

Supplementary Material

Phylogenetic analyses of the bacterial parasites detected in *A. terrestris* populations in Franche-Comte, France, collected in 2014-2018. Bacterial parasites were detected using high-throughput sequencing of the 16S rRNA bacterial gene extracted from organ
5 tissues of the hosts. OTU formation was conducted using the Swarm algorithm [1]; two clustering iterations are preformed - the first to denoise the data, and the second using the first iteration as seed sequences for grouping. Taxonomic assignment was conducted using blastn+ and the SILVA 16S database [2]. Sequence alignments were performed using the web implementation of Clustal Omega with default parameters
10 [3]. The following comparisons to reference sequences were performed using the seed sequences from the first clustering iteration.

Contents

1	Anaplasma	2
2	Bartonella	3
15	3 Bordetella	10
4	Borelliella	13
5	Filobacterium	14
6	Leptospira	15
7	Mycobacterium	20
20	8 Mycoplasma and Ureaplasma	22
9	Pasteurellaceae	28
10	Rickettsia	33
11	Yersinia	34
12	GenBank Accession Numbers	35
25	13 R Packages	36

1 Anaplasma

Anaplasma-001, the only OTU in this data set assigned to the *Anaplasma* genus, is identical to V4 16S rRNA reference sequences of *Anaplasma phagocytophilum* (formerly *Ehrlichia phagocytophilum*) and several unspecified *Ehrlichia* sp. sequences.

Table S1: Similarity, expressed as a percentage, and coverage (in parentheses) of the OTU Anaplasma-001 compared to a selection of reference strains of *Anaplasma phagocytophilum* and unspecified *Ehrlichia*. Sequences were compared using the blastn search tool.

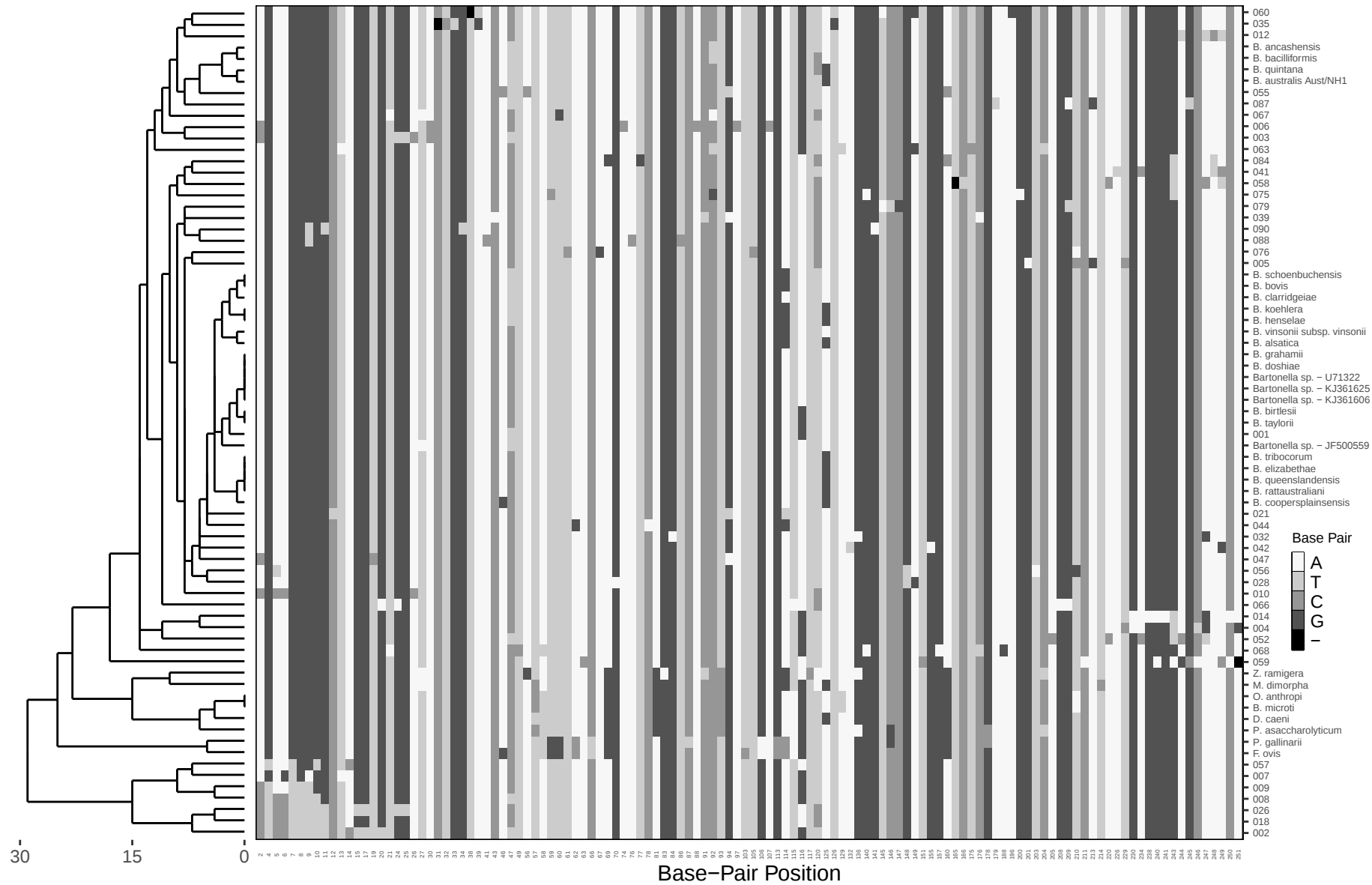
Reference Sequence	Similarity % (Coverage %)
Ehrlichia sp. - U10873	100 (100)
A. phagocytophilum - U02521	100 (100)
Ehrlichia sp. - U77389	100 (100)
Ehrlichia sp. - AJ242784	100 (100)
Ehrlichia sp. - AJ242783	100 (100)
A. phagocytophilum - HM439430	100 (100)
A. phagocytophilum str. JM - CP006617	100 (100)
A. phagocytophilum str. HZ2 - CP006616	100 (100)
Ehrlichia sp. - AF241532	100 (100)
A. phagocytophilum - GQ412339	100 (100)
A. phagocytophilum str. HZ - CP000235	100 (100)

30 2 **Bartonella**

The most prevalent of the Bartonella OTUs, Bartonella-001 (global prevalence of 43%), differed from reference sequences for *B. ancashensis*, *B. australis* Aust/NH1, *B. birtlesii*, *B. clarridgeiae*, and *B. taylorii* by one base pair (<1%). The remaining Bartonella OTU's differed from reference *Bartonella* sequences by as much as 1-6 % (2-14 base pairs) and
35 did not form clear affinities to previously-described *Bartonella* species (Figures S1 and S2).

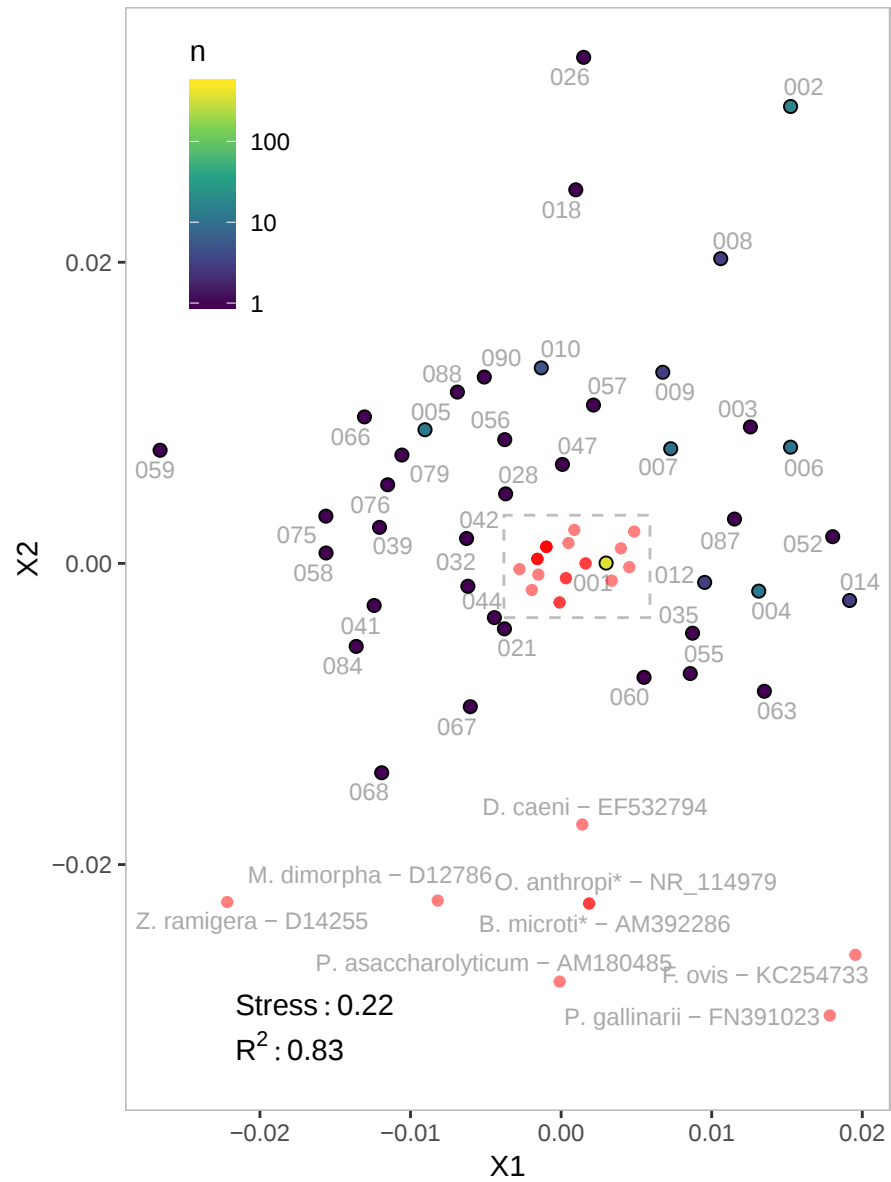
Table S2: Accession numbers, species, host and country of isolation for reference strains of *Bartonella* sp. used in comparisons.

