Unveiling genome plasticity and a novel phage in *Mycoplasma felis*: genomic investigations of four feline isolates

Sara M. Klose[1,2*], Alistair R. Legione[1*], Rhys N. Bushell[3], Glenn F. Browning[1] and Paola K. Vaz[1]

*S.M.K. and A.R.L. contributed equally to this work.

Affiliation:

1. Asia-Pacific Centre for Animal Health, Melbourne Veterinary School, Faculty of Science, The University of Melbourne, VIC, Australia

2. Institute of Molecular Medicine and Experimental Immunology, Faculty of Medicine, University of Bonn, NRW, Germany

3. Department of Veterinary Clinical Sciences, Melbourne Veterinary School, Faculty of Science, The University of Melbourne, VIC, Australia

Corresponding Author: Alistair Legione, legionea@unimelb.edu.au

Keywords: *Mycoplasma felis*, whole genome sequencing, genome plasticity, evolution, phage

Repositories: All sequence data and assemblies for data associated with this project can be found under the GenBank BioProject: PRJNA906261

Abstract:

Mycoplasma felis has been isolated from diseased cats and horses, but to date only a single fully assembled genome of this species, of an isolate from a horse, has been characterised. This study aimed to characterise and compare the completely assembled genomes of four clinical isolates of *M. felis* from three domestic cats, assembled with the aid of short and long read sequencing methods. The completed genomes encoded a median of 759 open reading frames (min, 743, max 777) and had a median average nucleotide identity (ANI) of 98.2% with the genome of the available equid origin reference strain. Comparative genomic analysis revealed the occurrence of multiple horizontal gene transfer (HGT) events and significant genome reassortment. This had resulted in the acquisition or loss of numerous genes within the Australian felid isolate genomes, encoding putative proteins involved in DNA transfer, metabolism, DNA replication, host cell interaction, and restriction modification systems. Additionally, a novel mycoplasma phage was detected in one Australian felid *M. felis* isolate by genomic analysis and visualised using cryo-transmission electron microscopy. This study has highlighted the complex genomic dynamics in different host environments. Furthermore, the sequences obtained in this work will enable the development of new diagnostic tools, and identification of future infection control and treatment options for the respiratory disease complex in cats.

Data summary:

All genome data for this study have been deposited in GenBank under BioProject PRJNA906261. Genome assemblies, as well as Illumina and Oxford Nanopore sequence reads for each isolate, can be found under their respective BioSamples: SAMN32182834 (isolate 047), SAMN32182835 (isolate 219), SAMN32182836 (isolate 329 and associated phage), and SAMN32182837 (isolate 632). The authors confirm all supporting data and protocols have been provided within the article.

Impact statement:

Mycoplasma felis is commonly associated with clinical cases of conjunctivitis and feline respiratory disease complex in cats, the leading cause of euthanasia in animal shelters. In the absence of vaccines, infection control is currently limited to the prolonged treatment with antimicrobials. Prior to this study there was only one complete genome assembly of an isolate of *M. felis*, which had been obtained from a horse. This study has provided the first high quality hybrid assembled genomes of

50 *M. felis* isolates from cats. This work adds four new genomes from clinical cases, as well as the
51 identification and validation of the presence of a novel phage that utilises the mycoplasma
52 translation code. The genomic data presented here can assist future projects investigating improved
53 diagnostics and development of new treatment options for this significant feline pathogen.
54

55 ## Introduction:

56 Mycoplasmas are the smallest free-living bacteria and have some of the smallest bacterial genomes
57 (0.58 – 1.38 Mb). They are thought to approach the minimal gene sets essential for independent life,
58 making them ideal model organisms for studies on the fundamental requirements of a living cell [1].
59 They evolved from Gram positive bacterial ancestors 600 million years ago [2] and have undergone
60 considerable genome reduction during their evolution [3, 4]. They lack a cell wall, are pleomorphic
61 and can infect and cause acute and chronic disease in a diverse range of animal and plant species [5].
62
63 Some mycoplasmas have been found to exhibit significant genetic plasticity and to evolve rapidly,
64 mainly driven by intra-species spontaneous mutation and recombination or horizontal gene transfer
65 (HGT) between and within species, facilitating swift adaptation to environmental changes, enhancing
66 survival and/or virulence [6-9]. Studies examining the genomic stability of *Mycoplasma synoviae* have
67 revealed that mutations occur frequently during infection in birds [10], and that rapid
68 thermoadaptive evolution can occur *in vitro* and *in vivo* [11], presumably to regain fitness and
69 pathogenicity [12, 13]. HGT appears to occur more frequently between species that share an
70 ecological niche, irrespective of their phylogenetic divergence. Examples of significant HGT include
71 the transfer of the genes encoding the phase variable VlhA cell surface lipoproteins between the
72 poultry pathogens *Mycoplasma gallisepticum* and *M. synoviae* [14], and the transfer of several cell
73 surface lipoproteins involved in host colonisation between the ruminant pathogens *Mycoplasma*
74 *capricolum* subsp. *capricolum* and *Mycoplasma mycoides* subsp. *Mycoides*, and *Mycoplasma*
75 *agalactiae* and *Mycoplasma bovis* [15].
76
77 At least four species of mycoplasmas have been isolated from cats (*Mycoplasma felis*, *Mycoplasma*
78 *gateae*, *Mycoplasma arginini*, and *Mycoplasma feliminutum*), in most cases as commensals. *M. felis*
79 in particular is primarily an opportunistic pathogen, able to infect and be isolated from the upper
80 respiratory tract. It has been associated with conjunctivitis and, increasingly, feline respiratory

