundancy of *Nocardia*,
616 this high amount of PAHs seems to inhibit the development of *Nocardia* species, while metal
617 trace elements can be an explicative factor of pathogenic species presence in the bottom zone
618 of this infiltration basin. Regarding the influence of PAHs in pathogenic *Nocardia*, the upper
619 zone, considered as the less contaminated zone, was the one with lowest counts in *N.*
620 *cyriacigeorgica*. So, we could consider that upper zones represent a low risk of infection by
621 aerial pathway as respiratory infections are the most usual in nocardiosis. However, other
622 pathways should not be excluded (*e.g.* direct contact with water, transmission to phreatic zone
623 below, etc.). Moreover the persistence of pathogenic *Nocardia* in humid areas could be seen

624 as a reservoir as it could be transmitted by aerosolization in eventually favorable
625 environmental conditions (*e.g.* dryer period).

626 Hsp65-based metabarcoding revealed as a powerful tool in the identification at species
627 level of Actinobacteria but showed some limitations evidenced by the number of reads that
628 could not be assigned to this class. This may be possibly due to the lack of hsp65 sequences in
629 used database for many Actinobacteria taxa. However, obtained results from this *hsp65*
630 metabarcoding analysis concerning species-level *Nocardia* taxa, showed that water content
631 was found to be explicative of some distribution patterns. Shannon indices highlighted a high
632 diversity within the sampled sediments but similar values could indicate a low variability
633 between the samples in a same area. The lower diversity within the SP2 sample in autumn
634 (according to its Shannon index) could be explained by the fact that this sample point was the
635 one with the highest humidity rate among those of the other sample points of the inflow area
636 (max humidity $\geq 115\%$). Extremely high humidity rates may favor the development of few
637 bacterial communities detrimental to others. In this study, we observed that a high water
638 content ($\geq 112\%$) was related to the higher number of *N. abscessus* and *N. nova hsp65* reads,
639 and intermediate conditions (around 72% sediment moisture) led to the higher number of *N.*
640 *cyriacigeorgica hsp65* reads. Most of the species found with the metabarcoding tool were also
641 detected by cultural method reinforcing the validity of this first approach. However, no
642 quantification nor comparison of the accuracy of these techniques was in the scope of this
643 work.

644 These observations suggest *N. cyriacigeorgica* to be the most worrisome species as its
645 environmental conditions for isolation are those present the most part of the year in the
646 studied system and besides, according to Zoropogui *et al.* (43) this species presents a likely
647 ongoing evolution towards a higher tropism for the human host. This led us to further

648 investigate the SIS *N. cyriacigeorgica* isolates. The phylogenetical analysis demonstrated
649 close relationships between the clinical and some of the SIS *N. cyriacigeorgica* of this study.
650 Moreover, a recent clinical strain from a patient living in the Lyon area was also positioned in
651 one of the MLSA-phylogroups harboring SIS's *N. cyriacigeorgica* strains, allowing us to
652 wonder of the source of infection of this person. In fact, clinical strains and environmental
653 ones could not be segregated into distinct clusters. This supports the idea that all strains of *N.*
654 *cyriacigeorgica* represent a human health hazard. Direct exposures through inhalation could
655 thus result in pulmonary nocardiosis. However, the degree of exposure with the human host
656 could have generated a gradient of virulence potentialities going from mild to severe. This
657 hypothesis was verified with the results of the genomic and 'in vivo' comparison of GUH-2
658 and EML446 strains.

659 Regarding the genomic analysis, the genome size of SIS *N. cyriacigeorgica* EML446
660 (6,530,670 bp) was found to be larger than the one of clinical GUH-2 (6,194,645 bp). In
661 general, opportunistic pathogens present in an environmental reservoir harbor a larger
662 genome, conferring a greater versatility in the use of nutrients and in the ability to resist at
663 certain environmental constraints (60). However, the differences in genome size between
664 EML446 and clinical GUH-2 are low, and not in favor of a size reduction related to greater
665 interactions with the human host. Furthermore, the content in virulence genes between these
666 strains did not differ substantially but its distribution is variable in both genomes. Indeed,
667 Zoropogui *et al.* (43) had already demonstrated that part of the virulence genes of clinical
668 GUH-2 strain is contained in the RGPs (region of genomic plasticity) (Figure 8b). In fact,
669 several CDSs found in *N. cyriacigeorgica* EML446 could play part in lung colonization. For
670 example, *mbt* is involved in the development of slow-growing bacteria in tissues under iron
671 limitations, *e.g.* in the lung (61). In addition, the presence of CDSs coding for lipoproteins
672 may contribute at the intracellular lifestyle of *N. cyriacigeorgica*. Actually, Li *et al.* (62)

673 demonstrated that a large amount of lipoproteins improves survival of another actinobacteria,
674 *Mycobacterium smegmatis*, in macrophage cells and murine lungs. Finally, CDSs such as
675 PE_PGRS30, PG_PGRS33 and PE_PGRS41, were shown to be essential for entry and
676 intracellular survival in macrophages. These virulence genes were also recorded in clinical *N.*
677 *cyriacigeorgica* GUH-2 strain. This suggests that virulence of these two strains on a mouse
678 model (or in human) should be similar.

679 Anyway, our work focused in a limited number of virulence genes (43) so, it would be
680 interesting further complete virulome studies including more genomes from other *Nocardia*
681 species and from other pathogenic microorganisms in order to obtain more information about
682 the genetic factors of the virulence in *Nocardia*.

683 However, changes in regulatory pathways might have generated some changes in the
684 fine-tuning of the expression of virulence genes, and modified the clinical outcomes. This is
685 hardly detected by simple genomic comparisons, and would require gene expression profile
686 analyses as performed by Cruz-Rabadán *et al.* (64). Nevertheless, it must be noted that the
687 GUH-2 MLSA-phylogroup (PII) diverged from the EML446 MLSA-phylogroup (PIII),
688 suggesting adaptation and selection for distinct context (human *vs* environment). This on-
689 going adaptation of the GUH-2 lineage is certainly due to repeated passed through humans,
690 and this hypothesis thus needed to be studied.

691 The pathophysiology of *N. cyriacigeorgica* has been extensively studied in
692 immunocompetent murine models (9). However, nocardiosis has a higher occurrence in
693 immunocompromised patients, and this led us to use a transient immunoparalysis (induced by
694 mild CLP) mouse model to compare the virulence potentialities of GUH-2 and EML446. This
695 is the first report describing the use of this model system with *Nocardia* cells, but previous
696 validations had been performed with *P. aeruginosa* cells (45). Compared on this latter study,

697 our *Nocardia* cell instillations at 1.0×10^6 CFU/ CLP mouse led to a lower survival at D8
698 (CLP-EML446 = 64%, CLP-GUH2 = 67%) than the one obtained with *P. aeruginosa* at the
699 same concentration after the same period of time (CLP=93%) (45). Furthermore, in our study,
700 thirty days after GUH-2 instillations, a second mortality wave decreased considerably this
701 survival rate (CLP-GUH-2=44%), which remained much lower than the one obtained with *P.*
702 *aeruginosa*. When extrapolated to humans, these differences in behavior become of high
703 importance in diseases in which both *N. cyriacigeorgica* and *P. aeruginosa* may cohabit. This
704 has been the case in some patients with cystic fibrosis, as reported by Rodriguez-Nava *et al.*
705 (2). Traditionally, in these cases, primary antibiotic treatments will target *P. aeruginosa*
706 because it is considered the main pathogen associated with lung deteriorations. Our study
707 shows that the infective process of *N. cyriacigeorgica* must not be underestimated and that a
708 treatment targeting both pathogens should be performed. This strategy has already been
709 applied, and improved the clinical outcomes (2).

710 It is to be noted that based on this study, a small window of immunoparalysis due to
711 CLP in mice was found sufficient to promote the colonization, persistence and dissemination
712 of *N. cyriacigeorgica*. This observation could also be extrapolated to humans for which
713 during a period of illness, the immune system of the patients is weak (diabetes, cancer, etc.)
714 allowing the colonization of *Nocardia* in the infected organ during this “immunosuppressive
715 window”. Indeed, some patients considered immunocompetent may have had a history of
716 disease (65). So, according to our observations, we can state that locations presenting
717 environmental conditions similar to the studied UIB in terms of humidity and metallic trace
718 elements, may suppose an infective risk for weak populations.

719 Our study presents some limitations. For example, we studied a single UIB with its
720 own geological, hydrological, chemical and vegetative characteristics. We must then be

721 prudent in extrapolating the obtained results to any other UIBs. In this work, we compared the
722 virulence of clinical and environmental strains of *N. cyriacigeorgica*, however just a single
723 representative of each has been chosen. So, obtained conclusions should be carefully
724 extrapolated to all *N. cyriacigeorgica* clinical and environmental strains.

725 However, our work has several strengths. For the first time, an urban, humid and
726 polluted environment (an UIB) was used to study the spatiotemporal distribution of pathogens
727 such as *N. cyriacigeorgica*. Moreover, we used the MLSA approach with precise gene
728 concatenation (*rrs-hsp65-sodA-secA1*) to identify clonal lineages between clinical and
729 environmental strains of *N. cyriacigeorgica*. In addition, according to the results obtained in
730 this approach, it may be possible that *N. cyriacigeorgica* were a complex and not a single
731 species, taking into account the presence of three MLSA-phylogroups whose taxonomy levels
732 remains to be specified. We used, for the first time, the *hsp65* marker for the metabarcoding
733 approach, allowing us to evaluate *Nocardia* biodiversity and to directly detect *N.*
734 *cyriacigeorgica* in the urban sediments of an UIB. This new marker could also be used to
735 track *Nocardia* in other environments. For the first time, the hazardousness of an
736 environmental strain of *N. cyriacigeorgica* isolated from an UIB was studied by complete
737 genome sequencing and by a murine model of transient immunoparalysis, which better
738 mimics the population most frequently targeted by this pathogen.