Species - GenBank Accession	Isolation Source	Country
<i>Bartonella alsatica</i> - AJ002139	<i>Oryctolagus cuniculus</i>	France
<i>Bartonella ancashensis</i> CP010401	<i>Homo sapiens</i>	Peru
<i>Bartonella australis</i> Aust/NH1 - DQ538394	<i>Macropus giganteus</i>	Australia
<i>Bartonella bacilliformis</i> - M65249		
<i>Bartonella birtlesii</i> - AF204274	<i>Apodemus</i> sp.	France
<i>Bartonella bovis</i> - AF293391	<i>Bos</i> sp.	France
<i>Bartonella clarridgeiae</i> - X89208	<i>Felis catus</i>	
<i>Bartonella coopersplainsensis</i> - EU111759	<i>Rattus leucopus</i>	Australia
<i>Bartonella doshiae</i> - Z31351	<i>Microtus agrestis</i>	UK
<i>Bartonella elizabethae</i> - L01260	<i>Homo sapiens</i>	USA
<i>Bartonella grahamii</i> - Z31349	<i>Myodes glareolus</i>	UK
<i>Bartonella grahamii</i> - CP001562	<i>Apodemus sylvaticus</i>	Sweden
<i>Bartonella henselae</i> - BX897699		
<i>Bartonella koehlera</i> - AF076237	<i>Felis catus</i>	USA
<i>Bartonella queenslandensis</i> - EU111754	<i>Melomys</i> sp.	Australia
<i>Bartonella quintana</i> - M73228	<i>Homo sapiens</i>	USA
<i>Bartonella rattaaustraliani</i> - EU111749	<i>Rattus tunneyi</i>	Australia
<i>Bartonella schoenbuchensis</i> - AJ278187	<i>Capreolus capreolus</i>	Germany
<i>Bartonella</i> sp. - U71322	<i>Peromyscus leucopus</i>	USA
<i>Bartonella</i> sp. - JF500559	<i>Miniopterus schreibersii</i>	Taiwan
<i>Bartonella</i> sp. AA131HXZ - KJ361606	<i>Apodemus agrarius</i>	China
<i>Bartonella</i> sp. CR93HXZ - KJ361625	<i>Myodes rutilus</i>	China
<i>Bartonella taylorii</i> - Z31350	<i>Apodemus</i> spp.	UK
<i>Bartonella tribocorum</i> - AM260525	<i>Rattus norvegicus</i>	France
<i>Bartonella vinsonii</i> subsp. <i>vinsonii</i> - L01259	Vole (unspecified)	Canada
<i>Brucella microti</i> - AM392286	<i>Microtus arvalis</i>	Czech Republic
<i>Daeguia caeni</i> - EF532794		
<i>Falsochrobactrum ovis</i> - KC254733	<i>Ovis aries</i>	
<i>Mycoplana dimorpha</i> - D12786		
<i>Ochrobactrum anthropi</i> - NR_114979	<i>Homo sapiens</i>	France
<i>Paenochrobactrum gallinarii</i> - FN391023		Germany
<i>Pseudochrobactrum asaccharolyticum</i> - AM180485	<i>Homo sapiens</i>	Sweden
<i>Zoogloea ramigera</i> - D14255		



9

Figure S1: Variable base-pair positions within the V4 region of the 16S rRNA bacterial gene of a selection of reference *Bartonella* sequences and *Bartonella* OTUs obtained from *Arvicola terrestris* in Franche-Comté, France. Tile colour indicates the base pair in each sequence at each position. Sequences were clustered using complete linkage of raw distances (ie. no assumptions were made about evolution rates). Black tiles indicate gaps (-) in the alignment



∞ Figure S2: nMDS ordination of differences between Bartonella OTU seed sequences (numbered, coloured to indicate frequency of detection within the animals sampled) and reference *Bartonella* sp. sequences (red). Additional reference sequences include *Mycoplana dimporpha*, *Zooglea ramigera*, *Daeguia caeni*, *Falsochromobacterium ovis*, *Paenochrobacterium gallinarii*, *Pseudochrobacterium asaccharolyticum*, *Brucella microti* and *Ochrobacterium anthropi*. See Figure S3 for expanded view of the ordination within the grey dashed box.

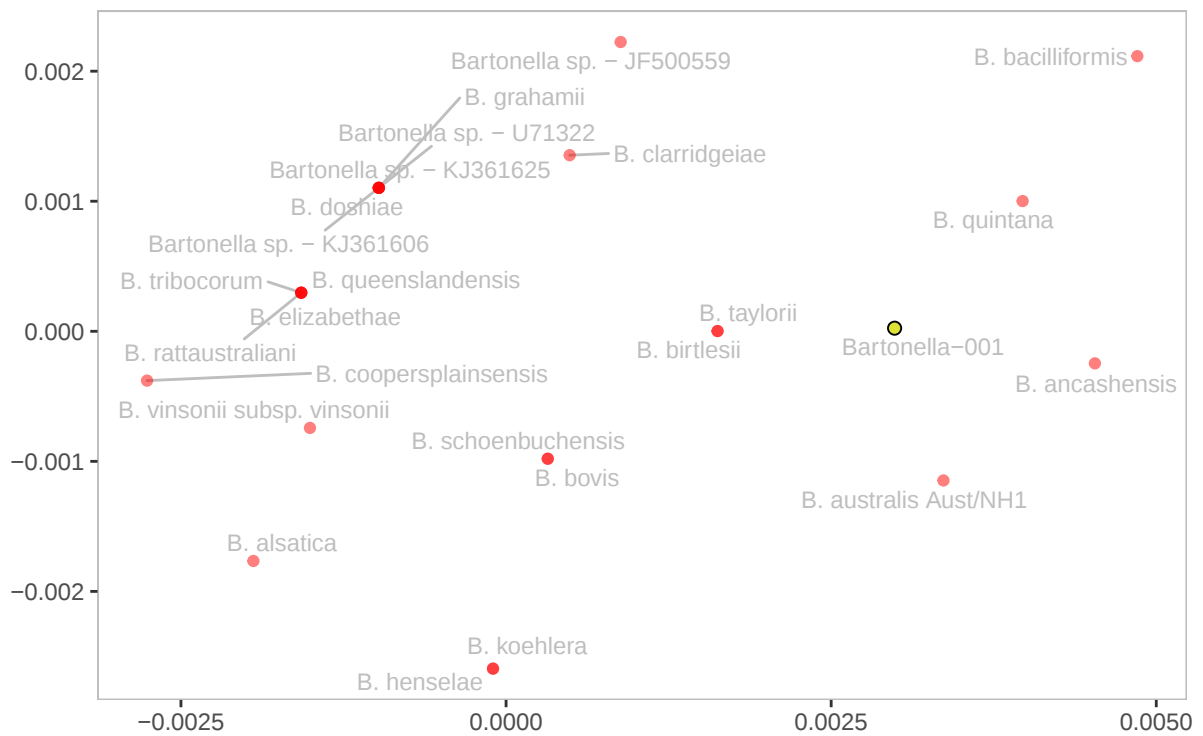


Figure S3: Inset of Figure S2 showing nMDS ordination of differences between the Bartonella-001 seed sequence (yellow) and reference *Bartonella* sp. sequences in red.

3 Bordetella

The seed sequence of Bordetella-001, the most prevalent of the 11 Bordetella OTUs in the data set, is identical to V4 16S rRNA reference sequences of *Bordetella parapertussis* and *B. bronchiseptica*. The seed sequence of Bordetella-003, the second-most prevalent Bordetella OTU, differs from these two reference sequences by 4 base-pairs. The remaining Bordetella OTU seed sequences differ from reference sequences by at least 4 base-pairs and do not show strong affinities with any of the reference sequences.

Table S3: Species and GenBank accession numbers, source and country of isolation for reference strains of *Bordetella* sp. used in comparisons.

Species - GenBank Accession	Isolation Source	Country
<i>B. avium</i> - AF177666	<i>Meleagris gallopavo</i>	Germany
<i>B. bronchiseptica</i> - U04948	<i>Canis lupus familiaris</i>	
<i>B. hinzii</i> - AF177667	<i>Gallus gallus</i>	Australia
<i>B. holmesii</i> - U04820	<i>Homo sapiens</i>	USA
<i>B. muralis</i> - LC053647	environment	Japan
<i>B. parapertussis</i> - U04949	<i>Homo sapiens</i>	
<i>B. sputigena</i> - KF601914	<i>Homo sapiens</i>	Sweden

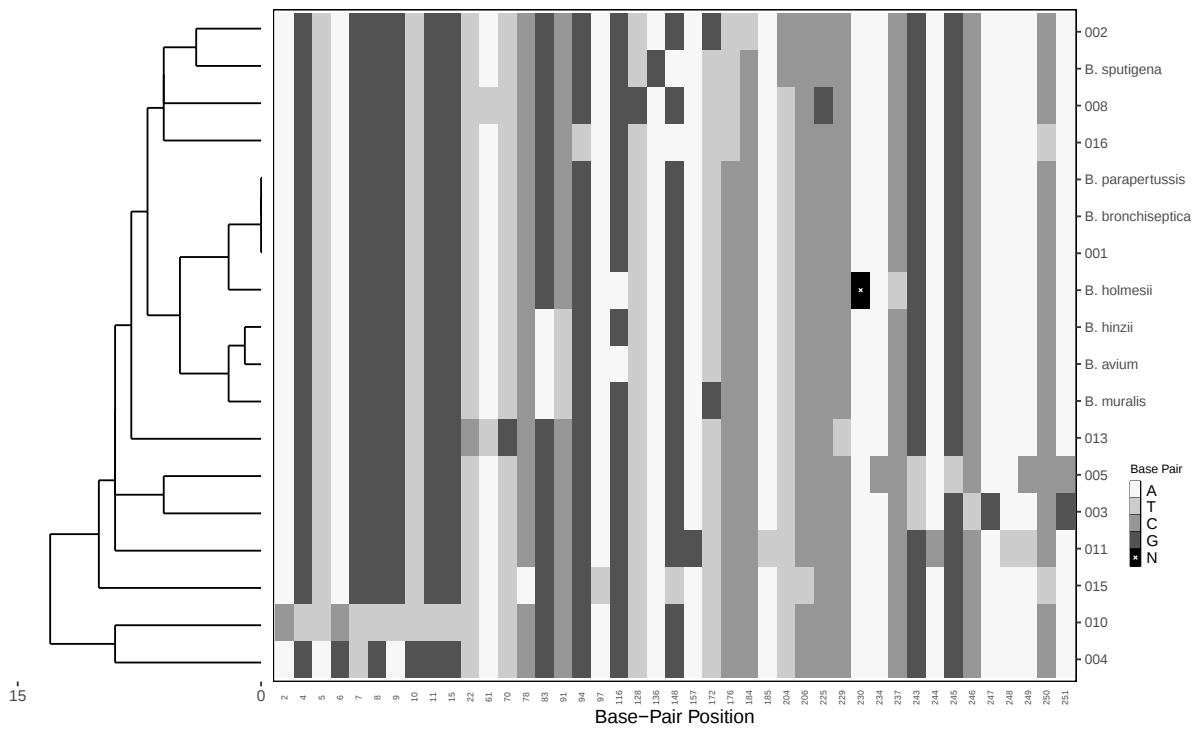


Figure S4: Variable base-pair positions within the V4 region of the 16S rRNA bacterial gene of a selection of reference *Bordetella* sequences and *Bordetella* OTU seed sequences (numbered) obtained from *Arvicola terrestris* in Franche-Comté, France. Tile colour indicates the base pair in each sequence at each position. Sequences were clustered using complete linkage of raw distances (ie. no assumptions were made about evolution rates). Black crossed tiles (N) indicate ambiguous base pairs in the sequence.