disease [16-19]. Less commonly, *M. felis* has also been associated with respiratory disease in horses [20] although it is unclear how bacterial infection mechanisms and pathogenesis may vary between hosts as infection by *M. felis* is under-explored in both cats and horses. Feline respiratory disease complex (FRDC) is the leading cause of euthanasia of cats in animal shelters [21] and in addition to *M. felis* is also commonly associated with coinfections with several viral (feline herpesvirus 1 and feline calicivirus) and bacterial pathogens (*Chlamydia felis*) [22, 23]. *M. felis* infections are common and difficult to eradicate in most cat populations, and the ability of this organism to persist in an animal, likely due to intracellular invasion, enhances opportunities for recurrent opportunistic infections. Vaccines against *M. felis* are not currently available and treatment of bacterial infection is limited to prolonged treatment of up to eight weeks with tetracyclines or fluoroquinolones, which is arguably a risk for antimicrobial resistance (AMR) development [24, 25]. Currently, full genomes of three isolates of this species have been published: strain Myco-2 (GenBank accession number: AP022325.1), an isolate from a case of respiratory disease in a horse in Japan [26], which has been completely assembled, and multiple assemblies of isolates 16-040612-4 (CP104192, CP104193, CP104194, CP11027) and 16-057065-4 (CP103988, CP103993, CP110269), which were each isolated from feline bronchoalveolar lavages in Canada [27]. Additionally, there is an incomplete genome of ATCC strain 23391, assembled into scaffolds (GenBank Biosample: SAMN02841218).

The aim of this study was to sequence and assemble four high quality, complete *M. felis* genomes, using hybrid assembly methods. These genomes represent the first complete assemblies of feline isolates of this species. The objective was to significantly expand on our limited understanding of the genomic structure and plasticity of *M. felis*, which may lead to the development of improved diagnostic tests and vaccines, as has been seen in other mycoplasma species.

Methods:

Mycoplasma isolation

Swabs were collected for bacteriological culture from clinical cases of respiratory disease (3 cases) and an unresolved post-surgical infection (1 case) (Table 1). Swabs were used to inoculate sheep blood agar (SBA) plates, which were incubated aerobically at 37°C. Following observation and presumptive identification of small mycoplasma-like colonies, single colonies were transferred to Mycoplasma Broth (MB) containing 10% swine serum (Sigma-Australia) and 0.01% nicotinamide

112 adenine dinucleotide (NAD) (Sigma-Australia), based on the formulation of Frey's medium, with
113 minor modifications [28]. Acidification of the MB medium was used to confirm the growth of
114 *Mycoplasma* spp. and the cultures were stored at -80°C for future investigation.
115
116 *DNA extraction, sequencing, and genome assembly*
117 Mycoplasma isolates were cultured axenically in MB for 18 h at 37°C prior to DNA extraction. Bacteria
118 were pelleted and the medium removed, and then resuspended in the appropriate extraction kit lysis
119 buffer. The DNA for short read sequencing was extracted using the DNeasy Blood and Tissue Kit
120 (QIAGEN), following the manufacturer's protocol, and 100 ng of the extracted DNA was used to
121 prepare sequencing libraries using Illumina's Nextera Flex DNA Library Prep Kit. Sequencing was
122 performed on the NovaSeq platform at Charles River Laboratories, Victoria, Australia, to generate
123 150 bp paired-end reads. Two rounds of Oxford Nanopore Technologies (ONT) long read sequencing
124 were performed. The DNA was extracted using the Promega Wizard DNA Purification Kit (Round 1)
125 or the Promega Wizard High Molecular Weight DNA Purification Kit (Round 2). DNA libraries were
126 generated using the ONT Rapid Barcoding (SQK-RBK004) Kit. Sequencing was performed on a MinION
127 Mk1b device fitted with a FLO-MIN106D Flow Cell (R9.4.1 chemistry). Base calling, de-multiplexing,
128 and adapter removal were performed using the Guppy (version 6.2.1+6588110a6) high accuracy base
129 caller with a minimum quality score of 9 (default settings). The Illumina reads were trimmed and
130 adaptors removed with fastp v0.20.1 [29] using the default settings. The ONT reads from the two
131 rounds were merged for each isolate and processed using Filtlong v0.2.0 [30] to remove all reads
132 shorter than 1000 bp, retaining 90% of input data (after removal of short and low-quality reads).
133 Subsequently, the processed short and long reads were used as input into Unicycler v0.5.0 [31] in
134 'normal' mode, using default settings besides a modified start gene database that included *M. felis*
135 and a minimum contig length of 200 bp. The assemblies where then polished using Polypolish v0.5.0
136 [32] using default settings, incorporating BWA-MEM v0.7.17-r1188 [33]. Full genome alignments
137 were generated using the progressive MAUVE algorithm [34] to identify regions of reassortment and
138 gene loss or gain.
139
140 *Genome annotation and sequencing analysis*
141 The genomes were annotated using the Prokaryotic Genome Annotation Pipeline v2022-12-
142 13.build6494 [35] for primary analysis. A *Tenericutes* specific protein database, containing both
143 Swiss-Prot (4464 sequences) and TrEMBL (343,598 sequences) annotations, was downloaded from