739 **5. Conclusion**

740 This study presents, for the first time, a complete inventory of *Nocardia* causative
741 agents of human nocardiosis found in a HAP-polluted urban infiltration system. This study
742 was made possible by the development of an innovative metabarcoding *hsp65*-based
743 approach. This led to the detection of the highly frequent species found in nocardiosis
744 worldwide such as *N. cyriacigeorgica*, *N. nova*, and *N. abscessus*. The number of reads per

745 species were found related with the field conditions. *N. cyriacigeorgica* was the most
746 abundant pathogenic species, which makes this species the most worrisome in this kind of
747 environments taking into account its epidemiologic characteristics. This led us to perform
748 rounds of isolation of this bacterial species, and investigate more deeply their molecular
749 epidemiology. A MLSA approach demonstrated a close proximity between the SIS *N.*
750 *cyriacigeorgica* isolates and the clinical ones. However, these strains were mostly distributed
751 in MLSA-phylogroup III, and not recorded in the MLSA-phylogroup II harboring the most
752 virulent isolate recorded so far *i.e.* GUH-2.. This suggested a significant diversification
753 between these strains that could be indicative of a distinct tropism for the human host. This
754 led us to compare the full genome of one SIS isolate with the one of GUH-2. No distinction in
755 their virulence gene contents could be made, suggesting similar virulence potentialities.
756 Experimentations were performed to test the virulence differences among these strains (and
757 MLSA-phylogroups). GUH-2 strain was shown to be the most virulent isolate in an
758 immunoparalysis CLP mice model but EML446, the SIS isolate, was also confirmed to be
759 significantly virulent. This fact further supports the idea that all strains of this species can be
760 pathogenic but with variable clinical outcomes. The GUH-2 lineage seems to be on-going an
761 adaptation for animal hosts including humans. These differentiations linking MLSA-
762 phylogroups and virulence will now need to be tested on a larger number of strains. The
763 presence of pathogenic *Nocardia* is positively correlated to metallic trace elements and they
764 can be suggested to be indicators of their presence. On the other hand, we have observed a
765 negative correlation between PAHs and relative abundance pathogenic *Nocardia*, so these
766 pollutants cannot be used as indicators of these bacteria. Overall, this study shows that humid
767 and polluted environments such as UIBs may represent a health hazard for adjacent
768 populations through either direct exposure or through an aerosolization of dusts harboring *N.*
769 *cyriacigeorgica* cells.

770 **Acknowledgment:** Florian Vautrin held a doctoral fellowship from the Auvergne-Rhône-
771 Alpes region. The authors thank the Institut Claude Bourgelat – Biovivo of VetAgro Sup for
772 animal facilities and the OTHU (Field Observatory for Urban Water Management,
773 <http://www.graie.org>) for access to the infiltration basin. The authors are grateful to Béatrice
774 Charton for her technical support on the *Nocardia* PCR detection, Audrey Dubost and Danis
775 Abrouk for their bioinformatics knowledge, Christophe Martin for the first isolation of *N.*
776 *cyriacigeorgica* environmental strain, Eva Gutbraut for her help with the electronic
777 microscopy image and Jordan Brun for his technical support in bacteriology experiment.

778 **Source of funding:** This work was partly funded by l'Agence Nationale de la Recherche
779 through ANR-17-CE04-0010 (Infiltron) project, by LabEx IMU (Intelligence des Mondes
780 Urbains), the Greater-Lyon Urban Community, the French national research program for
781 environmental and occupational health of Anses under the terms of project “Iouqmer” EST
782 2016/1/120, the School of Integrated Watershed Sciences H2O'LYON (ANR 17-EURE-
783 0018), and the Urban School of Lyon. Funders did not interfere in the elaboration of the
784 experiments or data analysis.

785 **Declaration of interests**

786
787 The authors declare that they have no known competing financial interests or personal
788 relationships that could have appeared to influence the work reported in this paper.

789
790 The authors declare the following financial interests/personal relationships which may be
791 considered as potential competing interests:
792

793
794

795 **References**

- 796 1. Brown-Elliott BA, Brown JM, Conville PS, Wallace RJ. 2006. Clinical and
797 Laboratory Features of the *Nocardia* spp. Based on Current Molecular Taxonomy. Clin
798 Microbiol Rev 19:259–282.
- 799 2. Rodriguez-Nava V, Durupt S, Chyderiotis S, Freydière A-M, Karsenty J, de Montclos
800 M, Reix P, Durieu I, Nove-Josserand R, Chiron R, Bremont F, Têtu L, Murriss M, Terru D,
801 Godreuil S, Bergeron E, Freney J, Boiron P, Vandenesch F, Marchandin H, Segonds C,
802 Doléans-Jordheim A. 2015. A French multicentric study and review of pulmonary *Nocardia*
803 spp. in cystic fibrosis patients. Med Microbiol Immunol (Berl) 204:493–504.
804 <https://doi.org/10.1007/s00430-014-0360-3>.
- 805 3. Heise ER. 1982. Diseases Associated with Immunosuppression. Environ Health
806 Perspect 11. <https://doi.org/10.1289/ehp.82439>.
- 807 4. Luo Q, Hiessl S, Poehlein A, Daniel R, Steinbüchel A. 2014. Insights into the
808 Microbial Degradation of Rubber and Gutta-Percha by Analysis of the Complete Genome of
809 *Nocardia nova* SH22a. Appl Environ Microbiol 80:3895–3907.
810 <https://doi.org/10.1128/AEM.00473-14>.
- 811 5. Luo Q, Hiessl S, Steinbüchel A. 2014. Functional diversity of *Nocardia* in
812 metabolism. Environ Microbiol 16:29–48. <https://doi.org/10.1111/1462-2920.12221>.
- 813 6. Quatrini P, Scaglione G, Pasquale CD, Riela S, Puglia AM. 2008. Isolation of Gram-
814 positive n-alkane degraders from a hydrocarbon-contaminated Mediterranean shoreline. J
815 Appl Microbiol 104:251–259. <https://doi.org/10.1111/j.1365-2672.2007.03544.x>
- 816 7. Nhi-Cong LT, Mikolasch A, Awe S, Sheikhany H, Klenk H-P, Schauer F. 2010.
817 Oxidation of aliphatic, branched chain, and aromatic hydrocarbons by *Nocardia*
818 *cyriacigeorgica* isolated from oil-polluted sand samples collected in the Saudi Arabian
819 Desert. J Basic Microbiol 50:241–253. <https://doi.org/10.1002/jobm.200900358>
- 820 8. Pujic P, Beaman BL, Ravalison M, Boiron P, Rodríguez-Nava V. 2015. Chapter 40 -
821 *Nocardia* and Actinomyces, p. 731–752. In Tang, Y-W, Sussman, M, Liu, D, Poxton, I,
822 Schwartzman, J (eds.), Molecular Medical Microbiology (Second Edition). Academic Press,
823 Boston. <https://doi.org/10.1016/B978-0-12-397169-2.00040-8>
- 824 9. Schlaberg R, Huard RC, Della-Latta P. 2008. *Nocardia cyriacigeorgica*, an Emerging
825 Pathogen in the United States. J Clin Microbiol 46:265–273.
826 <https://doi.org/10.1128/JCM.00937-07>
- 827 10. Valdezate S, Garrido N, Carrasco G, Medina-Pascual MJ, Villalón P, Navarro AM,
828 Saéz-Nieto JA. 2016. Epidemiology and susceptibility to antimicrobial agents of the main
829 *Nocardia* species in Spain. J Antimicrob Chemother dkw489.
830 <https://doi.org/10.1093/jac/dkw489>
- 831 11. Lebeaux D, Bergeron E, Berthet J, Djadi-Prat J, Mouniée D, Boiron P, Lortholary O,
832 Rodríguez-Nava V. 2018. Antibiotic susceptibility testing and species identification of

- 833 *Nocardia* isolates: a retrospective analysis of data from a French expert laboratory, 2010–
834 2015. Clin Microbiol Infect. <https://doi.org/10.1016/j.cmi.2018.06.013>
- 835 12. Bernardin-Souibgui C, Barraud S, Bourgeois E, Aubin J-B, Becouze-Lareure C, Wiest
836 L, Marjolet L, Colinon C, Lipeme Kouyi G, Cournoyer B, Blaha D. 2018. Incidence of
837 hydrological, chemical, and physical constraints on bacterial pathogens, *Nocardia* cells, and
838 fecal indicator bacteria trapped in an urban stormwater detention basin in Chassieu, France.
839 Environ Sci Pollut Res 25:24860–24881. <https://doi.org/10.1007/s11356-018-1994-2>
- 840 13. Rahmati M, Weihermüller L, Vanderborght J, Pachepsky YA, Mao L, Sadeghi SH,
841 Moosavi N, Kheirfam H, Montzka C, Van Looy K, Toth B, Hazbavi Z, Al Yamani W,
842 Albalasmeh AA, Alghzawi MZ, Angulo-Jaramillo R, Antonino ACD, Arampatzis G,
843 Armindo RA, Asadi H, Bamutaze Y, Batlle-Aguilar J, Béchet B, Becker F, Blöschl G, Bohne
844 K, Braud I, Castellano C, Cerdà A, Chalhoub M, Cichota R, Císlarová M, Clothier B, Coquet
845 Y, Cornelis W, Corradini C, Coutinho AP, de Oliveira MB, de Macedo JR, Durães MF,
846 Emami H, Eskandari I, Farajnia A, Flammini A, Fodor N, Gharaibeh M, Ghavimippanah MH,
847 Ghezzehei TA, Giertz S, Hatzigiannakis EG, Horn R, Jiménez JJ, Jacques D, Keesstra SD,
848 Kelishadi H, Kiani-Harchegani M, Kouselou M, Kumar Jha M, Lassabatere L, Li X, Liebig
849 MA, Lichner L, López MV, Machiwal D, Mallants D, Mallmann MS, de Oliveira Marques
850 JD, Marshall MR, Mertens J, Meunier F, Mohammadi MH, Mohanty BP, Pulido-Moncada M,
851 Montenegro S, Morbidelli R, Moret-Fernández D, Moosavi AA, Mosaddeghi MR, Mousavi
852 SB, Mozaffari H, Nabiollahi K, Neyshabouri MR, Ottoni MV, Ottoni Filho TB, Pahlavan-Rad
853 MR, Panagopoulos A, Peth S, Peyneau P-E, Picciafuoco T, Poesen J, Pulido M, Reinert DJ,
854 Reinsch S, Rezaei M, Roberts FP, Robinson D, Rodrigo-Comino J, Rotunno Filho OC, Saito
855 T, Sukanuma H, Saltalippi C, Sándor R, Schütt B, Seeger M, Sepehrnia N, Sharifi
856 Moghaddam E, Shukla M, Shutaro S, Sorando R, Stanley AA, Strauss P, Su Z, Taghizadeh-
857 Mehrjardi R, Taguas E, Teixeira WG, Vaezi AR, Vafakhah M, Vogel T, Vogelers I,
858 Votrubova J, Werner S, Winarski T, Yilmaz D, Young MH, Zacharias S, Zeng Y, Zhao Y,
859 Zhao H, Vereecken H. 2018. Development and analysis of the Soil Water Infiltration Global
860 database. Earth Syst Sci Data 10:1237–1263. <https://doi.org/10.5194/essd-10-1237-2018>
- 861 14. Sébastien C, Barraud S, Ribun S, Zoropogui A, Blaha D, Becouze-Lareure C, Kouyi
862 GL, Cournoyer B. 2014. Accumulated sediments in a detention basin: chemical and microbial
863 hazard assessment linked to hydrological processes. Environ Sci Pollut Res 21:5367–5378.
864 <https://doi.org/10.1007/s11356-013-2397-z>
- 865 15. Badin AL, Monier A, Volatier L, Geremia RA, Delolme C, Bedell J-P. 2011.
866 Structural Stability, Microbial Biomass and Community Composition of Sediments Affected
867 by the Hydric Dynamics of an Urban Stormwater Infiltration Basin: Dynamics of Physical
868 and Microbial Characteristics of Stormwater Sediment. Microb Ecol 61:885–897.
869 <https://doi.org/10.1007/s00248-011-9829-4>
- 870 16. Marti R, Becouze-Lareure C, Ribun S, Marjolet L, Bernardin Souibgui C, Aubin J-B,
871 Lipeme Kouyi G, Wiest L, Blaha D, Cournoyer B. 2017. Bacteriome genetic structures of
872 urban deposits are indicative of their origin and impacted by chemical pollutants.
873 <https://doi.org/10.1038/s41598-017-13594-8>
- 874 17. Kaushik R, Balasubramanian R, Cruz AA de la. 2012. Influence of Air Quality on the
875 Composition of Microbial Pathogens in Fresh Rainwater. Appl Environ Microbiol 78:2813–
876 2818. <https://doi.org/10.1128/AEM.07695-11>