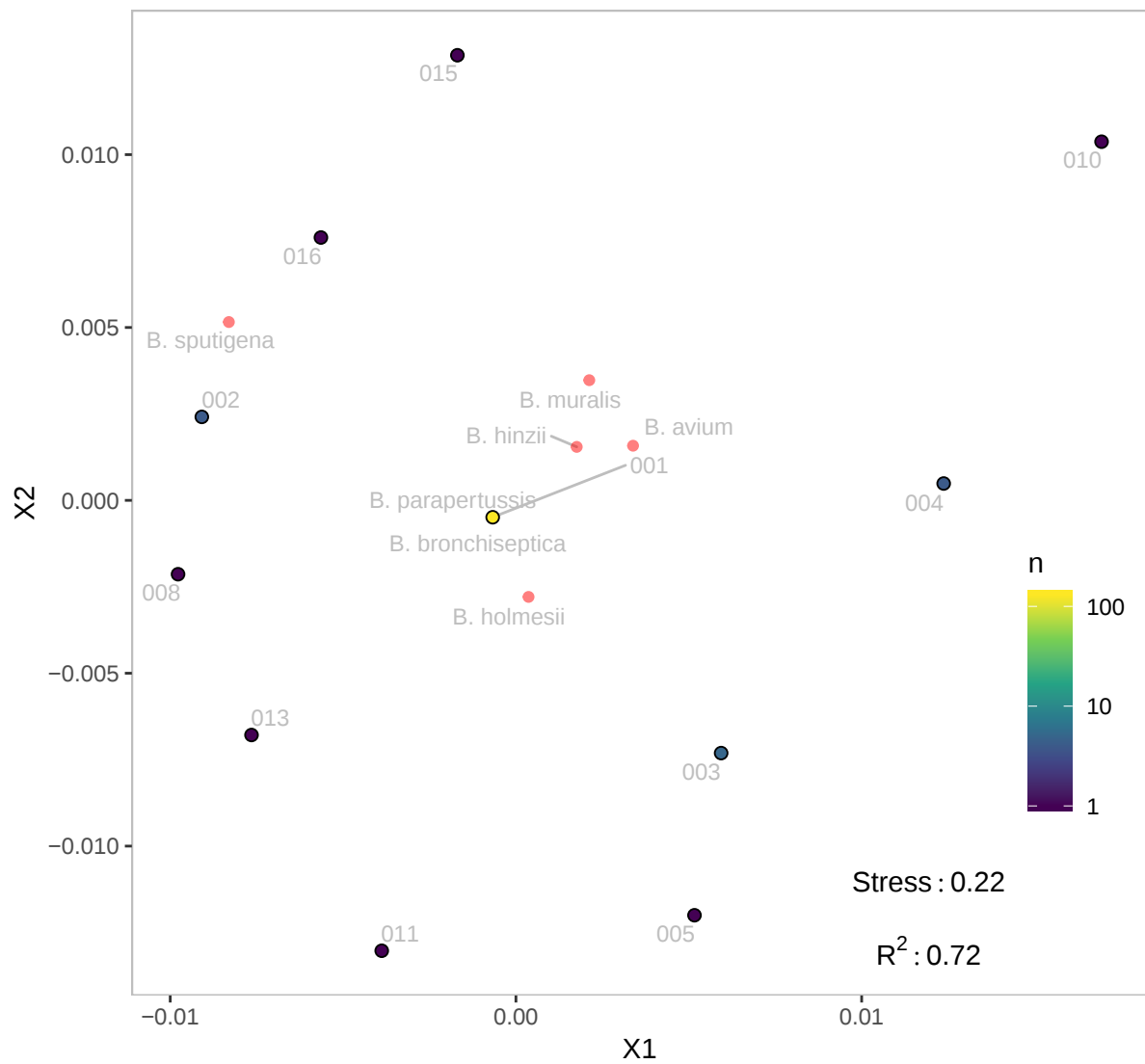


Figure S5: nMDS ordination of differences between *Bordetella* OTU seed sequences (numbered, coloured to indicate frequency of detection within the animals sampled) and reference *Bordetella* sp. sequences (red).

4 Borelliella

- 45 The seed sequence of *Borreliella*-001, the only *Borreliella* OTU in this data set, is identical to reference sequences of *Borelliella* (*Borellia*) *burgdorferi* as well as an unspecified *Borelliella* strain.

Table S4: Similarity, expressed as a percentage, and coverage (in parentheses) of the OTU *Borreliella*-001 compared to a selection of reference strains of *Borreleilla* (*Borrelia*) *burgdorferi* [4]. Sequences were compared using the blastn search tool.

Reference Sequence	Similarity % (Coverage %)
B. burgdorferi (Lyme disease spirochete) - U03396	100 (100)
B. burgdorferi (Lyme disease spirochete) - X85204	100 (100)
B. sp. - AF467971	100 (100)
B. burgdorferi JD1 - CP002312	100 (100)
B. burgdorferi N40 - CP002228	100 (100)
B. burgdorferi ZS7 - CP001205	100 (100)

5 Filobacterium

23 OTUs were assigned to *Filobacterium*; due to the low number of reference sequences
 50 available, they were compared to the two reference sequences of *Filobacterium rodentium* deposited in GenBank (accession numbers LC036397 and LC055729). The seed sequence of Filobacterium-001 is identical to both reference sequences; the remaining OTU seed sequences range in similarity from 96.8% and 99.6%.

Table S5: Similarity (expressed as a percentage) and coverage of the *Filobacterium* OTU's compared to two reference *Filobacterium* strains from the GenBank database. Sequences were compared using the blast+n search tool during OTU procession

OTUlabel	F. rodentium - LC036397	F. rodentium - LC055729
Filobacterium-001	100 (100)	100 (100)
Filobacterium-002	99.59 (96.8)	99.59 (96.8)
Filobacterium-003	99.58 (94.4)	99.58 (94.4)
Filobacterium-004	97.99 (99.2)	97.99 (99.2)
Filobacterium-006	98.01 (100)	98.01 (100)
Filobacterium-007	97.61 (100)	97.61 (100)
Filobacterium-008	97.61 (100)	97.61 (100)
Filobacterium-009	96.41 (100)	96.41 (100)
Filobacterium-013	98.79 (98.4)	98.79 (98.4)
Filobacterium-017	98.41 (100)	98.41 (100)
Filobacterium-018	98.01 (100)	98.01 (100)
Filobacterium-023	98.39 (98.8)	98.39 (98.8)
Filobacterium-026	98.01 (100)	98.01 (100)
Filobacterium-027	98.41 (100)	98.41 (100)
Filobacterium-029	96.81 (100)	96.81 (100)
Filobacterium-030	97.21 (100)	97.21 (100)
Filobacterium-031	97.61 (100)	97.61 (100)
Filobacterium-032	98.41 (100)	98.41 (100)
Filobacterium-034	97.61 (100)	97.61 (100)
Filobacterium-040	98.01 (100)	98.01 (100)
Filobacterium-043	98.41 (100)	98.41 (100)
Filobacterium-044	98.01 (100)	98.01 (100)
Filobacterium-045	98.01 (100)	98.01 (100)

6 Leptospira

55 The seed sequence of Leptospira-001, the most globally prevalent Leptospira OTU, is identical to the type sequences for *L. interrogans* and *L. noguchii* Panama. The majority of the remaining Leptospira OTU seed sequences differ from reference sequences by less than 10 base pairs, but few clear affinities between seed sequences and references are present (Figures S6 and S7)

Table S6: Accession numbers, species, host and country of isolation and associated publications for reference strains of *Leptospira* sp. used in comparisons. *Mayotte is a island department of France located in the Indian Ocean

Species - Accession	Isolation Source	Country
<i>Leptospira alexanderi</i> serovar Manhao 3 - AY631880		China
<i>Leptospira alstonii</i> serovar Sichuan - AY631881		
<i>Leptospira broomii</i> - AY796065	<i>Homo sapiens</i>	
<i>Leptospira idonii</i> - AB721966		
<i>Leptospira interrogans</i> - Z12817	<i>Homo sapiens</i>	
<i>Leptospira kmetyi</i> serovar Malaysia - AB279549		
<i>Leptospira licerasiae</i> serovar Varillal - EF612284	<i>Homo sapiens</i>	Peru
<i>Leptospira meyeri</i> serovar Ranarum - AY631878	<i>Rana pipiens</i>	USA
<i>Leptospira noguchii</i> serovar Panama - AY631886	<i>Didelphis marsupialis</i>	Panama
<i>Leptospira parva</i> serovar Parva - AY293856		
<i>Leptospira santarosai</i> serovar Shermani - AY631883	<i>Proechimys semispinosus</i>	Panama
<i>Leptospira wolbachii</i> serovar Codice - AY631879		USA
<i>Leptospira wolffii</i> serovar Khorat - EF025496	<i>Homo sapiens</i>	Thailand
<i>Leptospira mayottensis</i> - KJ847187	<i>Homo sapiens</i>	France (Mayotte, Indian Ocean)
<i>Leptospira borgpetersenii</i> serovar Tarassovi - AM050577		
<i>Leptospira</i> sp. E156 - LC196101		
<i>Leptospira</i> sp. E152 - LC196099		