144 UniProt and converted to a Diamond format protein database [36]. This database was used as input
145 to Pydeel (https://github.com/alegione/pydeel) to determine the coding ratio of our annotated
146 genomes in comparison with *M. felis* genomes on GenBank. This tool identifies the coding ratio of
147 annotated genes compared to their best hit within the provided database to identify the proportion
148 of annotated genes that may have erroneous insertions or deletions as a by-product of the use of
149 long read sequencing. From here, annotated genomes were processed through Panaroo v1.3.0 [37],
150 along with previously sequenced *M. felis* genomes (Myco-2 and ATCC23391), to identify the core
151 genome. The core genome results were visualised using Pagoo [38]. Gene alignments generated by
152 Panaroo were filtered based on their presence in at least four genomes and present as a single copy
153 per genome before analysis for genetic diversity using DnaSP v6.12.03 [39]. The annotated genomes
154 from this study were investigated for pathogenicity islands and integrative conjugative elements
155 using Island Viewer 4 [40] and ICE finder [41], respectively. Antimicrobial resistance genes and
156 virulence factors were detected by generating a Diamond database of the CARD [42] and full
157 virulence factor databases [43], and using Diamond (v 2.1.9) BLASTp to identify best hits with a cut-
158 off bit score of 50 (all other parameters were left as their default). The UniProt webserver was used
159 to identify homologs of specific open reading frames (ORFs) in other genomes
160 (https://www.uniprot.org).
161
162 *Phage isolation and visualisation*
163 A phage-like contig was identified in the annotation results of one of the assemblies (MF329). In
164 order to confirm whether the detection of a bacteriophage genomic sequence corresponded to the
165 presence of an integrated prophage or to replicative virions, bioinformatic methods were first
166 utilised. The probable terminal ends were investigated using short and long read mapping to the
167 contig with BWA-MEM and minimap2 v2.26-r1175 [44], respectively, and the contig's starting point
168 adjusted based on the depth of coverage. The tool PhageTerm [45] was also used to investigate the
169 potential phage start and end points. To investigate whether this contig was a prophage, the long
170 reads were mapped, using minimap2, to the updated phage contig. Any reads that had overhangs at
171 the 5' and 3' ends of the contig were extracted and subsequently mapped to the assembled MF329
172 genome to determine whether there was a common point of overlap. To detect the presence of
173 bacteriophage virions, transmission electron microscopy was performed on a culture of MF329.
174 Isolate MF329 was cultured in 50 mL of Mycoplasma Broth, as described above, for 20 hours. The
175 culture was centrifuged for 15 mins at 10,000 x g at 4°C and the medium fraction separated from the

176 bacterial cell pellet. The bacterial cells were gently lysed by resuspension in 150 µL of sterile water,
177 followed by centrifugation at 10,000 x g at 4°C for 5 mins to remove cellular debris. This was repeated
178 twice, with the supernatant fraction captured each time for ultracentrifugation. The culture medium
179 and the cell lysate supernatant fractions were separately ultra-centrifugated at 100,000 x g for 90
180 min at 4°C, and the pellets resuspended in 0.2 M HEPES buffer, pH 7.7. Negatively stained samples
181 were examined at the Ian Holmes Imaging Centre (Melbourne, Australia) using a FEI Talos L120C
182 cryoTEM (Talos L120C) microscope.
183