- 877 18. Barraud S, Gibert J, Winiarski T, Krajewski J-LB. 2002. Implementation of a
878 monitoring system to measure impact of stormwater runoff infiltration. *Water Sci Technol*
879 45:203–210. <https://doi.org/10.2166/wst.2002.0080>
- 880 19. Winiarski T. 2014. Fonction filtration d'un ouvrage urbain - consequence sur la
881 formation d'un anthroposol. 2015. Available: [http://temis.documentation.developpement-](http://temis.documentation.developpement-durable.gouv.fr/docs/Temis/0081/Temis-0081850/21970_A.pdf)
882 [durable.gouv.fr/docs/Temis/0081/Temis-0081850/21970_A.pdf](http://temis.documentation.developpement-durable.gouv.fr/docs/Temis/0081/Temis-0081850/21970_A.pdf) 199 p [accessed 19 April
883 2019].
- 884 20. Li J, Shang X, Zhao Z, Tanguay RL, Dong Q, Huang C. 2010. Polycyclic aromatic
885 hydrocarbons in water, sediment, soil, and plants of the Aojiang River waterway in Wenzhou,
886 China. *J Hazard Mater* 173:75–81. <https://doi.org/10.1016/j.jhazmat.2009.08.050>
- 887 21. Muntean N, Muntean E, Duda MM. 2015. Polycyclic Aromatic Hydrocarbons in Soil.
888 ResearchGate.
- 889 22. Rodriguez-Nava V, Couble A, Devulder G, Flandrois J-P, Boiron P, Laurent F. 2006.
890 Use of PCR-Restriction Enzyme Pattern Analysis and Sequencing Database for *hsp65* Gene-
891 Based Identification of *Nocardia* Species. *J Clin Microbiol* 44:536–546.
892 <https://doi.org/10.1128/JCM.44.2.536-546.2006>.
- 893 23. Telenti A, Marchesi F, Balz M, Bally F, Böttger EC, Bodmer T. 1993. Rapid
894 identification of mycobacteria to the species level by polymerase chain reaction and
895 restriction enzyme analysis. *J Clin Microbiol* 31:175–178.
- 896 24. Sánchez-Herrera K, Sandoval H, Mouniee D, Ramírez-Durán N, Bergeron E, Boiron
897 P, Sánchez-Saucedo N, Rodríguez-Nava V. 2017. Molecular identification of *Nocardia*
898 species using the *sodA* gene. *New Microbes New Infect* 19:96–116.
899 <https://doi.org/10.1016/j.nmni.2017.03.008>
- 900 25. Conville PS, Zelazny AM, Witebsky FG. 2006. Analysis of *secA1* Gene Sequences for
901 Identification of *Nocardia* Species. *J Clin Microbiol* 44:2760–2766.
902 <https://doi.org/10.1128/JCM.00155-06>
- 903 26. Laurent FJ, Provost F, Boiron P. 1999. Rapid Identification of Clinically Relevant
904 *Nocardia* Species to Genus Level by 16S rRNA Gene PCR. *J Clin Microbiol* 37:99–102.
- 905 27. Verzani J. 2004. Using R for Introductory Statistics. Chapman and Hall/CRC.
906 <https://doi.org/10.4324/9780203499894>
- 907 28. Charif D, Thioulouse J, Lobry JR, Perri re G. 2005. Online synonymous codon usage
908 analyses with the *ade4* and *seqinR* packages. *Bioinformatics* 21:545–547.
909 <https://doi.org/10.1093/bioinformatics/bti037>
- 910 29. Lê Cao K-A, Boitard S, Besse P. 2011. Sparse PLS discriminant analysis: biologically
911 relevant feature selection and graphical displays for multiclass problems. *BMC*
912 *Bioinformatics* 12:253. <https://doi.org/10.1186/1471-2105-12-253>
- 913 30. Friendly M. 2002. Corrgrams: Exploratory Displays for Correlation Matrices. *Am Stat*
914 56:316–324. <https://doi.org/10.1198/000313002533>
- 915

- 916 31. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski
917 RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ,
918 Weber CF. 2009. Introducing mothur: Open-Source, Platform-Independent, Community-
919 Supported Software for Describing and Comparing Microbial Communities. *Appl Environ*
920 *Microbiol* 75:7537–7541. <https://doi.org/10.1128/AEM.01541-09>
- 921 32. Rodriguez-Nava V, Couble A, Devulder G, Flandrois J-P, Boiron P, Laurent F. 2006.
922 Use of PCR-Restriction Enzyme Pattern Analysis and Sequencing Database for *hsp65* Gene-
923 Based Identification of *Nocardia* Species. *J Clin Microbiol* 44:536–546.
924 <https://doi.org/10.1128/JCM.44.2.536-546.2006>
- 925 33. Beaman BL, Maslan S. 1977. Effect of cyclophosphamide on experimental *Nocardia*
926 *asteroides* infection in mice. *Infect Immun* 16:995–1004.
- 927 34. Yassin AF, Rainey FA, Steiner U. 2001. *Nocardia cyriacigeorgici* sp. nov.. *Int J Syst*
928 *Evol Microbiol* 51:1419–1423. <https://doi.org/10.1099/00207713-51-4-1419>
- 929 35. Gouy M, Guindon S, Gascuel O. 2010. SeaView Version 4: A Multiplatform
930 Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. *Mol Biol*
931 *Evol* 27:221–224. <https://doi.org/10.1093/molbev/msp259>
- 932 36. McTaggart LR, Richardson SE, Witkowska M, Zhang SX. 2010. Phylogeny and
933 Identification of *Nocardia* Species on the Basis of Multilocus Sequence Analysis. *J Clin*
934 *Microbiol* 48:4525–4533. <https://doi.org/10.1128/JCM.00883-10>
- 935 37. Rudramurthy SM, Kaur H, Samanta P, Ghosh A, Chakrabarti A, Honnavar P, Ray P.
936 2015. Molecular identification of clinical *Nocardia* isolates from India. *J Med Microbiol*
937 64:1216–1225. <https://doi.org/10.1099/jmm.0.000143>
- 938 38. Xiao M, Pang L, Chen SC-A, Fan X, Zhang L, Li H-X, Hou X, Cheng J-W, Kong F,
939 Zhao Y-P, Xu Y-C. 2016. Accurate Identification of Common Pathogenic *Nocardia* Species:
940 Evaluation of a Multilocus Sequence Analysis Platform and Matrix-Assisted Laser
941 Desorption Ionization-Time of Flight Mass Spectrometry. *PLOS ONE* 11:e0147487.
942 <https://doi.org/10.1371/journal.pone.0147487>
- 943 39. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics
944 Analysis Version 7.0 for Bigger Datasets. *Mol Biol Evol* 33:1870–1874.
945 <https://doi.org/10.1093/molbev/msw054>
- 946 40. Gilbert B, McDonald IR, Finch R, Stafford GP, Nielsen AK, Murrell JC. 2000.
947 Molecular Analysis of the *pmo* (Particulate Methane Monooxygenase) Operons from Two
948 Type II Methanotrophs. *Appl Environ Microbiol* 66:966–975.
949 <https://doi.org/10.1128/AEM.66.3.966-975.2000>
- 950 41. Vallenet D, Calteau A, Cruveiller S, Gachet M, Lajus A, Josso A, Mercier J, Renaux
951 A, Rollin J, Rouy Z, Roche D, Scarpelli C, Médigue C. 2017. MicroScope in 2017: an
952 expanding and evolving integrated resource for community expertise of microbial genomes.
953 *Nucleic Acids Res* 45:D517–D528. <https://doi.org/10.1093/nar/gkw1101>
- 954 42. Vautrin F, Bergeron E, Dubost A, Abrouk D, Martin C, Cournoyer B, Louzier V,