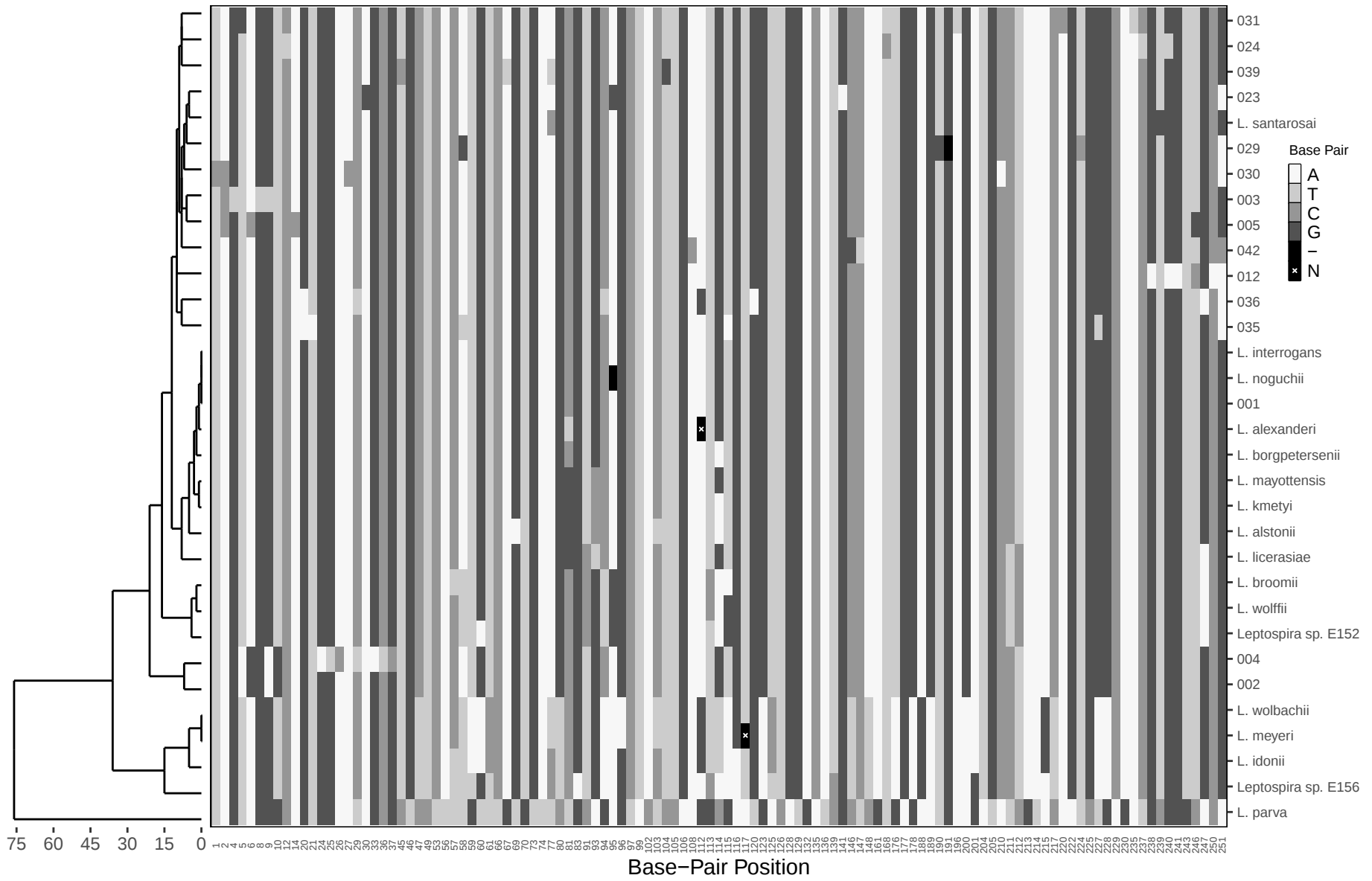


Figure S6: Variable base-pair positions within the V4 region of the 16S rRNA bacterial gene of a selection of reference *Leptospira* sequences and *Leptospira* OTU seed sequences (numbered) obtained from *Arvicola terrestris* in Franche-Comté, France. Tile colour indicates the base pair in each sequence at each position. Sequences were clustered using complete linkage of raw distances (ie. no assumptions were made about evolution rates). Black crossed tiles (N) indicate ambiguous base pairs in the sequence, and black tiles indicate gaps (-) in the alignment

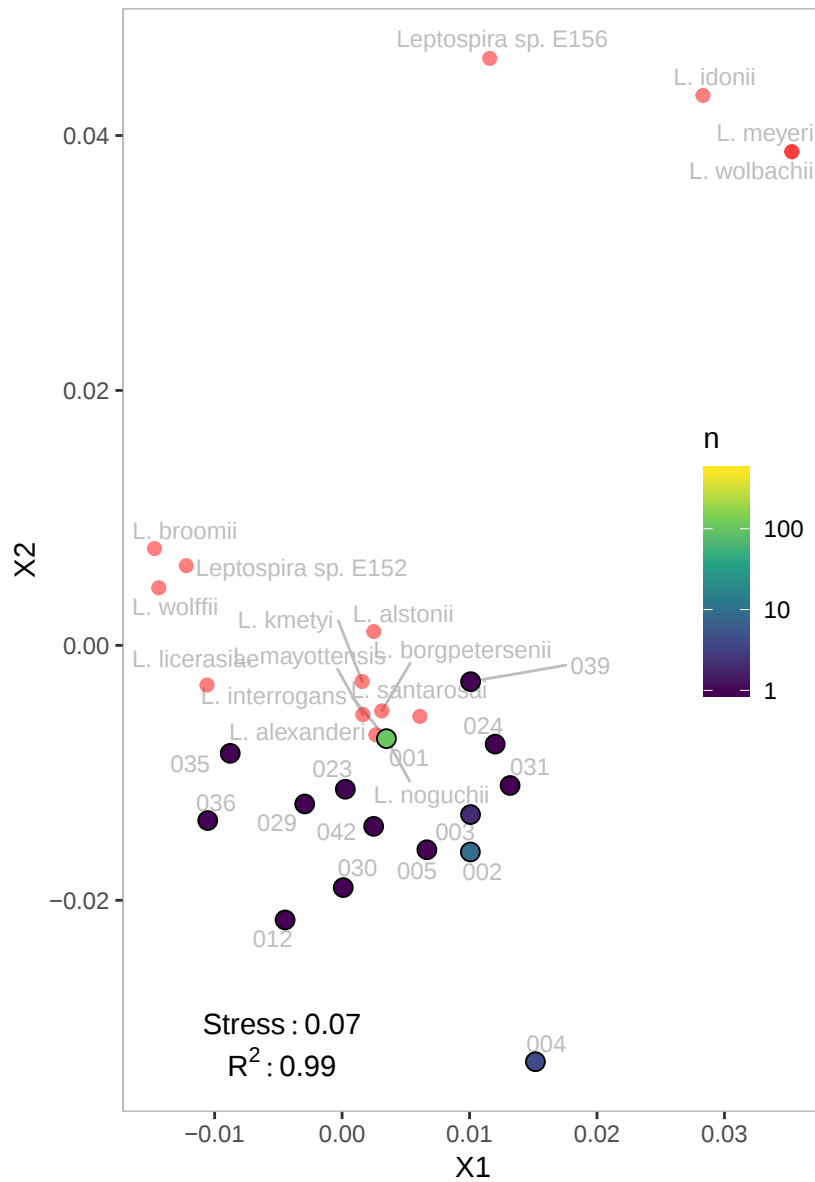


Figure S7: nMDS ordination of differences between *Leptospira* OTU seed sequences (numbered, coloured to indicate frequency of detection within the animals sampled) and reference *Leptospira* sp. sequences (red).

60 7 **Mycobacterium**

The seed sequences of Mycobacterium-001 and -002 are identical to several reference sequences from the *Mycobacterium* genus; Mycobacterium-001 is identical to *M. llatzerense*, *M. mucogenicum*, and *M. phlei*, while Mycobacterium-002 is identical to *M. florentinum*, *M. heidelbergense*, *M. lentiflavum*, *M. kubicae*, *M. montefiorensense*, *M. parmense*, *M. stom-*
65 *atepiae*, *M. triplex*.

Table S7: Similarity, expressed as a percentage, and coverage (in parentheses) of the *Mycobacterium* OTU's compared to a selection of typ strains of the genus *Mycobacterium*. Sequences were compared using the blastn search tool.

Reference Sequence	Mycobacterium-001	Mycobacterium-002
<i>M. chelonae</i> - AY457072	99.203 (100)	95.618 (100)
<i>M. florentinum</i> - AJ616230	96.414 (100)	100 (100)
<i>M. heidelbergense</i> - AJ000684	96.414 (100)	100 (100)
<i>M. insubricum</i> - EU605695	100 (100)	96.414 (100)
<i>M. kubicae</i> - AF133902	96.414 (100)	100 (100)
<i>M. lentiflavum</i> - AF480583	96.414 (100)	100 (100)
<i>M. llatzerense</i> - AJ746070	100 (100)	96.414 (100)
<i>M. montefiorensense</i> - AF330038	96.414 (100)	100 (100)
<i>M. mucogenicum</i> - AY457074	100 (100)	96.414 (100)
<i>M. paratuberculosis</i> - X52931	96.016 (100)	99.602 (100)
<i>M. parmense</i> - AF466821	96.414 (100)	100 (100)
<i>M. phlei</i> - AF480603	100 (100)	96.414 (100)
<i>M. stomatepiae</i> - AM884331	96.414 (100)	100 (100)
<i>M. triplex</i> - U57632	96.414 (100)	100 (100)
<i>M. tuberculosis</i> - X60070	96.016 (100)	99.602 (100)

8 Mycoplasma and Ureaplasma

The majority of the Mycoplasma OTU seed sequences are most similar to the hemotrophic mycoplasmas, with two matching exactly to reference sequences; Mycoplasma-032 is identical to *M. microti*, and Mycoplasma-002 is identical to an unspecified hemotrophic mycoplasma isolated from *Akodon* sp. (Figures S8 and S9). The remaining hemotrophic mycoplasma OTUs do not form clear affinities to described mycoplasma species, with Mycoplasma-001, the most prevalent, differing from *Mycoplasma haemomurus* by 18 base-pairs. Two OTUs appear more similar to non-hemotrophic mycoplasmas, with Mycoplasma-009 and Mycoplasma-016 most similar to *M. neurolyticum* (differing by 17 and 14 base-pairs respectively).