184 **Results:**

185 *Characterisation and comparisons of genomes*

186 All four isolates were successfully assembled into circularised, complete genomes. In addition to
187 these four genomes, a novel *M. felis* phage genome was identified that utilised the mycoplasma
188 translation code. Metrics for long and short read sequencing output are provided in Supplementary
189 Table 1. The assembly statistics for the four isolates are provided in Supplementary Table 2. Briefly,
190 the circularised isolates were significantly larger than the reference genome (Myco-2; 841,695 bp),
191 with a median length of 940,935 bp (range: 905,741 – 948,716) and a median of 759 coding domain
192 sequences (CDSs) (range: 743 – 777), compared to the reference genome which has 740 CDSs. The
193 NCBI Prokaryotic Genome Annotation Pipeline incorporates a CheckM [46] completeness analysis
194 and an average nucleotide identity taxonomic classification step to identify any potential issues with
195 an assembled genome. For all four genomes the CheckM completeness mirrored that of the
196 reference Myco-2 genome (99.21% completeness), and all assemblies were classified as *Mycoplasma*
197 *felis* with high confidence (Supplementary Table 2). The coding ratios of annotated assemblies were
198 investigated to determine the frequency of potential indel errors. The two published *M. felis*
199 genomes (Myco-2 and ATCC23391) had mean coding ratios of 1.0 (standard deviation ±0.02) and 0.96
200 (±0.32), respectively. Our completed genomes, MF47, MF219, MF329, and MF632, had mean coding
201 ratios of 1.0 (±0.28), 0.99 (±0.30), 1.0 (±0.27), and 1.0 (±0.26), respectively. Seven recently sequenced
202 *M. felis* isolates that utilised only Oxford Nanopore sequencing had a substantial number of indel
203 errors in their annotations based on coding ratio analysis (mean coding ratios between 0.44 and 0.47)
204 and were not used in further comparisons for this reason (Supplementary Figure 1).
205

206 A progressive MAUVE alignment of the four genome assemblies in comparison to that of Myco-2
207 (Figure 1) identified two significant genome inversions in the reference sequence compared to the
208 isolates assembled here, as well as more localised inversions within each of the isolates. These
209 inversions, and the genomic regions present in individual isolates, were typically flanked by
210 transposases. Comparison of the average nucleotide identities (ANIs) of the four new assemblies,
211 and with previously published *M. felis* genomes (Supplementary Table 3) found that the best match
212 for each of the four Australian isolates was another isolate from this study, with the exception of
213 MF329, noting that MF047 and MF219 were obtained from the same animal two months apart.
214
215 The core genome of *M. felis*, as determined by Panaroo (genes present in >99% of genomes)
216 contained 594 genes, with 316 shell genes (genes present in 15% - 95% of genomes). Collapsing
217 paralogues reduced the number of shell genes to 237. Principal component analysis (PCA) highlighted
218 clustering of the two published genomes, as well as clustering of isolates that were obtained from
219 the same individual (Figure 2). All genes, except those annotated as hypothetical proteins, identified
220 in the newly sequenced isolates but not present in the reference sequence are listed in Table 2, with
221 the homologs found in other genomes. The variations between the genomes were mostly observed
222 in sequences encoding putative proteins involved in DNA transfer, metabolism, DNA replication,
223 mycoplasma-host cell interactions, and restriction modification systems (Figure 3). At a gene level,
224 an average of 2.41% of each coding sequence was a segregation site, and an average nucleotide
225 identity per gene of 98.8% (Supplementary Table 4). Only four genes were identified as having a
226 significant Tajima's D value by DnaSP. These genes encoded for two hypothetical proteins (Myco-2
227 loci 000416 and 000586), 23S rRNA (pseudouridine^{1915}-*N*^{3})-methyltransferase (Myco-2 locus 000375),
228 and a Sua5/YciO/YrdC/YwlC family protein (Myco-2 locus 000629). Each of these coding regions had
229 a Tajima's D value below 1, suggesting neutral selection (Supplementary Table 4).
230
231 The genomes of the newly sequenced isolates had 1-3 secretory systems (Type IV secretory systems),
232 whilst that of Myco-2 has no annotated secretory systems. The newly annotated genomes also had
233 an increased number of transposases, typically annotated as IS*1634* or IS*256*, or as IS*3*. In
234 comparison, the Myco-2 genome only had IS*3* and IS*1634* family transposase annotations.
235
236 <u>*Genomic islands, virulence genes and antimicrobial resistance genes*</u>