- 955 Winiarski T, Rodriguez-Nava V, Pujic P. 2019. Genome Sequences of Three *Nocardia*
956 *cyriacigeorgica* Strains and One *Nocardia asteroides* Strain. *Microbiol Resour Announc*
957 8:e00600-19, /mra/8/33/MRA.00600-19.atom. <https://doi.org/10.1128/MRA.00600-19>
- 958 43. Zoropogui A, Pujic P, Normand P, Barbe V, Belli P, Graindorge A, Roche D, Vallenet
959 D, Mangenot S, Boiron P, Rodriguez-Nava V, Ribun S, Richard Y, Cournoyer B, Blaha D.
960 2013. The *Nocardia cyriacigeorgica* GUH-2 genome shows ongoing adaptation of an
961 environmental Actinobacteria to a pathogen's lifestyle. *BMC Genomics* 14:286.
962 <https://doi.org/10.1186/1471-2164-14-286>
- 963 44. Restagno D, Venet F, Paquet C, Freyburger L, Allaouchiche B, Monneret G, Bonnet J-
964 M, Louzier V. 2016. Mice Survival and Plasmatic Cytokine Secretion in a "Two Hit" Model
965 of Sepsis Depend on Intratracheal *Pseudomonas Aeruginosa* Bacterial Load. *PLoS ONE* 11.
966 <https://doi.org/10.1371/journal.pone.0162109>
- 967 46. Nemzek JA, Xiao H-Y, Minard AE, Bolgos GL, Remick DG. 2004. Humane
968 endpoints in shock research. *Shock* Augusta Ga 21:17–25.
969 <https://doi.org/10.1097/01.shk.0000101667.49265.fd>
- 970 47. Feldman AT, Wolfe D. 2014. Tissue Processing and Hematoxylin and Eosin Staining,
971 p. 31–43. *In* *Histopathology*. Humana Press, New York, NY. [https://doi.org/10.1007/978-1-](https://doi.org/10.1007/978-1-4939-1050-2_3)
972 [4939-1050-2_3](https://doi.org/10.1007/978-1-4939-1050-2_3)
- 973 48. Wang Q, Garrity GM, Tiedje JM, Cole JR. 2007. Naïve Bayesian Classifier for Rapid
974 Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl Environ Microbiol*
975 73:5261–5267. <https://doi.org/10.1128/AEM.00062-07>
- 976 49. Cui Q, Fang T, Huang Y, Dong P, Wang H. 2017. Evaluation of bacterial pathogen
977 diversity, abundance and health risks in urban recreational water by amplicon next-generation
978 sequencing and quantitative PCR. *J Environ Sci* 57:137–149.
- 979 50. Chulan F, Yanpeng L, Pengxia L, Feifei M, Zhengsheng X, Rui L, Yuzhen Q, Beibei
980 W, Cheng J. 2019. Characteristics of airborne opportunistic pathogenic bacteria during
981 autumn and winter in Xi'an, China. *Sci Total Environ*
982 <https://doi.org/10.1016/j.scitotenv.2019.03.412>. <https://doi.org/10.1016/j.jes.2016.11.008>
- 983 51. Arrache D, Zait H, Rodriguez-Nava V, Bergeron E, Durand T, Yahiaoui M,
984 Grenouillet F, Amrane A, Chaouche F, Baiod A, Madani K, Hamrioui B. 2018. Nocardiose
985 cérébrale et pulmonaire à *Nocardia abscessus* chez un patient algérien immunocompétent. *J*
986 *Mycol Médicale* 28:531–537. <https://doi.org/10.1016/j.mycmed.2018.04.010>
- 987 52. Garcia-Bellmunt L, Sibila O, Solanes I, Sanchez-Reus F, Plaza V. 2012. Pulmonary
988 Nocardiosis in Patients With COPD: Characteristics and Prognostic Factors. *Arch*
989 *Bronconeumol Engl Ed* 48:280–285. <https://doi.org/10.1016/j.arbr.2012.06.006>
- 990 53. Steinbrink J, Leavens J, Kauffman CA, Miceli MH. 2018. Manifestations and
991 outcomes of nocardia infections: Comparison of immunocompromised and
992 nonimmunocompromised adult patients. *Medicine (Baltimore)* 97:e12436.
993 <https://doi.org/10.1097/MD.00000000000012436>

- 994 54. Eshraghi SS, Heidarzadeh S, Soodbakhsh A, Pourmand M, Ghasemi A, GramiShoar
995 M, Zibafar E, Aliramezani A. 2014. Pulmonary nocardiosis associated with cerebral abscess
996 successfully treated by co-trimoxazole: a case report. *Folia Microbiol (Praha)* 59:277–281.
997 <https://doi.org/10.1007/s12223-013-0298-7>
- 998 55. Budzinski H, Garrigues Ph, Connan J, Devillers J, Domine D, Radke M, Oudins JL.
999 1995. Alkylated phenanthrene distributions as maturity and origin indicators in crude oils and
1000 rock extracts. *Geochim Cosmochim Acta* 59:2043–2056. [https://doi.org/10.1016/0016-7037\(95\)00125-5](https://doi.org/10.1016/0016-7037(95)00125-5)
- 1002 56. Yunker MB, Macdonald RW, Vingarzan R, Mitchell RH, Goyette D, Sylvestre S.
1003 2002. PAHs in the Fraser River basin: a critical appraisal of PAH ratios as indicators of PAH
1004 source and composition. *Org Geochem* 33:489–515. [https://doi.org/10.1016/S0146-6380\(02\)00002-5](https://doi.org/10.1016/S0146-6380(02)00002-5)
- 1006 57. Nádudvari Á, Fabiańska MJ. 2015. Coal-related sources of organic contamination in
1007 sediments and water from the Bierawka River (Poland). *Int J Coal Geol* 152:94–109.
1008 <https://doi.org/10.1016/j.coal.2015.11.006>
- 1009 58. Radke M, Welte DH, Willsch H. 1986. Maturity parameters based on aromatic
1010 hydrocarbons: Influence of the organic matter type. *Org Geochem* 10:51–63.
1011 [https://doi.org/10.1016/0146-6380\(86\)90008-2](https://doi.org/10.1016/0146-6380(86)90008-2)
- 1012 59. Sébastien C, Barraud S, Ribun S, Zoropogui A, Blaha D, Becouze-Lareure C, Kouyi
1013 GL, Cournoyer B. 2014. Accumulated sediments in a detention basin: chemical and microbial
1014 hazard assessment linked to hydrological processes. *Environ Sci Pollut Res Int* 21:5367–
1015 5378. <https://doi.org/10.1007/s11356-013-2397-z>
- 1016 60. Moran NA. 2002. Microbial Minimalism: Genome Reduction in Bacterial Pathogens.
1017 *Cell* 108:583–586. [https://doi.org/10.1016/S0092-8674\(02\)00665-7](https://doi.org/10.1016/S0092-8674(02)00665-7)
- 1018 61. McMahon MD, Rush JS, Thomas MG. 2012. Analyses of MbtB, MbtE, and MbtF
1019 Suggest Revisions to the Mycobactin Biosynthesis Pathway in *Mycobacterium tuberculosis*. *J*
1020 *Bacteriol* 194:2809–2818. <https://doi.org/10.1128/JB.00088-12>
- 1021 62. Li F, Feng L, Jin C, Wu X, Fan L, Xiong S, Dong Y. 2018. LpqT improves
1022 mycobacteria survival in macrophages by inhibiting TLR2 mediated inflammatory cytokine
1023 expression and cell apoptosis. *Tuberculosis* 111:57–66.
1024 <https://doi.org/10.1016/j.tube.2018.05.007>
- 1025 63. Cruz-Rabadán JS, Miranda-Ríos J, Espín-Ocampo G, Méndez-Tovar LJ, Maya-Pineda
1026 HR, Hernández-Hernández F. 2017. Non-Coding RNAs are Differentially Expressed by
1027 *Nocardia brasiliensis* in Vitro and in Experimental Actinomycetoma. *Open Microbiol J*
1028 11:112–125. <https://doi.org/10.2174/1874285801711010112>
- 1029 65. Singh I, West FM, Sanders A, Hartman B, Zappetti D. 2015. Pulmonary Nocardiosis
1030 in the Immunocompetent Host: Case Series. *Case Rep Pulmonol* 2015:1–6.
1031 <https://doi.org/10.1155/2015/314831>
- 1032

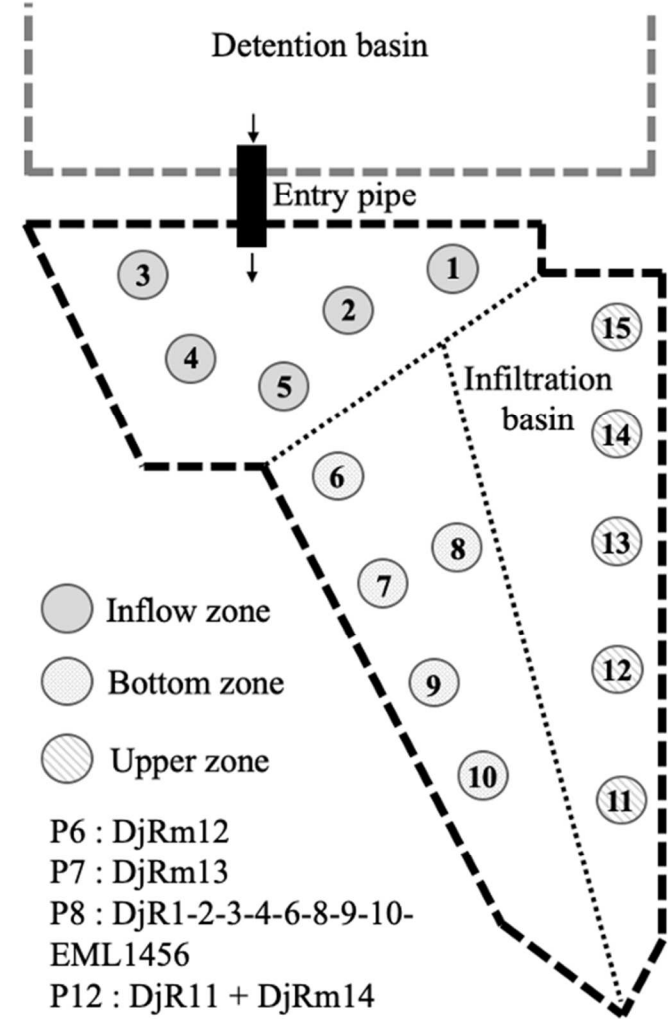
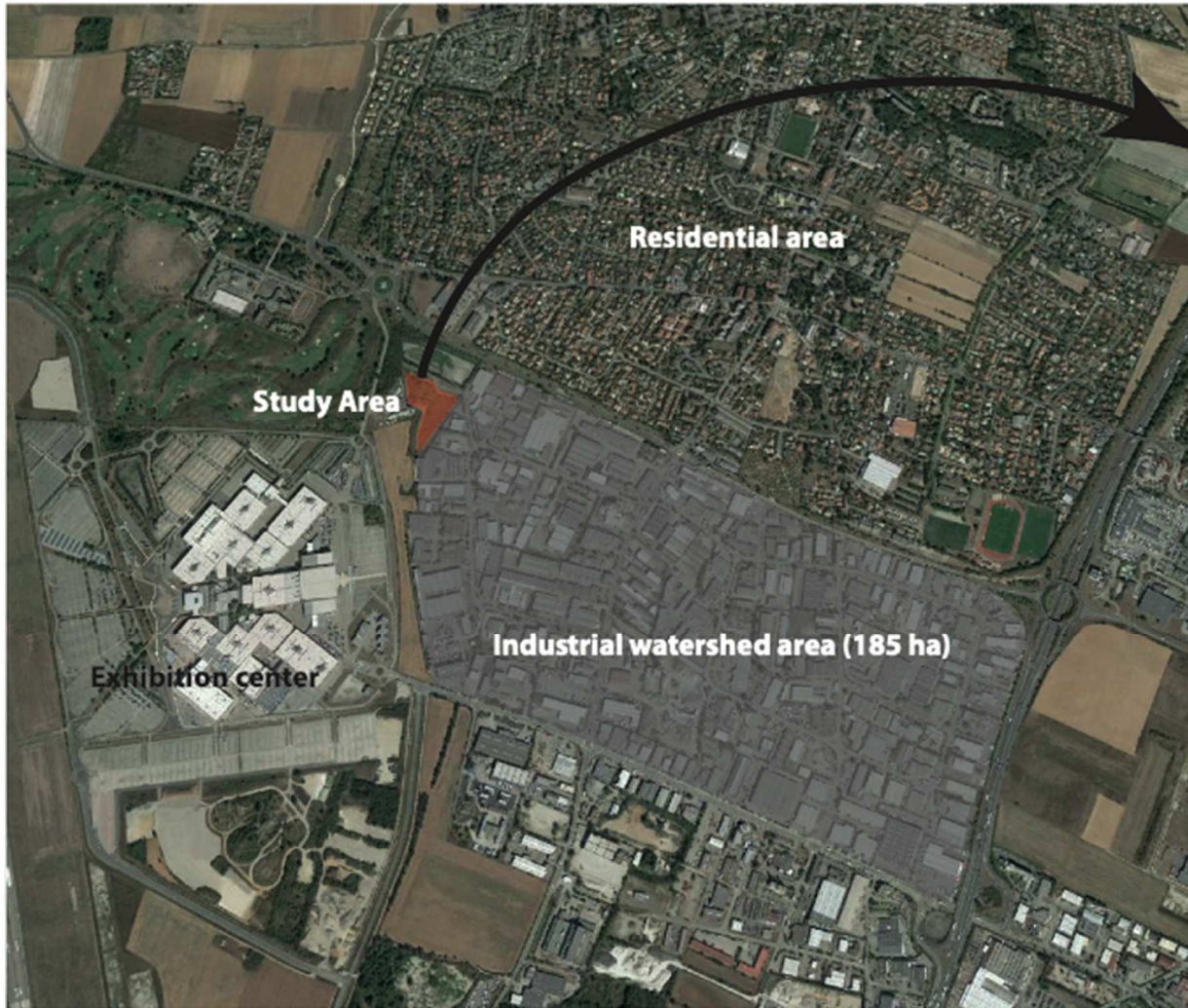