Table S8: Species and GenBank accession numbers, source and country of isolation for reference strains of *Mycoplasma* sp. used in comparisons.

Species - GenBank Accession	Isolation Source	Country
<i>Anaeroplasma varium</i> - M23934		
<i>Asteroleplasma anaerobium</i> - M22351		
Candidatus <i>Mycoplasma haemomuris</i> subsp. <i>musculi</i> - AB758440	<i>Apodemus argenteus</i>	Japan
Candidatus <i>Mycoplasma haemomuris</i> subsp. <i>ratti</i> - AB758439	<i>Rattus rattus</i>	Japan
<i>Haloplasma contractile</i> - EF999972		
<i>Mesoplasma tabanidae</i> - AY187288	<i>Tabanus abactor</i>	USA
<i>Mycoplasma arginini</i> - AF125581	<i>Panthera leo</i>	
<i>Mycoplasma arthritidis</i> - M24580	<i>Rattus</i> sp.	
<i>Mycoplasma caviae</i> - AF221111	<i>Cavia porcellus</i>	
<i>Mycoplasma cavipharyngis</i> - AF125879	<i>Cavia porcellus</i>	
<i>Mycoplasma citelli</i> - AF412973	<i>Citellus richardsonii</i>	
<i>Mycoplasma coccoides</i> - AY171918	<i>Mus</i> sp.	
<i>Mycoplasma collis</i>		
ATCC 35278 - AF538681	<i>Rattus</i> sp.	UK
<i>Mycoplasma cricetuli</i>		
ATCC 35279 - AF412976	<i>Cricetulus griseus</i>	
<i>Mycoplasma haemofelis</i> - AF178677	<i>Felis catus</i>	USA
<i>Mycoplasma haemofelis</i> - U88563	<i>Apodemus argenteus</i>	Japan
<i>Mycoplasma haemofelis</i> - U95297	<i>Felis catus</i>	USA
<i>Mycoplasma haemomuris</i> - U82963	<i>Apodemus argenteus</i>	Japan
<i>Mycoplasma microti</i> - AF212859	<i>Microtus ochrogaster</i>	USA
<i>Mycoplasma muris</i> - M23939	<i>Mus</i> sp.	
<i>Mycoplasma neurolyticum</i> - M23944	<i>Mus</i> sp.	
<i>Mycoplasma ovipneumoniae</i> - U44771	<i>Ovis aries</i>	
<i>Mycoplasma oxoniensis</i> - AF412987	<i>Cricetulus griseus</i>	
<i>Mycoplasma pneumoniae</i>		
ATCC 15531 - AF132740	<i>Homo sapiens</i>	
<i>Mycoplasma pulmonis</i> - AF125582	<i>Rattus</i> sp.	
<i>Mycoplasma putrefaciens</i> - U26055	<i>Capra aegagrus</i>	USA
<i>Spiroplasma sabaudiense</i> - AY189308	<i>Aedes sticticus/vexans</i>	France
uncultured <i>Mycoplasma</i> sp. - KT215637	<i>Akodon</i> sp.	Brazil (Atlantic Forest)
uncultured <i>Mycoplasma</i> sp. - KT215638	<i>Necromys lasiurus</i>	Brazil (Atlantic Forest)
<i>Ureaplasma gallorale</i> - U62937	<i>Gallus gallus domesticus</i>	
<i>Ureaplasma urealyticum</i>		
ATCC27618 - AF073450	<i>Homo sapiens</i>	

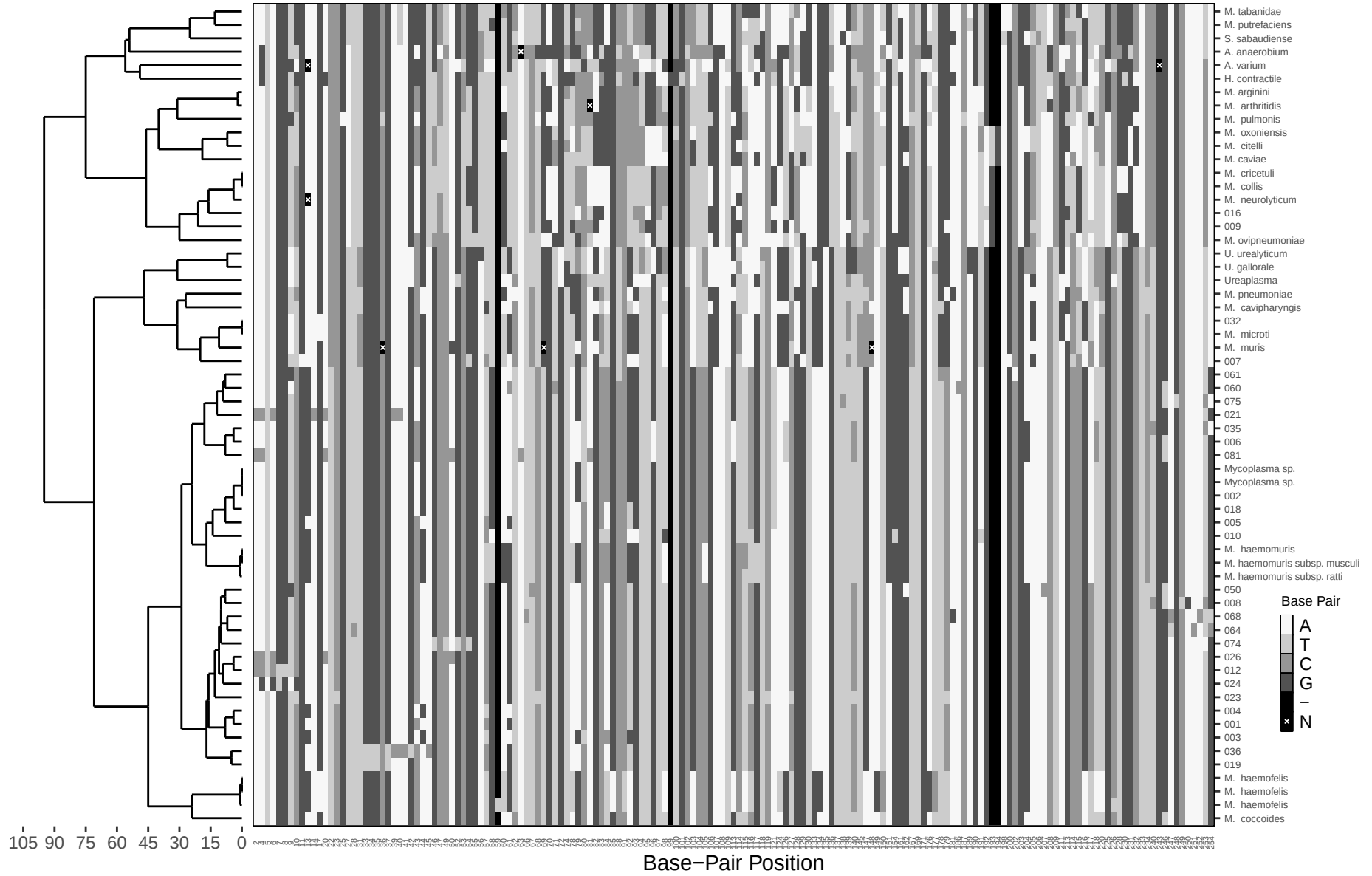


Figure S8: Variable base-pair positions within the V4 region of the 16S rRNA bacterial gene of a selection of reference *Mycoplasma* sequences and *Mycoplasma* OTUs obtained from *Arvicola terrestris* in Franche-Comté, France. Tile colour indicates the base pair in each sequence at each position. Sequences were clustered using complete linkage of raw distances (ie. no assumptions were made about evolution rates). Black crossed tiles (N) indicate ambiguous base pairs in the sequence, solid black tiles indicate gaps (-) in the alignment

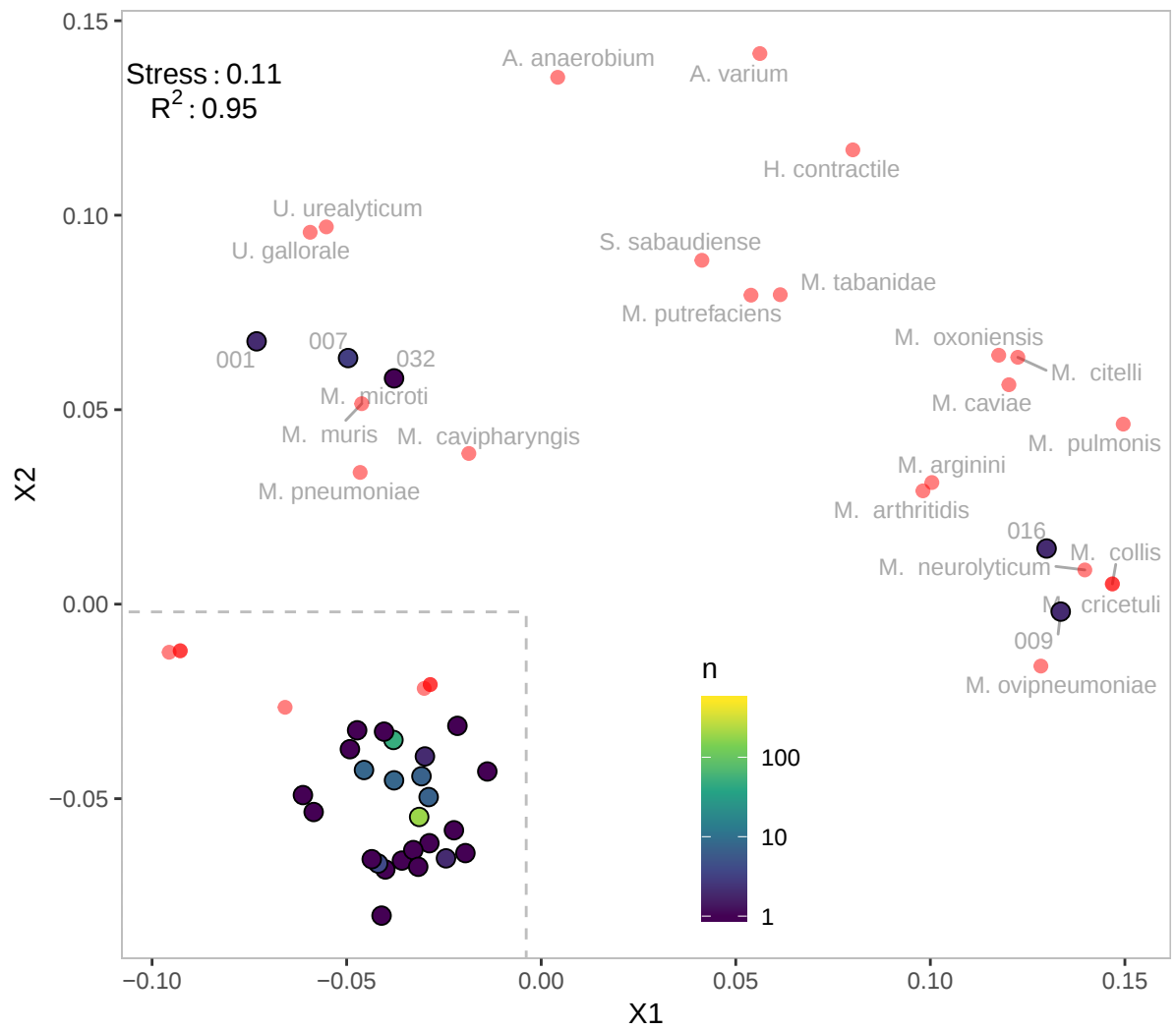


Figure S9: nMDS ordination of differences between Mycoplasma OTU seed sequences (coloured to indicate frequency within the data set) and references Mycoplasma sequences (red). See Figure S10 for expanded view of the ordination within the grey dashed box.