237 Genomic islands were detected in three of the four newly sequenced genomes (**Error! Reference**
238 **source not found.**). The average length of these islands was 22,145 bp (± 5818 bp) and each
239 contained an average of 16 (±5) annotated genes. In comparison, the Myco-2 strain of *M. felis* had
240 two predicted genomic islands of 3052 bp and 10,027 bp. Further inspection of these predicted
241 genomic islands identified matching regions present in isolates in alternative locations. For example,
242 MF329 had a predicted genomic island of 24 kb (nucleotide position [nt] 311,409 – 335,810)
243 containing 19 genes, including a type IV secretion system and an insertion sequence (IS*1634*) at its
244 midpoint. Extraction and remapping of the sequence of this genomic island showed that it was
245 repeated three times within a wider 122 kb region (nt 213,995 – 335,810) of the MF329 genome.
246 This same genomic region occurred, without the IS element present, in the genomes of MF047
247 (twice), MF219 (once), and MF632 (once). In the case of MF632, the region was also a predicted to
248 be a genomic island, but in this instance was flanked with restriction endonuclease genes and either
249 predicted transposases (3') or an IS*256* element (5').
250 Diamond BLAST searches for antimicrobial resistance genes (CARD database) and virulence genes
251 (Virulence Factor Database) identified few high identity hits (Table 4 and Supplementary Tables 5 and
252 6). Only one hit with the virulence factor database had an amino acid sequence identity of greater
253 than 90%, and this was to the gene for the elongation factor Tu in *Mycoplasma synoviae* (90.1%
254 average identity), which has been found to be an adherence factor and is present in all genomes in
255 this study. Only five genes had hits with >50% amino acid sequence identity. The closest hits to these
256 genes were *oppF* (59.9% identity; *Mycoplasma mycoides* subsp. *mycoides*; immune modulation), *eno*
257 (58.9%; *Streptococcus pneumoniae*; exoenzyme), *gap* (57.9%; *Mycoplasma genitalium*; adherence),
258 *dnaK* (53.9%; *Chlamydia trachomatis*; adherence), and *sugC* (50.4%; *Mycobacterium marinum*;
259 nutritional/metabolic factor). No genes had hits with >50% amino acid sequence identity with
260 antimicrobial resistance genes in the CARD database.
261
262 *Phage detection and identification*
263 A 33.6 kb contig (vB Mfe PM329; *Mycoplasma felis* strain PM329) was assembled and identified as a
264 likely phage or prophage in the sequenced *M. felis* MF329 isolate. It had an average depth of coverage
265 of 40,847.7 reads, with 82.6% of the MF329 Illumina reads mapping to the phage genome.
266 Interrogation of the Nanopore read dataset to determine its genome integration status detected only
267 63/18134 (0.35%) reads that mapped to both the phage and the bacterial genome. Of these, most
268 were found to map across the bacterial genome at mostly single-read depth, with the exception of

269 five reads that clustered mid-genome in the region of nt 456,500. If this was indicative of a true
270 prophage insertion site it would disrupt ORF MF329_000403, which encodes a putative 16S rRNA
271 pseudouridine synthase gene. The phage genome had a 27.1% G+C content and encoded 36
272 predicted ORFs using the mycoplasma translation code. It shared 94.8% global pairwise nucleotide
273 sequence identity with two contigs in the publicly available unassembled *M. felis* contig dataset (NCBI
274 Biosample: SAMN30346842), and there was 96.8% and 97.4% pairwise amino acid sequence identity
275 between each of their encoded putative recombinase and portal proteins, respectively. Interrogation
276 of the read datasets of the other three isolates sequenced in this study with the phage genome
277 identified only 4/9515 (0.04%; MF047), 19/19128 (0.1%; MF632) and 22/65383 (0.03%; MF219)
278 mapped Nanopore reads, within the range of barcode error. No Illumina reads mapped to the phage
279 genome, indicating that the isolates did not carry the phage. Phylogenetic analyses using VipTree
280 [47] and vContact2 [48] was not able to assign vB Mfe PM329 to any known phage genus or family,
281 but its gene arrangement was similar to those of the MAgV1 phage from *M. agalactiae* and MAgV1-
282 like prophage sequences detected in mycoplasma species in the Hominis phylogenetic group. In
283 particular, both the protein and gene arrangement similarity were greatest to those of prophages
284 found in *Mycoplasma mustelae* and *Mycoplasma molare*, including the central position of the
285 recombinase gene. BLASTX analysis of predicted ORFs detected pairwise protein sequence identities
286 of 48% to 69% between core conserved ORFs encoded within the *M. mustelae* and *M. molare*
287 prophage sequences [49] (Supplementary Table 7).
288
289 Confirmation of productive virus formation and, therefore, active replication, was achieved using
290 cryoTEM on medium and cell lysate supernatant fractions, which revealed the presence of non-
291 enveloped contractile-tailed polyhedral viral particles approximately 100 nm in length, with a 50 nm
292 diameter head and tails of varying lengths. Higher numbers of virions were observed within the cell
293 lysate supernatant fraction than in the culture medium fraction, with clusters of virions observed
294 seemingly embedded within bacterial membrane debris (Figure 4). These long tailed polyhedral viral
295 structures were consistent with those seen in siphoviruses and myoviruses within the recently
296 ratified *Caudoviricetes* class of bacteriophage. The contractile tail and polyhedral head structure is
297 consistent across all visualised mycoplasma phages, with the main differences being the head
298 diameter and tail lengths [50][51].
299