Figure 1. Aerial image of the Django-Reinhardt infiltration basin (DRIB) and position of the sampling points in the DRIB and placement of the three different sampling areas (inflow zone, bottom zone and upper zone). Px: DRIB sample point in which *N. cyriacigeorgica* was isolated and respective reference code.

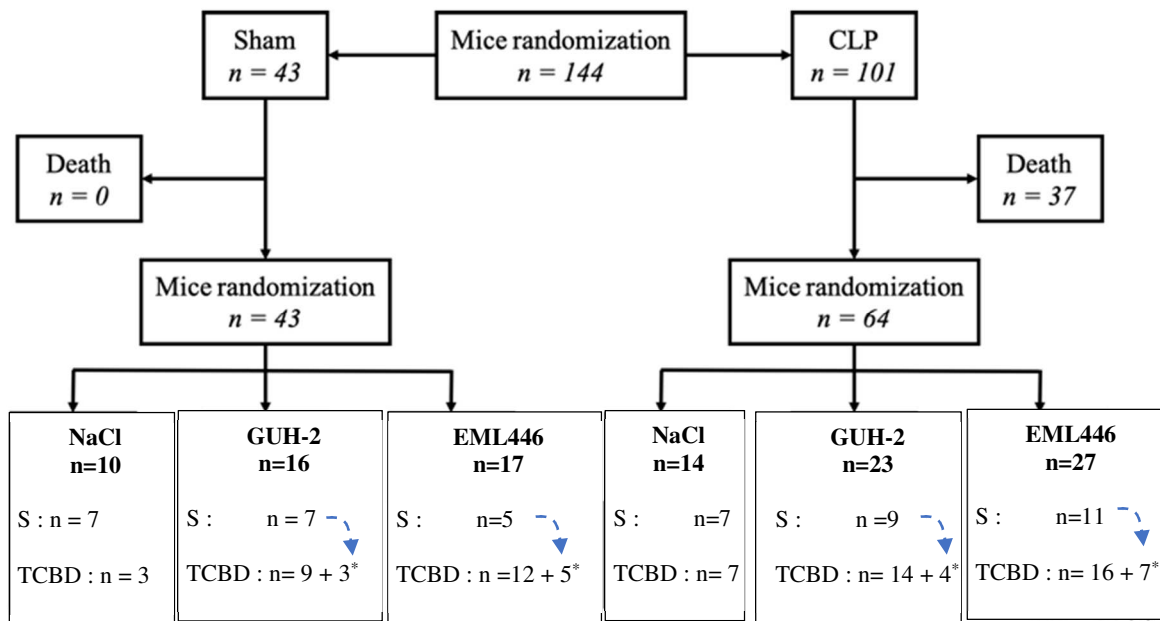


Figure 2A. Study design describing the experiment made with the immunosuppressed CLP (cecal ligation and puncture) (30%) murine model to compare the virulence of the SIS (stormwater infiltration system) and clinical *N. cyriacigeorgica* representative strains. Sham = mice that underwent laparotomy with exposition of the cecum but without CLP. CLP = intestinal tract externalization and puncture performed as described by Restagno et al. (40). S= Survival experiment. TCBD = Time Course Bacterial Detection experiment. n = number of individuals in each group. * = surviving mice at the end of the ‘S’ experiment retrieved for ‘TCBD’ one at D33 (GUH-2) or D41 (EML446).

Group	Day	Mice number for TCBD
	4	5
Sham GUH-2	10	3
	33	1 + 3*
	4	5
Sham EML446	10	5
	33	2
	41	5 ^{*, a)}
	4	5
CLP GUH-2	10	5
	33	4 + 4*
	4	7 ^{b)}
CLP EML446	10	4
	33	5
	41	7 ^{*, c)}

Figure 2B: Mice distribution for the Time Course Bacterial Detection experiment (TCBD). *

= Surviving mice from the end of the Survival experiment retrieved at the end of the TCBD

one. a) lungs of two mice will be reserved for histological analysis, b) lungs of one mouse

will be reserved for histological analysis. c) lungs of three mice and brains of two mice will

be reserved for histological analysis.

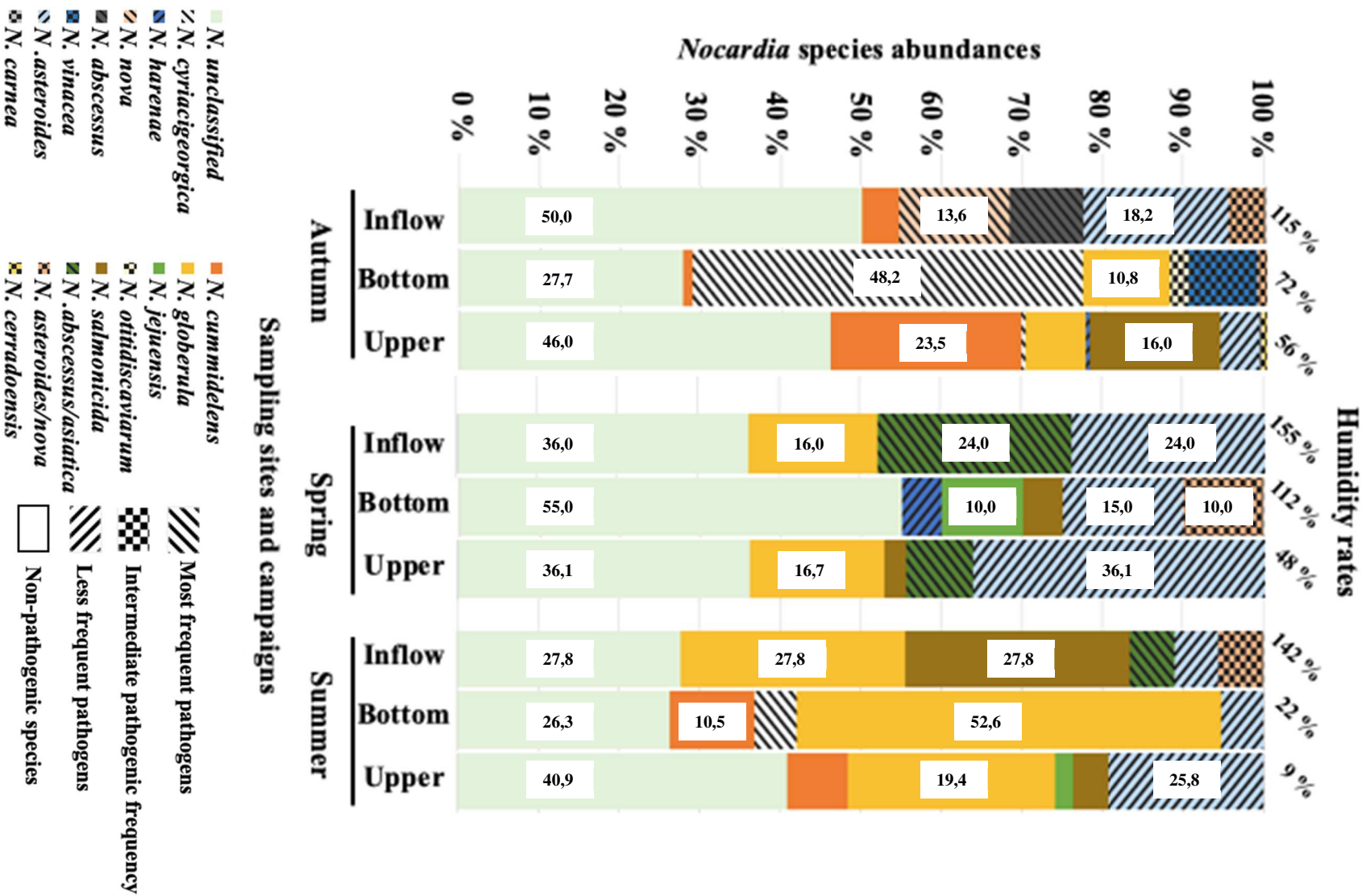


Figure 3. DNA metabarcoding analysis of *Nocardia* genetic diversity among SIS (stormwater infiltration system) sediment samples using the *hsp65* target. A total of 7,128 sequences per sample was analyzed. The relative proportion of the detected *Nocardia* species is presented in %. Species representing more than 1% over the total number of reads are indicated and only values > 10% are indicated. The classification of most/intermediate/less frequent pathogens was based on the most clinically relevant species in France epidemiology according to Lebeaux et al. (43).