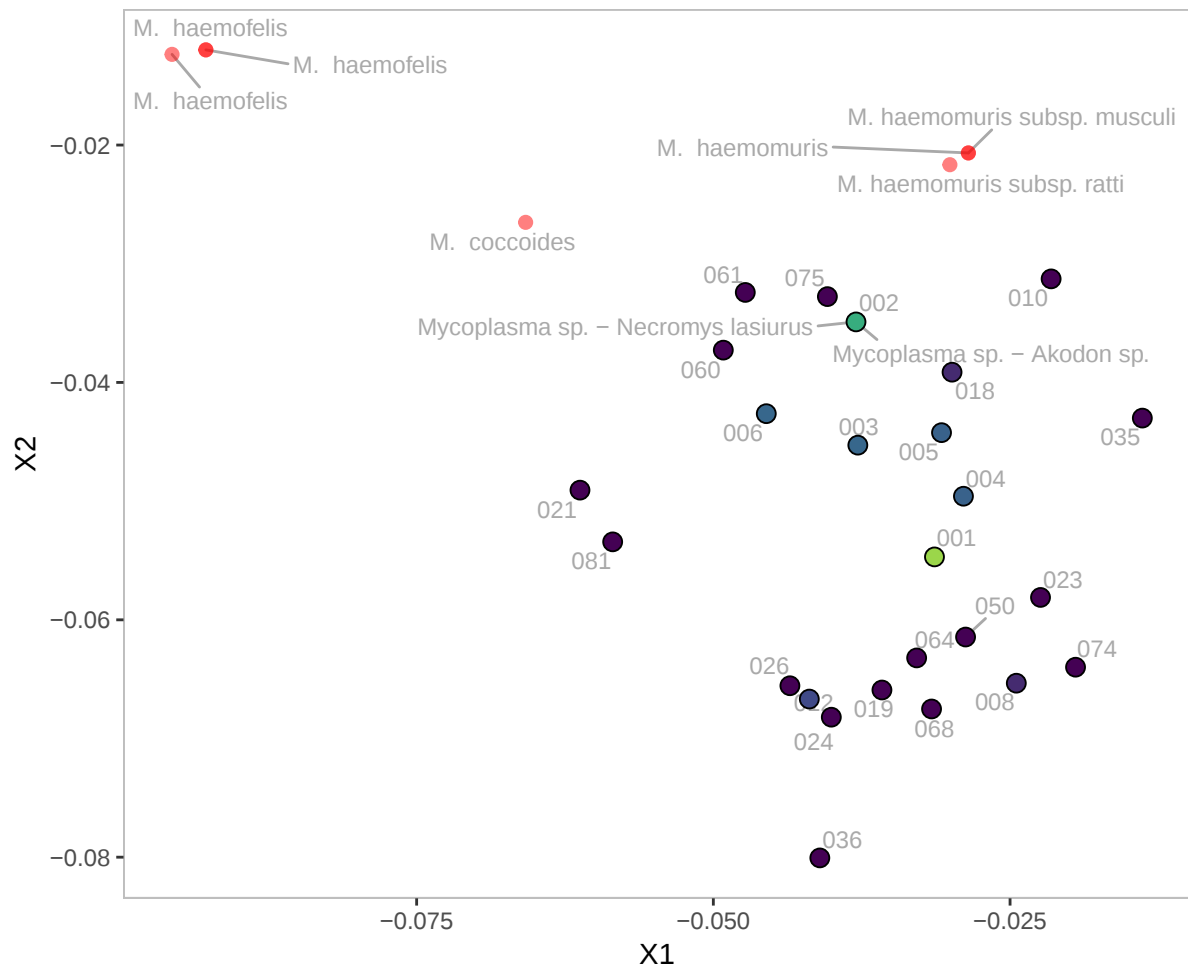


Figure S10: Inset of Figure S9 showing nMDS ordination of differences between *Mycoplasma* OTU seed sequence (numbered, coloured) and reference *Mycoplasma* sp. sequences in red.

9 Pasteurellaceae

The *Pasteurellaceae* as a family is currently taxonomically ambiguous; with a myriad of undescribed species and strains poorly characterized, new taxonomies are produced relatively frequently, often with radically different topologies depending on the loci
80 under consideration. That being said, we don't attempt to assign genera to the majority of our Pasteurellaceae OTUs unless the seed sequence matches a described species exactly (Actinobacillus-001, for example).

Table S9: Accession numbers, species, host and country of isolation and associated publications for reference strains of *Pasteurellaceae* sp. used in comparisons. *Members of the Rodent Group described in [5].

Species - Accession	Isolation Source	Country
<i>Muribacter muris</i> * - AF024526	mouse (unspecified)	USA
<i>Rodentibacter rarus</i> * - AF024529	<i>Rattus</i> sp.	USA
<i>Haemophilus influenzae-murium</i> * - AF024530	mouse (unspecified)	
<i>Mannheimia granulomatis</i> - AF053902	<i>Bos</i> sp.	Brazil
<i>Haemophilus felis</i> - AF224292	<i>Felis catus</i>	USA
<i>Pasteurella</i> sp. MCCM 00235 * - AF224300		
<i>Pasteurellaceae bacterium</i> * - AF224301	rodent (unspecified)	
<i>Mesocricetibacter intestinalis</i> * - AF224302	<i>Mesocricetus auratus</i>	Germany
<i>Necropsobacter rosorum</i> * - AF224303	<i>Cavia porcellus</i>	Germany
<i>Pasteurella</i> sp. MCCM 02120* - AF224304		
<i>Gallibacterium anatis</i> - AF228001	<i>Anas platyrhynchos domesticus</i>	Denmark
<i>Pasteurella multocida</i> subsp. <i>multocida</i> - AF294410	<i>Sus scrofa domesticus</i>	Canada
<i>Pasteurella multocida</i> subsp. <i>septica</i> - AF294411	<i>Homo sapiens</i>	France
<i>Pasteurella skyensis</i> - AJ243202	<i>Salmo salar</i>	UK
<i>Haemophilus pittmaniae</i> - AJ290755	<i>Homo sapiens</i>	Denmark
<i>Haemophilus parahaemolyticus</i> - AJ295746	<i>Homo sapiens</i>	USA
<i>Pasteurella multocida</i> subsp. <i>multocida</i> - AY078999	<i>Sus scrofa domesticus</i>	Canada
<i>Haemophilus aegyptius</i> - AY362905	<i>Homo sapiens</i>	USA
<i>Haemophilus haemoglobinophilus</i> - AY362907	<i>Canis lupus familiaris</i>	Denmark
<i>Haemophilus parainfluenzae</i> - AY362908	<i>Homo sapiens</i>	UK
<i>Glaesserella parasuis</i> - AY362909		
<i>Pasteurella bettyae</i> - AY362917	<i>Homo sapiens</i>	UK
<i>Avibacterium gallinarum</i> - AY362921	<i>Gallus gallus</i>	Denmark
<i>Pasteurella langaaensis</i> - AY362922		Denmark
<i>Pasteurella mairii</i> - AY362923	<i>Sus scrofa domesticus</i>	UK
<i>Bibersteinia trehalosi</i> - AY362927	<i>Ovis aries</i>	UK
<i>Avibacterium paragallinarum</i> - AY498868	<i>Gallus gallus</i>	Germany
<i>Haemophilus influenzae</i> - M35019	<i>Homo sapiens</i>	UK
<i>Haemophilus ducreyi</i> - M63900	<i>Homo sapiens</i>	France
<i>Aggregatibacter actinomycetemcomitans</i> - M75039	<i>Homo sapiens</i>	UK
<i>Aggregatibacter aphrophilus</i> - M75042	<i>Homo sapiens</i>	UK
<i>Aggregatibacter segnis</i> - M75043	<i>Homo sapiens</i>	Denmark
<i>Pasteurella oralis</i> - M75052	<i>Homo sapiens</i> bitten by <i>Felis catus</i>	Germany
<i>Avibacterium avium</i> - M75058	<i>Gallus gallus</i>	Germany
<i>Haemophilus paracuniculus</i> - M75061	<i>Oryctolagus cuniculus</i>	USA
<i>Avibacterium volantium</i> - M75070	<i>Galloanserae</i> sp.	UK
<i>Actinobacillus pleuropneumoniae</i> - M75074	<i>Sus scrofa domesticus</i>	Germany
<i>Haemophilus paraphrohaemolyticus</i> - M75076	<i>Homo sapiens</i>	UK
<i>Mannheimia haemolytica</i> - M75080	<i>Ovis aries</i>	UK
<i>Rodentibacter pneumotropicus</i> * - M75083	<i>Mus musculus</i> (Swiss mouse)	USA
<i>Haemophilus massiliensis</i> - NR149208	<i>Homo sapiens</i>	Senegal