300 ## Discussion

301 In this study, the genomes of four new *M. felis* isolates were characterised and compared; three of
302 the isolates were derived from clinical cases of respiratory disease and the remaining one from an
303 unresolved post-surgical infection. Comparisons between these genomes and with several published
304 datasets derived from other isolates of *M. felis* detected genomic variations and inversions between
305 the Australian felid isolates and the only assembled *M. felis* reference isolate, from a horse. These
306 variations and inversions were flanked by transposases, which are commonly involved in HGT and
307 genome reassortment in other mycoplasmas. There were differences in the number of copies of
308 putative transposases from several families, including the IS*1634*, IS*256* and IS*3* (Figure 3). Insertion
309 sequences (ISs) are transposable elements that are often present in multiple copies in a genome and
310 involved in chromosomal rearrangements by deletion, insertion, or inversion. The substantial
311 similarity between various isoforms of ISs enables them to facilitate homologous recombination [52].
312 IS families encoded in mycoplasma genomes play an important role in genomic plasticity with several
313 reported instances of IS exchange between strains or species [15, 53-56]. This is in contrast to early
314 research that suggested that some ISs, such as IS*1634*, which was found in abundance in the
315 sequenced genomes of our isolates, were unique to specific mycoplasma species [57]. The
316 differences between our Australian *M. felis* genomes, which were obtained from three different cats,
317 based on gene presence/absence and the number of coding sequences, underscores the genetic
318 diversity present amongst the sequenced felid *M. felis* isolates (Figure 2). Multiple gene insertions,
319 relative to the equid *M. felis* reference genome, were detected within the cluster of Australian field
320 isolates (Table 2). These included genes involved in DNA transfer, replication, metabolism, and
321 restriction modification (RM) systems, as well as a number of proteins of unknown function (Figure
322 3).
323
324 *Gene gains compared to the equid M. felis reference strain*
325 Of particular interest are the gene acquisitions involved in DNA transfer, such as the multiple copies
326 of genes encoding putative type IV secretory system (T4SS) conjugative DNA transfer family proteins
327 within the Australian felid genomes, which were not detected in the publicly available equid and felid
328 *M. felis* genomes. T4SS conjugation systems translocate DNA through a process dependent on cell to
329 cell contact and have been identified as one of the core genes within mycoplasma Integrative and
330 Conjugative Elements (ICEs), which play a significant role in genomic plasticity in the species in which

331 they have been found [58, 59]. ICEs have been shown to be self-transmissible in *M. agalactiae*,
332 providing conjugative properties from ICE-positive cells to ICE-negative cells, while concurrent
333 chromosomal transfers (CTs) occur in the opposite direction [60]. Protein sequence analysis of the
334 putative T4SS identified in the Australian felid *M. felis* isolates detected 43.2% amino acid sequence
335 identity with ORF5 of the ICE in *Mycoplasma cynos*. Whilst *M. cynos* predominantly infects the
336 respiratory tract of dogs, it has also been detected in cats in association with conjunctivitis and upper
337 respiratory tract disease [61]. Taken together, it is plausible that a potential HGT event may have
338 resulted in transfer of the T4SS conjugative element of an ICE from a *M. cynos*-like species to *M. felis*
339 in a shared feline host. This T4SS was part of a defined region of genes that, along with 15
340 hypothetical proteins and a DUF87 domain-containing protein, was repeated three times in the
341 genome of isolate MF329. DUF87 domain-containing proteins are associated with conjugative
342 elements in other organisms, and include TraJ in the virulence plasmid of *Salmonella enterica* serovar
343 Typhimurium [62], adding further weight to the likelihood that this region is an ICE, even though the
344 bioinformatics tools used here did not detect it.
345
346 Several genes involved in metabolism were newly detected in the felid *M. felis* isolates sequenced in
347 this study. A type 2 glycerol-3-phosphate oxidase (GlpO) was identified within the Australian *M. felis*
348 genomes, while in the equid *M. felis* isolate, only its FAD-dependent oxidoreductase domain was
349 detected, accompanied by two transposases, located upstream. GlpO converts glycerol-3-phosphate
350 to dihydroxyacetone phosphate (DHAP), releasing hydrogen peroxide as a by-product, which has
351 been shown to play a role in cytotoxicity *in vitro* in *M. gallisepticum* and *Mycoplasma pneumoniae*
352 [63, 64], and in the virulence of *M. mycoides* subsp. *mycoides* small colony type and mosquito-
353 associated *Spiroplasma* species [65, 66]. The significant reduction or complete loss of H_2O_2
354 production in highly passaged *M. bovis*, *M. agalactiae* and *M. gallisepticum* strains, and in attenuated
355 vaccine strains of *M. gallisepticum*, compared to wild type strains suggests that glycerol metabolism
356 can rapidly become redundant in some mycoplasmas [63, 67, 68]. It is possible that the sequences
357 coding for GlpO were lost in the equid *M. felis* isolate, possibly because of differing levels of available
358 glycerol in horses compared to cats. Likewise, variations were seen in the counts of genes coding for
359 nutrient transport systems, including ATP-binding cassette (ABC) transporters and the PEP-
360 dependent phosphotransferase transport system (PTS), as well as numerous other metabolic
361 enzymes. A number of these nutrient transport systems, and many cell surface enzymes, are
362 multifunctional in mycoplasmas, playing critical roles in pathogenesis and survival in the host [13, 69-