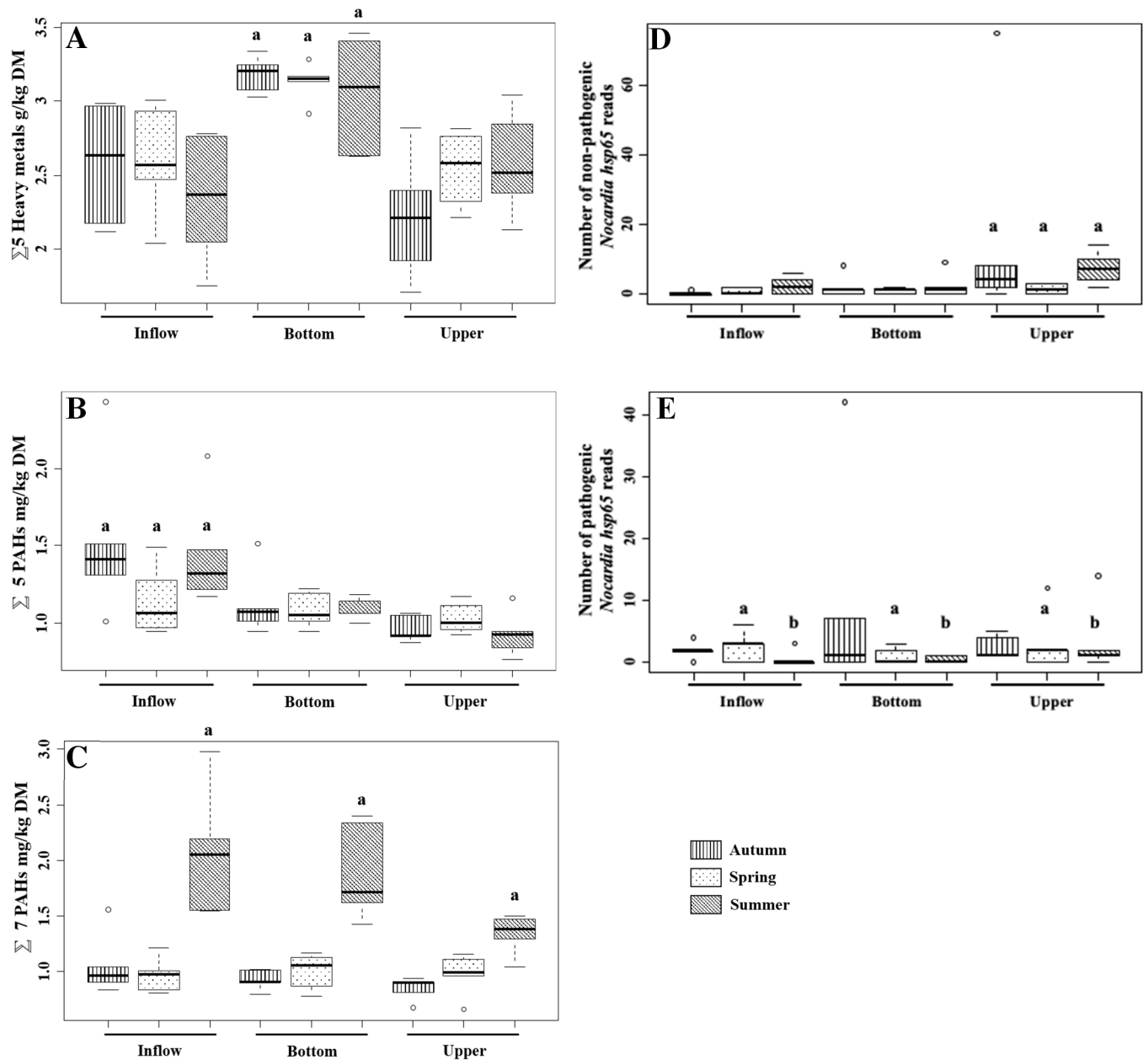


Figure 4. Boxplots explicative of the between-class analysis (BCA). **A.** 5 trace element metals: Cd (Cadmium); Cu (Copper); Hg (Mercury); Pb (Lead) and Zn (Zinc). P-value a = 0.0195. **B.** 7 PAHs (polycyclic aromatic hydrocarbons): naphthalene, acenaphthene, fluorene, benzo(b)fluoranthene, dibenzo(a,h)anthracene, benzo(ghi)perylene, indeno(1,2,3-cd)pyrene. P-value a = 6.97×10^{-6} . **C.** 5 PAHs: fluoranthene, pyrene, phenanthrene, benzo(a)anthracene,

benzo(a)pyrene. P-value a = 0.0157. **D.** Non-pathogenic *Nocardia* species: *N. cummidelens*, *N. globerula*, *N. harenae*, *N. iowensis*, *N. jejuensis*, *N. pseudovaccinii*, *N. salmonicida*, *N. soli*. P-value a = 0.0294. **E.** Pathogenic *Nocardia* species: *N. abscessus*, *N. abscessus/asiatica*, *N. anaemiae*, *N. asteroides*, *N. brasiliensis*, *N. carnea*, *N. cerradoensis*, *N. cyriacigeorgica*, *N. ninae*, *N. nova*, *N. otitiscaviarum*, *N. shimofusensis*, *N. sienata*, *N. vinaceae*. P-values a = 0.04055 and b = 0.02963.

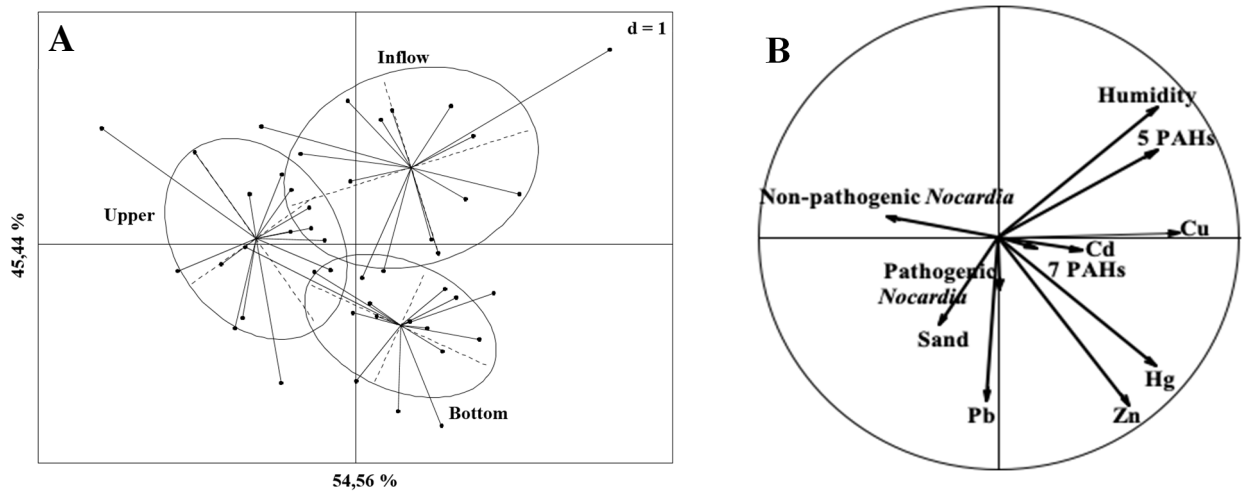


Figure 5 A. Between-Class Analysis (BCA) performed on physical-chemical parameters and *Nocardia* relative abundance from sediments of the Django-Reinhardt infiltration basin (DRIB). **B.** Explicative factors of the BCA. 7 PAHs (polycyclic aromatic hydrocarbons): naphthalene, acenaphthene, fluorene, benzo(b)fluoranthene, dibenzo(a,h)anthracene, benzo(ghi)perylene, indeno(1,2,3-cd)pyrene ; 5 PAHs: fluoranthene, pyrene, phenanthrene, benzo(a)anthracene, benzo(a)pyrene; Cd (Cadmium); Cu (Copper); Hg (Mercury); Pb (Lead) and Zn (Zinc). Pathogenic *Nocardia* species: *N. abscessus*, *N. abscessus/asiatica*, *N. anaemiae*, *N. asteroides*, *N. brasiliensis*, *N. carnea*, *N. cerradoensis*, *N. cyriacigeorgica*, *N. ninae*, *N. nova*, *N. otitiscaviarum*, *N. shimofusensis*, *N. sienata*, *N. vinaceae* ; non-pathogenic *Nocardia* species: *N. cummidelens*, *N. globerula*, *N. harenae*, *N. iowensis*, *N. jejuensis*, *N. pseudovaccinii*, *N. salmonicida*, *N. soli*.