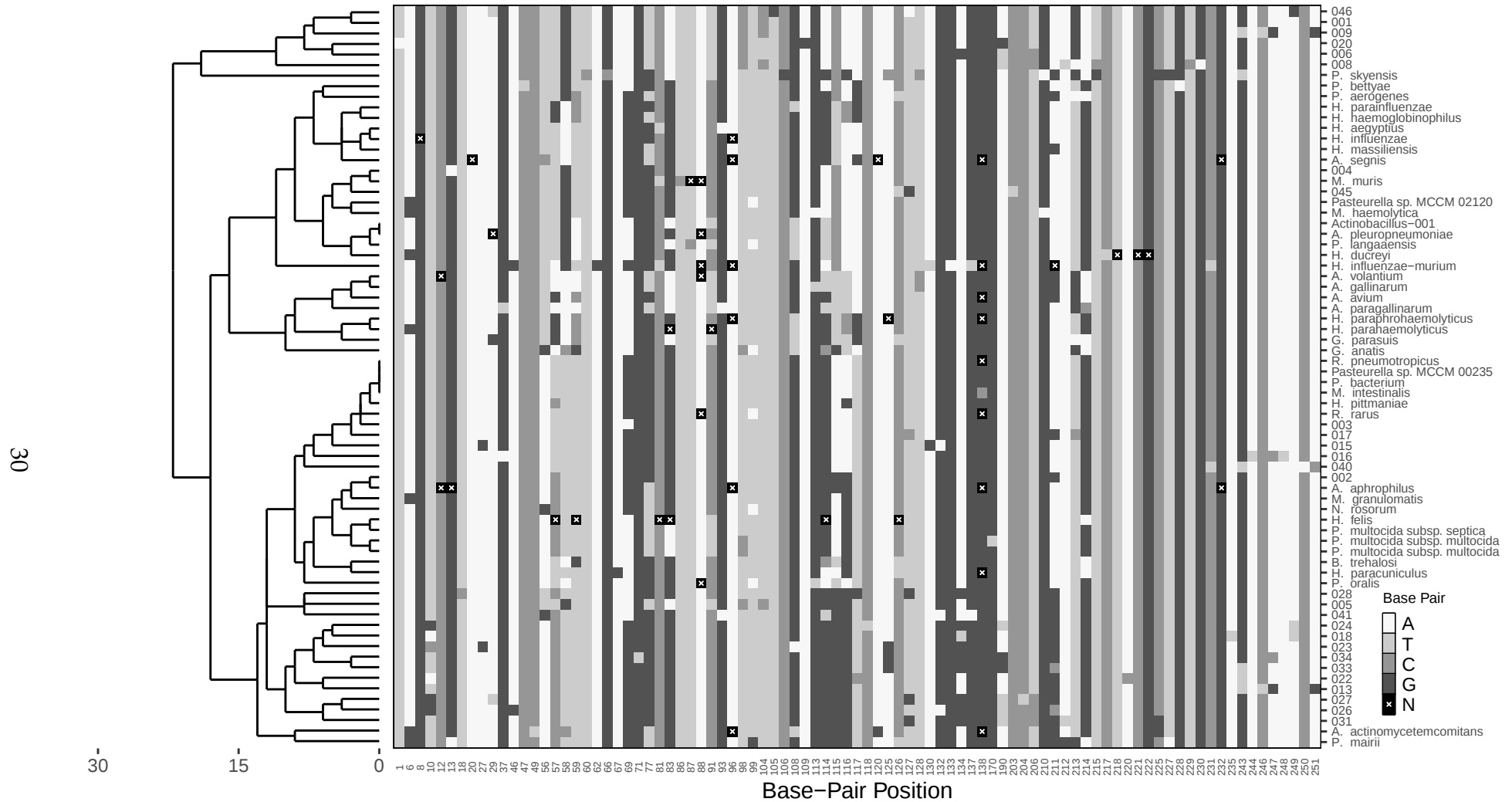


Figure S11: Variable base-pair positions within the V4 region of the 16S rRNA bacterial gene of a selection of reference *Pasteurellaceae* sequences and *Pasteurellaceae* OTUs obtained from *Arvicola terrestris* in Franche-Comté, France. Tile colour indicates the base pair in each sequence at each position. Sequences were clustered using complete linkage of raw distances (ie. no assumptions were made about evolution rates). Black crossed tiles (N) indicate ambiguous base pairs in the sequence.

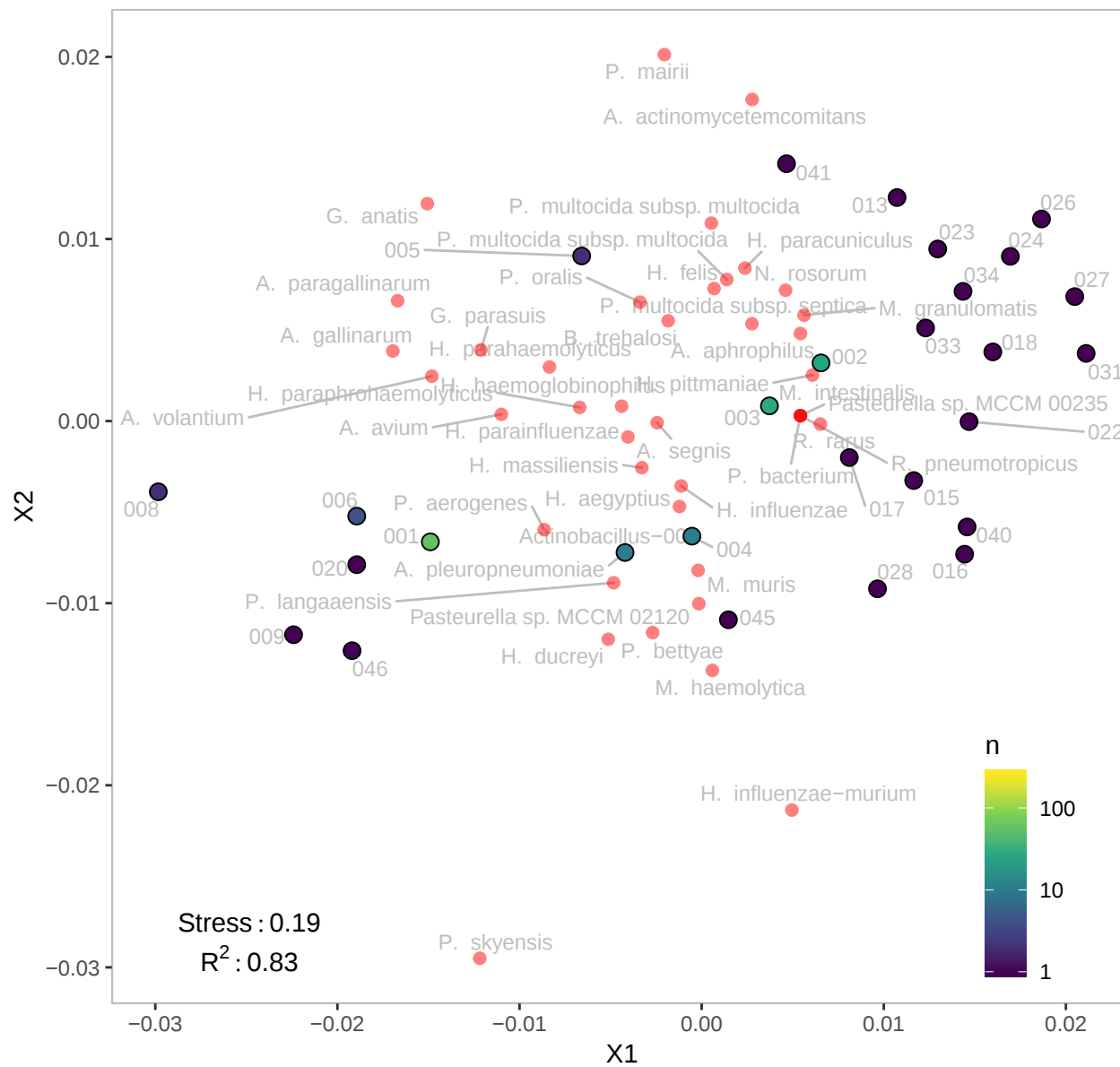


Figure S12: nMDS ordination of differences between Pasteurellaceae OTU seed sequences (numbered, coloured to indicate frequency within the data set) and references *Pasteurellaceae* sequences (red).

10 Rickettsia

Only one OTU in this data set was assigned to *Rickettsia*; the seed sequences is most
85 similar to several members of the tick-borne spotted fever group.

Table S10: Similarity, expressed as a percentage, and coverage (in parentheses) of the OTU Rickettsia-001 compared to a selection of reference strains of *Rickettsia* sp. Sequences were compared using the blastn search tool.

OTUlabel	Reference Sequence	Similarity % (Coverage %)
Rickettsia-001	R. canadensis str. CA410 - CP003304	98.01 (100)
Rickettsia-001	R. canadensis - L36104	98.01 (100)
Rickettsia-001	R. rhipicephali - L36216	98.01 (100)
Rickettsia-001	R. canadensis str. McKiel - CP000409	98.01 (100)

11 *Yersinia*

Only one OTU in this data set was assigned to *Yersinia*; the seed sequence is identical to a number of species in this genus, as well as the type sequence for *Serratia liquefaciens*.

Table S11: Similarity, expressed as a percentage, and coverage (in parentheses) of the *Yersinia* OTU compared to a selection of reference strains of *Yersinia* sp. and *Serratia liquefaciens*. Sequences were compared using the blastn search tool.