363 74]. The differences observed in the PTS and ABC transporters of the Australian felid *M. felis* isolates,
364 in contrast to the equid *M. felis* isolate, may also be indicative of rapid metabolic gene gains and
365 losses in response to the varying nutrient availabilities within felid and equid ecological niches.
366
367 In the Australian felid *M. felis* genomes, five new coding sequences related to replication were
368 identified that were absent in the equid origin *M. felis* genome, including genes encoding the putative
369 anti-termination protein NusB, a Hsp70 family protein, a C5 methylase, the 16S rRNA
370 methyltransferase RsmG, and an HNH endonuclease. The presence of these additional proteins
371 suggests potential differences in bacterial growth within the feline and equine hosts, potentially due
372 to distinct replication strategies. Future experiments comparing growth patterns of these isolates in
373 felid or equid derived cells are indicated. The potential origins of these additional sequences have
374 been detected in mycoplasmas isolated from mink, pigs, and dogs, which share a common ecological
375 environment with cats, or pigeons, which are a common prey of cats. The HGT events resulting in the
376 acquisition of these genes by the Australian felid *M. felis* isolates may have been facilitated by the
377 close ecological proximity or may have occurred within the feline gastrointestinal tract (GI) following
378 the ingestion of infected prey. HGT is frequently observed between diverse microorganisms within
379 the mammalian gastrointestinal tract, possibly due to the high population density in this niche, which
380 presumably facilitates genetic exchange [75]. Furthermore, the conjugation mechanism favoured for
381 HGT within the GI tract protects DNA from nucleases and heavy metals, and thus is more likely to
382 result in a successful genetic exchange [76].
383
384 Two additional restriction endonucleases (Eco47II and Sau3AI) were annotated in the Australian felid
385 *M. felis* genomes, and they lost a DpnII family type II restriction endonuclease, compared to the equid
386 isolate. Restriction modification (RM) systems play a pivotal role in regulating bacterial HGT by
387 protecting the genome against incoming mobile genetic elements, or by facilitating recombination
388 with the introduced DNA [77]. A strong correlation has been reported between the presence of active
389 Type II RM systems and the presence of mobile genetic elements and ICEs within *M. agalactiae*
390 genomes, indicating the facilitative role of Type II RM systems in HGT in mycoplasmas [78]. The
391 diversification of RM target recognition sites can regulate HGT by promoting genetic exchanges
392 among those organisms with RM systems that recognise the same target motifs, and by reducing
393 genetic transfer between lineages that possess distinct RM systems [79]. Taken together, it is

394 plausible that following a host switch into horses, *M. felis* may have diversified its RM systems to
395 enhance the level of genetic exchange from phylogenetically distant bacteria.
396
397 *Gene losses compared to the equid M. felis reference strain*
398 A sequence encoding a putative ComEC/Rec2 family competence protein was absent in the
399 Australian felid *M. felis* isolates. This protein is one of the components of DNA translocation via
400 transformation, enabling DNA transport by forming a channel across the cytoplasmic membrane [80].
401 While the ComEC component has been reported to be essential for DNA transport in Gram positive
402 bacteria, deletion of homologous genes did not affect DNA uptake in bacteria without a cell wall [81].
403 Furthermore, the competence domain seems to be missing in many mycoplasmas, possibly due to
404 the reductive evolution of these bacteria [82]. Characterisation of two mycoplasma genomes
405 recovered from gut microbiota of a deep-sea isopod detected the *comEC* gene adjacent to genes
406 encoding enzymes producing dTMP from thymidine or dUMP, suggesting that the role of the ComEC
407 protein might be importing extracellular DNA as a nutrient source [83]. The presence of ComEC in
408 the equid *M. felis* isolate, and its absence in the Australian felid *M. felis* isolates suggests that during
409 evolutionary adaptation DNA translocation via a membrane channel might have become redundant
410 and perhaps was replaced with T4SS conjugation-mediated DNA transfer, as discussed above.
411
412 The Australian felid *M. felis* genomes exhibited a total absence of four coding sequences associated
413 with host cell interactions, those encoding a putative variable surface lipoprotein, a fibronectin type
414 III domain-containing protein, a FIVAR domain-containing protein, and a GA module-containing
415 protein, in contrast to the equid *M. felis* genome. The FIVAR (Found In Various Architectures Region)
416 and GA domains are essential constituents of extracellular matrix-binding proteins (Embp),
417 facilitating biofilm formation and attachment to fibronectin in some bacteria such as *Staphylococcus*
418 *epidermidis* and *S. aureus* [84, 85]. The presence of more than 30% amino acid sequence identity
419 between the FIVAR and GA domains in the equid *M. felis* reference isolate Myco-2 and these domains
420 in Embp of *S. aureus* raises the possibility that the equid *M. felis* isolate might have potential for
421 biofilm formation and/or fibronectin attachment. Conversely, the absence of these proteins, and of
422 the surface lipoprotein and fibronectin type III domain-containing protein, suggests that these
423 capacities may have become obsolete in the Australian felid *M. felis* isolates, possibly indicating
424 different modes of host cell interaction in the two hosts.
425