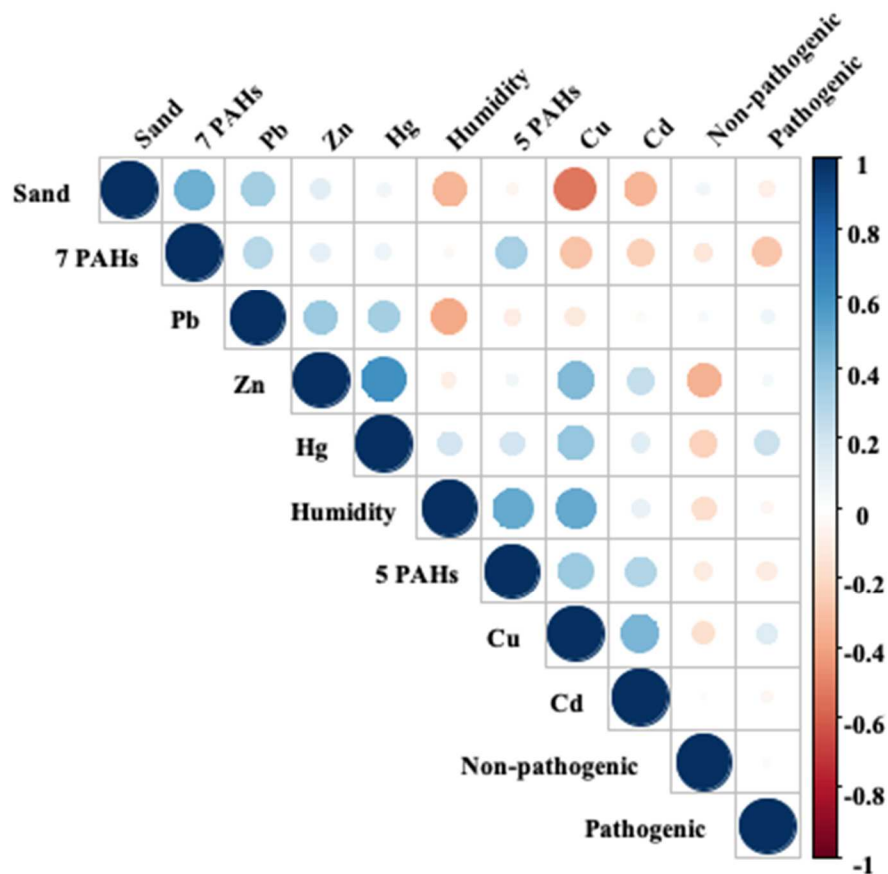
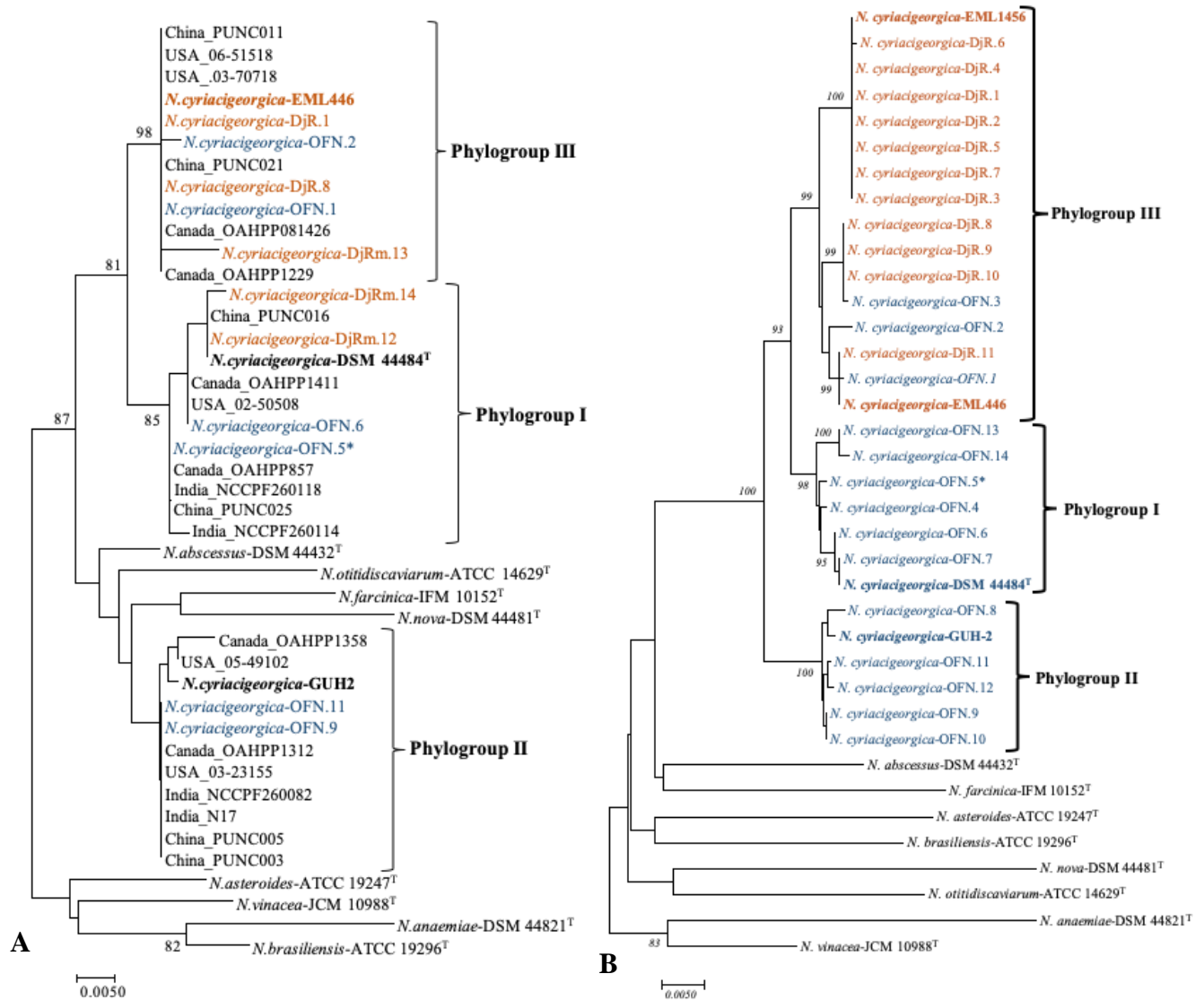


Figure 6. Correlogram of the physical-chemical parameters and *Nocardia* pathogenic and non-pathogenic species relative abundance. 7 PAHs (polycyclic aromatic hydrocarbons): naphthalene, acenaphthene, fluorene, benzo(b)fluoranthene, dibenzo(a,h)anthracene, benzo(ghi)perylene, indeno(1,2,3-cd)pyrene ; 5 PAHs: fluoranthene, pyrene, phenanthrene, benzo(a)anthracene, benzo(a)pyrene; trace element metals: Cd (Cadmium); Cu (Copper); Hg (Mercury); Pb (Lead) and Zn (Zinc).

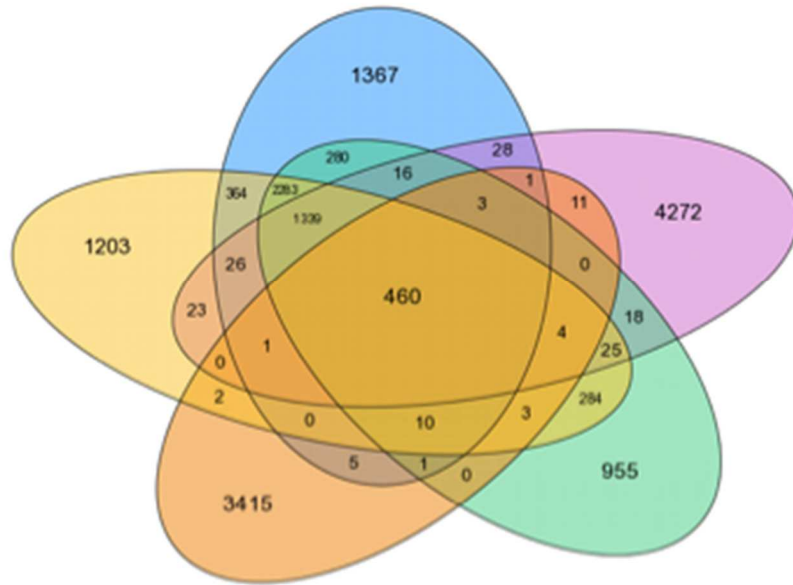







1

2 **Figure 7.** Molecular phylogeny of *Nocardia cyriacigeorgica*. **A)** Unrooted phylogenetic tree
3 based on the *hsp65* sequences (401 bp) from USA (Schlaberg et al., 2008), Canada
4 (MacTaggart et al., 2010), India (Rudramurthy et al., 2015) and China (Xiao et al., 2016)
5 isolates, and representative *N. cyriacigeorgica* sequences from the metabarcoding analysis,
6 and the clinical and SIS strains reported in this study. OFN.5* corresponds to a patient from
7 Lyon (France). DjR.1, DjR.8 and EML446 corresponding *hsp65* sequences were used to
8 represent the environmental cluster. **B)** Unrooted phylogenetic tree based on the *rrs-hsp65-*
9 *sodA-secA1* concatenated sequences (1,845 bp). The maximum likelihood tree was

10 constructed using MEGA software (version 7.0.16) after having aligned the sequences with
11 ClustalW. The bootstrap values were calculated from 1,000 replicates, and those higher than
12 80% are given at the corresponding nodes. Clinical strains are colored in blue, and
13 environmental strains are colored in orange

a



-  *N. cyriacigeorgica* DSM 44484^T
-  *N. cyriacigeorgica* EML446
-  *N. cyriacigeorgica* GUH-2
-  *N. farcinica* IFM 10152^T
-  *M. tuberculosis* H37Rv

b

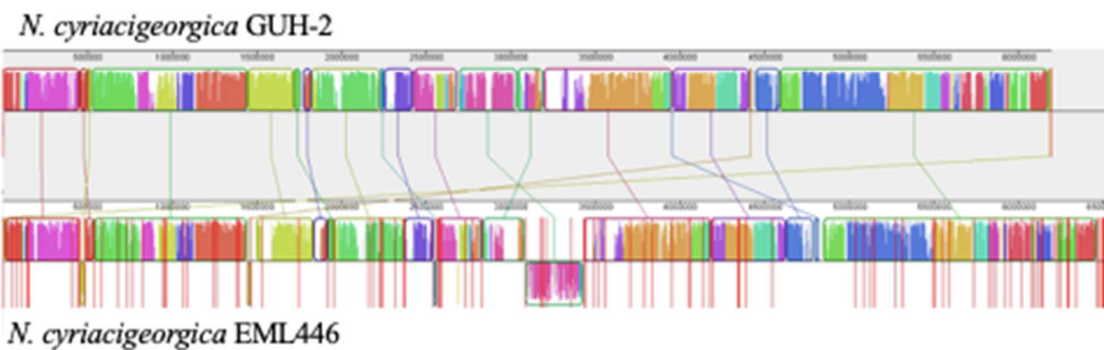


Figure 8. Comparison of *Nocardia cyriacigeorgica* EML446 genome with other actinomycetal reference genomes. (a) Venn Diagram representing the number of shared CDSs between whole genomes of *N. cyriacigeorgica* DSM44484^T, *N. cyriacigeorgica* EML446, *N.*

cyriacigeorgica GUH-2, *N. farcinica* IFM10152^T and *Mycobacterium tuberculosis* H37Rv.

(b) MAUVE comparison between EML446 and GUH-2 genomes showing the position of RGP (regions of genomic plasticity).