Reference Sequence	Similarity % (Coverage %)
<i>Serratia liquefaciens</i> - CP011303	100 (100)
<i>Y. kristensenii</i> - CP009997	100 (100)
<i>Y. pestis</i> A1122 - CP009840	100 (100)
<i>Y. enterocolitica</i> - CP009838	100 (100)
<i>Y. intermedia</i> - CP009801	100 (100)
<i>Y. aldovae</i> 670-83 - CP009781	100 (100)
<i>Y. rohdei</i> - CP009787	100 (100)
<i>Y. pseudotuberculosis</i> - CP009786	100 (100)

12 GenBank Accession Numbers

Table S12: GenBank accession numbers for OTU seed sequences presented in the text

OTUlabel	accession	OTUlabel	accession	OTUlabel	accession
Actinobacillus-001	MN594323	Borrelia-001	MN594321	Mycoplasma-021	MN594372
Anaplasma-001	MN594320	Filobacterium-001	MN594312	Mycoplasma-023	MN594351
Bartonella-001	MN594311	Filobacterium-002	MN594370	Mycoplasma-024	MN594354
Bartonella-002	MN594344	Filobacterium-003	MN594366	Mycoplasma-026	MN594368
Bartonella-003	MN594343	Filobacterium-004	MN594376	Mycoplasma-032	MN594345
Bartonella-004	MN594337	Filobacterium-006	MN594378	Mycoplasma-035	MN594373
Bartonella-005	MN594338	Filobacterium-007	MN594381	Mycoplasma-036	MN594375
Bartonella-006	MN594347	Filobacterium-008	MN594382	Mycoplasma-050	MN594400
Bartonella-007	MN594341	Filobacterium-009	MN594383	Mycoplasma-060	MN594407
Bartonella-008	MN594355	Filobacterium-013	MN594387	Mycoplasma-061	MN594408
Bartonella-009	MN594356	Filobacterium-017	MN594416	Mycoplasma-064	MN594409
Bartonella-010	MN594350	Filobacterium-018	MN594417	Mycoplasma-068	MN594410
Bartonella-012	MN594339	Filobacterium-023	MN594438	Mycoplasma-074	MN594411
Bartonella-014	MN594353	Filobacterium-026	MN594440	Mycoplasma-075	MN594412
Bartonella-018	MN594364	Filobacterium-027	MN594448	Mycoplasma-081	MN594419
Bartonella-021	MN594380	Filobacterium-029	MN594449	Pasteurellaceae-001	MN594319
Bartonella-026	MN594374	Filobacterium-030	MN594450	Pasteurellaceae-002	MN594317
Bartonella-028	MN594379	Filobacterium-031	MN594451	Pasteurellaceae-003	MN594318
Bartonella-032	MN594385	Filobacterium-032	MN594452	Pasteurellaceae-004	MN594328
Bartonella-035	MN594388	Filobacterium-034	MN594453	Pasteurellaceae-005	MN594358
Bartonella-039	MN594389	Filobacterium-040	MN594454	Pasteurellaceae-006	MN594365
Bartonella-041	MN594390	Filobacterium-043	MN594455	Pasteurellaceae-008	MN594346
Bartonella-042	MN594391	Filobacterium-044	MN594456	Pasteurellaceae-009	MN594357
Bartonella-044	MN594392	Filobacterium-045	MN594457	Pasteurellaceae-013	MN594377
Bartonella-047	MN594398	Leptospira-001	MN594314	Pasteurellaceae-015	MN594386
Bartonella-052	MN594401	Leptospira-002	MN594336	Pasteurellaceae-016	MN594393
Bartonella-055	MN594402	Leptospira-003	MN594367	Pasteurellaceae-017	MN594394
Bartonella-056	MN594404	Leptospira-004	MN594361	Pasteurellaceae-018	MN594403
Bartonella-057	MN594406	Leptospira-005	MN594363	Pasteurellaceae-020	MN594420
Bartonella-058	MN594413	Leptospira-012	MN594371	Pasteurellaceae-022	MN594433
Bartonella-059	MN594414	Leptospira-023	MN594396	Pasteurellaceae-023	MN594434
Bartonella-060	MN594415	Leptospira-024	MN594397	Pasteurellaceae-024	MN594435
Bartonella-063	MN594422	Leptospira-029	MN594418	Pasteurellaceae-026	MN594436
Bartonella-066	MN594423	Leptospira-030	MN594421	Pasteurellaceae-027	MN594437
Bartonella-067	MN594424	Leptospira-031	MN594458	Pasteurellaceae-028	MN594439
Bartonella-068	MN594425	Leptospira-035	MN594462	Pasteurellaceae-031	MN594441
Bartonella-075	MN594426	Leptospira-036	MN594463	Pasteurellaceae-033	MN594442
Bartonella-076	MN594427	Leptospira-039	MN594464	Pasteurellaceae-034	MN594443
Bartonella-079	MN594428	Leptospira-042	MN594465	Pasteurellaceae-040	MN594444
Bartonella-084	MN594429	Mycobacterium-001	MN594325	Pasteurellaceae-041	MN594445
Bartonella-087	MN594430	Mycobacterium-002	MN594329	Pasteurellaceae-045	MN594446
Bartonella-088	MN594431	Mycoplasma-001	MN594313	Pasteurellaceae-046	MN594447
Bartonella-090	MN594432	Mycoplasma-002	MN594315	Rickettsia-001	MN594322
Bordetella-001	MN594316	Mycoplasma-003	MN594332	Ureaplasma-001	MN594330
Bordetella-002	MN594340	Mycoplasma-004	MN594333	Yersinia	MN594327
Bordetella-003	MN594349	Mycoplasma-005	MN594334		
Bordetella-004	MN594359	Mycoplasma-006	MN594324		
Bordetella-005	MN594395	Mycoplasma-007	MN594331		
Bordetella-006	MN594342	Mycoplasma-008	MN594352		
Bordetella-008	MN594384	Mycoplasma-009	MN594326		
Bordetella-010	MN594399	Mycoplasma-010	MN594348		
Bordetella-011	MN594405	Mycoplasma-012	MN594360		
Bordetella-013	MN594459	Mycoplasma-016	MN594335		
Bordetella-015	MN594460	Mycoplasma-018	MN594362		
Bordetella-016	MN594461	Mycoplasma-019	MN594369		

90 13 R Packages

The following R packages were used in the analyses and preparation of the manuscript:

- ape Paradis, E. and Schliep, K. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. – *Bioinformatics* 35: 526–528
- cooccur Griffith, D. M. et al. 2016. cooccur: Probabilistic species co-occurrence analysis in R. – *Journal of Statistical Software, Code Snippets* 69: 1–17
- dplyr Wickham, H. et al. 2020. dplyr: A Grammar of Data Manipulation
- dunn.test Dinno, A. 2017. dunn.test: Dunn’s Test of Multiple Comparisons Using Rank Sums
- ecodist Goslee, S. C. and Urban, D. L. 2007. The ecodist package for dissimilarity-based analysis of ecological data. – *Journal of Statistical Software* 22: 1–19
- egg Auguie, B. 2019. egg: Extensions for ‘ggplot2’: Custom Geom, Custom Themes, Plot Alignment, Labelled Panels, Symmetric Scales, and Fixed Panel Size
- emmeans Lenth, R. 2020. emmeans: Estimated Marginal Means, aka Least-Squares Means
- english Fox, J. et al. 2020. english: Translate Integers into English
- ggdendro de Vries, A. and Ripley, B. D. 2016. ggdendro: Create Dendrograms and Tree Diagrams Using ‘ggplot2’
- ggplot2 Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. – Springer-Verlag New York
- ggrepel Slowikowski, K. 2020. ggrepel: Automatically Position Non-Overlapping Text Labels with ‘ggplot2’
- ggsn Santos Baquero, O. 2019. ggsn: North Symbols and Scale Bars for Maps Created with ‘ggplot2’ or ‘ggmap’
- gtable Wickham, H. and Pedersen, T. L. 2019. gtable: Arrange ‘Grobs’ in Tables
- iNEXT Hsieh, T. C. et al. 2020. iNEXT: Interpolation and Extrapolation for Species Diversity
- lubridate Grolemund, G. and Wickham, H. 2011. Dates and times made easy with lubridate. – *Journal of Statistical Software* 40: 1–25
- MASS Venables, W. N. and Ripley, B. D. 2002. Modern Applied Statistics with S. Fourth edn. – Springer, New York
- multcomp Hothorn, T. et al. 2008. Simultaneous inference in general parametric models. – *Biometrical Journal* 50: 346–363
- plyr Wickham, H. 2011. The split-apply-combine strategy for data analysis. – *Journal of Statistical Software* 40: 1–29

purrr Henry, L. and Wickham, H. 2020. purrr: Functional Programming Tools
 raster Hijmans, R. J. 2020. raster: Geographic Data Analysis and Modeling
 rcompanion Mangiafico, S. 2020. rcompanion: Functions to Support Extension Education Program Evaluation
 readr Wickham, H. et al. 2018. readr: Read Rectangular Text Data
 rgdal Bivand, R. et al. 2019. rgdal: Bindings for the 'Geospatial' Data Abstraction Library
 rgeos Bivand, R. and Rundel, C. 2019. rgeos: Interface to Geometry Engine - Open Source ('GEOS')
 sp Bivand, R. S. et al. 2013. Applied spatial data analysis with R, Second edition. – Springer, NY
 stplanr Robin Lovelace and Richard Ellison 2018. stplanr: A Package for Transport Planning. – The R Journal 10
 stringr Wickham, H. 2019. stringr: Simple, Consistent Wrappers for Common String Operations
 tibble Mller, K. and Wickham, H. 2020. tibble: Simple Data Frames
 xtable Dahl, D. B. et al. 2019. xtable: Export Tables to LaTeX or HTML

References

- [1] Mahé, F., Rognes, T., Quince, C., de Vargas, C., Dunthorn, M.: Swarm: robust and
95 fast clustering method for amplicon-based studies. *PeerJ* **2**, 593 (2014)
- [2] Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glockner, F.O., Yarza, P.: The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* **41**, 590–596 (2013). doi:10.1093/nar/gks1219
- 100 [3] Madeira, F., Park, Y.M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., Basutkar, P., Tivey, A.R., Potter, S.C., Finn, R.D., *et al.*: The embl-ebi search and sequence analysis tools apis in 2019. *Nucleic acids research* **47**(W1), 636–641 (2019)
- [4] Adeolu, M., Gupta, R.S.: A phylogenomic and molecular marker based proposal for the division of the genus *borrelia* into two genera: the emended genus *borrelia* containing only the members of the relapsing fever *borrelia*, and the genus *borreliella* gen. nov. containing the members of the lyme disease *borrelia* (*borrelia burgdorferi* sensu lato complex). *Antonie Van Leeuwenhoek* **105**(6), 1049–1072
105 (2014)
- [5] Olsen, I., Dewhirst, F.E., Paster, B.J., Busse, H.: Pasteurellaceae. In: Whitman, W.B., Rainey, F., Kämpfer, P., Trujillo, M., Chun, J., DeVos, P., Hedlund, B., Dedysh, S. (eds.) *Bergey’s Manual of Systematics of Archaea and Bacteria*, pp. 1–9. John Wiley & Sons, Inc. in association with Bergey’s Manual Trust,, ??? (2015).
110 doi:10.1002/9781118960608.fbm00230