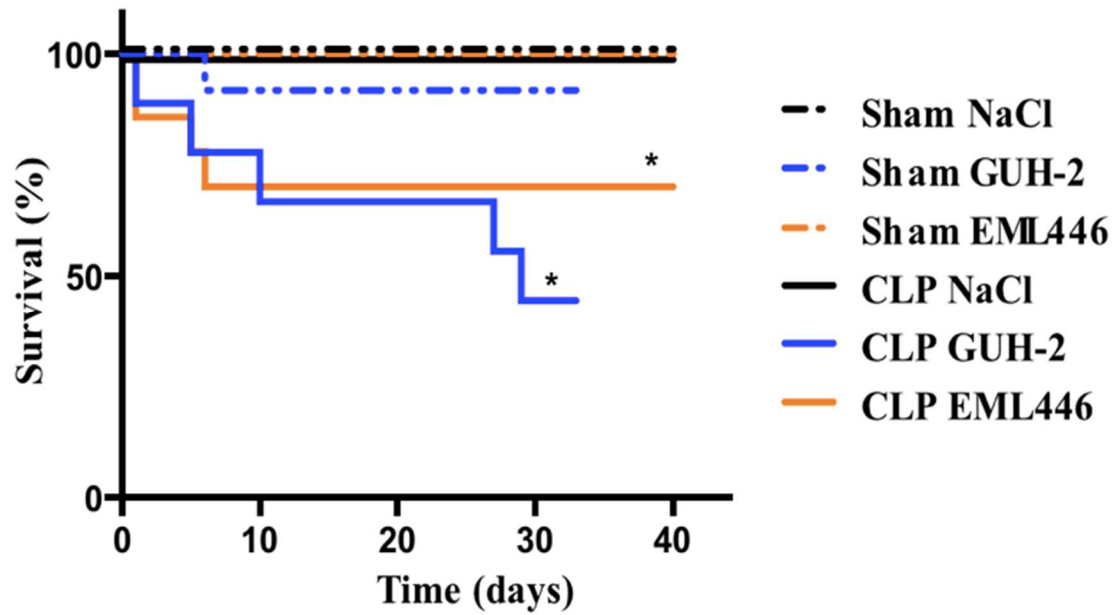
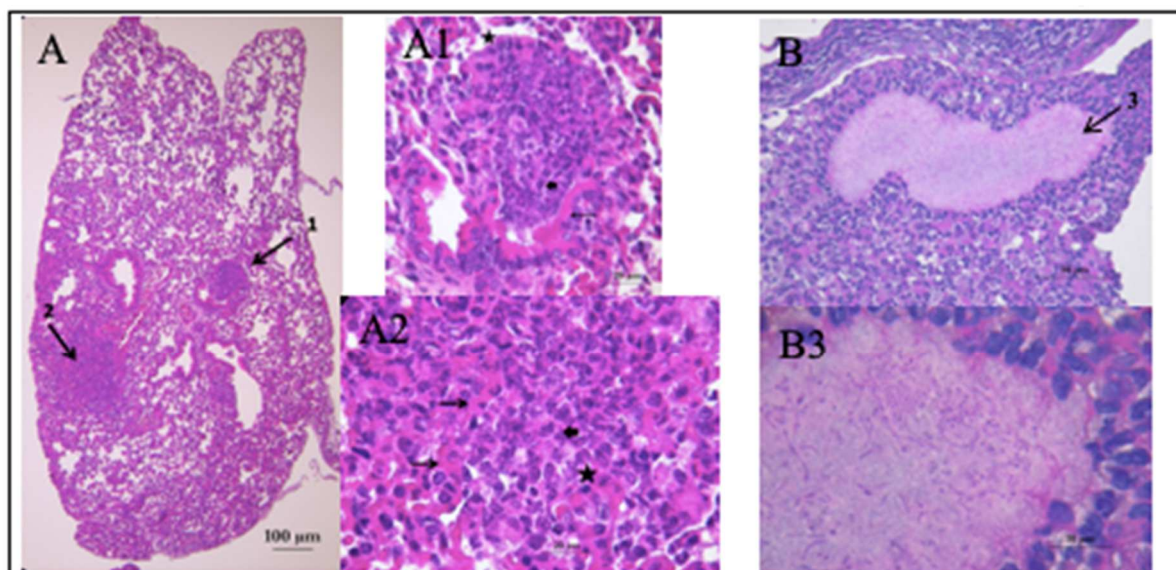
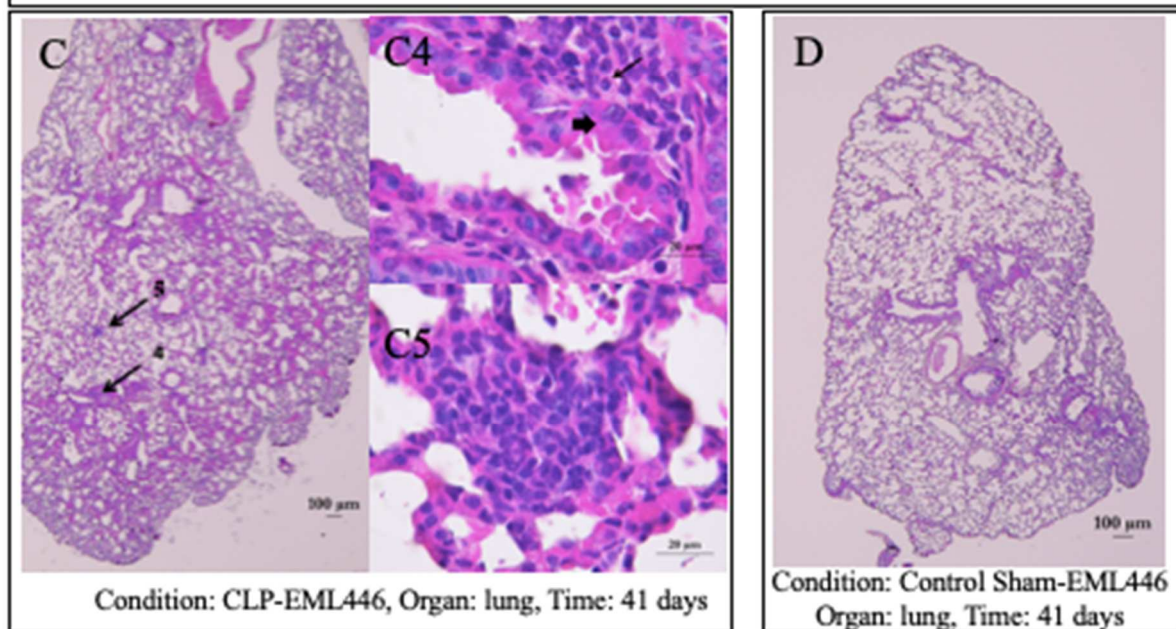


Figure 9. Survival rate after *N. cyriacigeorgica* cells instillations of Sham and CLP mice. Five days post-CLP (*i.e.* D0), mice were challenged with an intratracheal administration of *Nocardia* GUH-2 or EML446 at 1.0×10^6 CFU/mouse. Either NaCl (physiological saline solution) (Sham n=7, CLP n=7), GUH-2 (Sham n=7, CLP n=9) or EML446 (Sham n=5, CLP n=11) were instilled. Results are expressed as Kaplan-Meier survival curves. * p-value < 0.05 was considered statistically significant compared to the respective control groups.

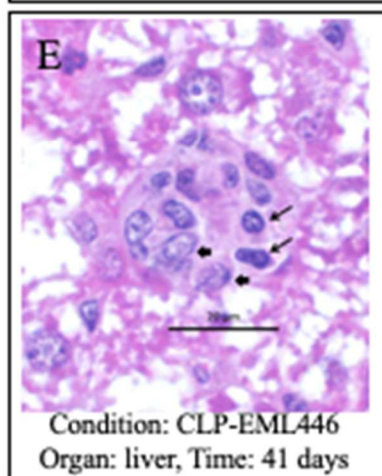


Condition: CLP-EML446, Organ: lung, Time: 4 days



Condition: CLP-EML446, Organ: lung, Time: 41 days

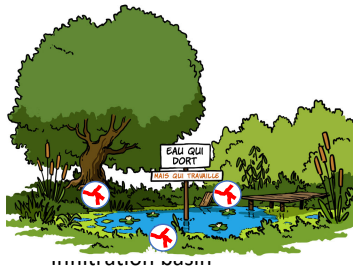
Condition: Control Sham-EML446
Organ: lung, Time: 41 days



Condition: CLP-EML446
Organ: liver, Time: 41 days

Figure 10. Histological lung and liver sections of CLP-EML446 mouse at days 4 and 41 after intratracheal instillation. **A, B: CLP-EML446 mouse lung at day 4.** Light micrographs of mouse lung section with evidence of granulomatous process characterized by both inflammatory response (**A**) and necrosis usually present in nocardiosis (**B**). **A:** Two early pyogranulomas disseminated in the pulmonary parenchyma (arrows). Original magnification: obj. 4 X, lungs, postcaval lobe. **A1:** enlargement of the pyogranulomas (arrow 1). Multiple degenerated neutrophilic polymorphonuclear cells into a lung alveolus lumen admixed with few macrophages. Some inflammatory cells overflow in the lumen of a bronchus. Thin arrow: bronchus epithelium. Large arrow: neutrophilic polymorphonuclear cells. Star: macrophages. Original magnification: obj. 40 X. **A2:** enlargement of the pyogranulomas (arrow 2). Multiple degenerated neutrophilic polymorphonuclear cells into a lung alveolus lumen admixed with few macrophages. Thin arrows: alveolar wall. Large arrow: neutrophilic polymorphonuclear cells. Star: macrophages. original magnification: obj. 40 X. **B:** The lungs showed cavitory lesions constituted by central necrotic material surrounded by some polymorphonuclear and numerous macrophages. Macrophage alveolitis and interstitial lymphocytic infiltration are also observed. Light micrographs; hematoxylin–eosin staining; original magnification: obj. 40 X. **B3:** High power magnification showing histologic sections of the lesion characterized by caseous necrotic central area with filamentous bacteria inside. Light micrographs; hematoxylin–eosin staining; original magnification: obj. 100 X. **C: CLP-EML446 mouse lung at day 41:** **C:** At day 41, some lymphocyte aggregates are present in the interstitium (arrows). Original magnification: obj. 4 X, lungs, postcaval lobe. **C4:** enlargement of the aggregate (arrow 4). Some small lymphocytes (thin arrow) infiltrate the interstitium beneath the bronchial epithelium (large arrow). Original magnification: obj. 100 X. **C5:** enlargement of the aggregate (arrow 5). Some small lymphocytes infiltrate the interstitium of the interalveolar wall. Original magnification: obj. 100 X. **D: Sham EML446 mouse lung at day**

41. no lesions. Original magnification obj. 4 X, lungs, postcaval lobe. **E: CLP-EML446**
mouse liver at day 41. Mature granuloma disseminated at random into liver lobules. Thin
arrows: lymphocytes. Large arrows: macrophages. Original magnification obj. 100 X.



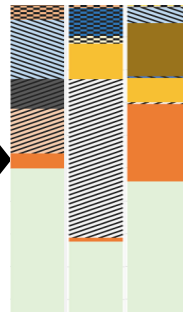
Question 1: Could anthropized urban environments host virulent strains of *Nocardia*?

Environmental *Nocardia cyriacigeorgica* strain EML446

versus

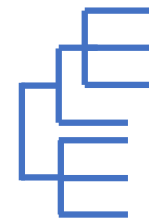
Clinical *Nocardia cyriacigeorgica* strain GUH-2

Diversity



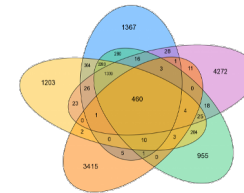
Spatio-temporal *Nocardia* diversity

Phylogeny



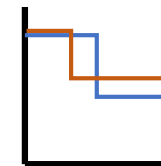
rrs-hsp65-sodA-secA1

Comparative genomics

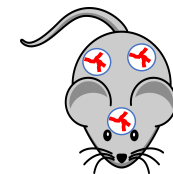


Venn diagram

Pathophysiology

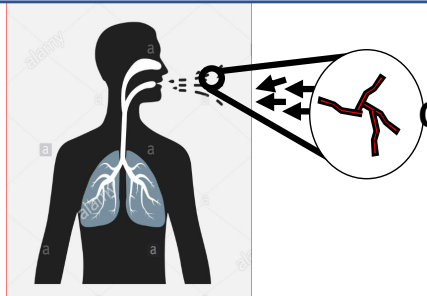


survival



spread

Comparative characteristics for clinical and environmental isolates



Question 2: Which are the potential sources of inhaled *Nocardia*?

Statement: *Nocardia cyriacigeorgica* is part of the most implicated *Nocardia* species in the worldwide nocardiosis epidemiology : N°1 in Spain (Valdezate et al., 2016) and Italy (Mazzaferri et al., 2018), N°2 in the USA (Schlaberg et al., 2014) and Australia (Paige & Spelman, 2019) or N°4 in France (Lebeaux et al., 2018)