426 *Rapid evolution within two months in one host*

427 Two of the Australian *M. felis* isolates sequenced in this study, MF047 and MF219, were obtained
428 from separate bronchoalveolar lavages (BAL) of a feline patient with clinical signs consistent with
429 pneumonia. Isolate MF047 was obtained from a BAL collected during the initial consultation,
430 whereafter the cat was treated with 25 mg of doxycycline twice a day orally for approximately two
431 months. Because the response to treatment was unsatisfactory, a second BAL was collected and
432 isolate MF219 was obtained. While no genome reassortment was detected between the MF047 and
433 MF219 genomes, minor differences were observed in the counts of some genes encoding proteins
434 involved in RM systems, DNA transfer, and methylation (Figure 3). These variations may suggest rapid
435 host adaptation of *M. felis* after only two months. The growing utilisation of whole genome
436 sequencing in mycoplasma diagnostics is revealing links between genetic mutations and
437 antimicrobial resistance. Genomic comparisons of MF047 and MF219 revealed the presence of five
438 single nucleotide polymorphisms (SNPs) within the gene encoding the ribosomal RNA small subunit
439 methyltransferase A (RsmA/KsgA), along with one SNP in the gene encoding the 16S rRNA and
440 another SNP in the gene encoding the 23S rRNA. In other studies, mutations within the 16S RNA
441 methyltransferase family have been reported to confer aminoglycoside resistance in *M. bovis* and *M.*
442 *gallisepticum* [86, 87], while mutations in the 16S rRNA or 23S rRNA genes have been shown to be
443 associated with resistance to tetracyclines and macrolides, respectively, in *M. pneumoniae*, *M.*
444 *hominis*, *M. genitalium*, and *M. bovis* [86, 88-90]. MF047 and MF219 both contained five protein
445 coding regions with low confidence matches to tetracycline resistance genes, with the highest bit
446 score of the five associated with a predicted elongation factor G protein with amino acid sequence
447 identity/similarity to *tet*(T) from *Streptococcus pyogenes* of 25.8%/47.7%. It is important to
448 emphasise that the application of novel genome sequencing techniques for mycoplasma diagnostics,
449 particularly in the context of antimicrobial resistance, is still in its early stages, and the availability of
450 data for comprehensive analysis of all mycoplasma species (and particularly animal mycoplasmas) is
451 currently limited. Future research determining the phenotypic susceptibility of isolates MF047 and
452 MF219 to tetracyclines is necessary to determine whether there is a correlation between the
453 genomic variations observed and tetracycline resistance.

454

455 *Phage in mycoplasmas*

456 In other bacteria, phages have been demonstrated to provide critical virulence factors to their hosts,
457 and play a key role in the pathogenesis of diseases associated with a diverse range of bacteria (e.g.

458 cholera, diphtheria, botulism, and haemolytic-uremic syndrome). Approximately 20 mycoplasma
459 phages/prophages have been reported across diverse ruminant, avian, canine, pig, human and
460 rodent mycoplasmas [49, 91]. Of these, most are unclassified, only three have had their virion
461 structure visualised (P1, Hr1 and Br1), and only one has been classified by the International
462 Committee for the Taxonomy of Viruses (ICTV), the temperate *Mycoplasma pulmonis* phage P1 (12
463 kb genome), as the sole member of the floating genus *Delislevirus* [92]. The well-described temperate
464 MAV1 phage found in *Mycoplasma arthritidis* (16 kb genome) has not yet (at the time of publication)
465 been classified within the new system. In other mycoplasmas, phages have demonstrated dynamic
466 mobility within the bacterial chromosome, with loss of integrated prophages seen with clonal
467 passage (MFV1 in *Mycoplasma fermentans*), and variable expression of unique phage-encoded
468 membrane surface proteins (MFV1 and MAV1 in *M. fermentans* and *M. arthritidis*, respectively) [93],
469 and may have a role in bacterial adaptation and survival. In *M. felis*, vb MFe PM329 appears to also
470 be a temperate phage, capable of being virion-productive and potentially have a lysogenic phase. No
471 known selective advantage or disadvantage of the presence of this phage within *M. felis* was able to
472 be determined in this study because of the low number of *M. felis* genomes available. Detection of a
473 contig similar to that of vb MFe PM329 within *M. felis* 16-057065-4 (NCBI Biosample:
474 SAMN30346842) from Canada indicates that this phage is probably carried by *M. felis* strains globally,
475 although in that study it was not identified as a phage nor prophage. Future epidemiological studies
476 on isolates from healthy and clinical disease cases are indicated to investigate the potential role of
477 phage carriage in feline mycoplasma-associated disease.
478
479 Conclusions
480 This study has provided a comprehensive analysis of the first high-quality, complete genomes of *M.*
481 *felis* isolated from cats, and characterised a novel mycoplasma phage. It provides valuable insights
482 into the intricate genomic dynamics of *M. felis*, showcasing a complex interplay of gene gains and
483 losses, horizontal gene transfer events, and adaptations to new host environments. These findings
484 suggest divergent evolutionary paths in different hosts and offer compelling avenues for further
485 research into the evolutionary mechanisms and ecological factors that influence the genomic
486 diversity of mycoplasmas. Such studies include investigating the variation in metabolism, replication,
487 virulence, antimicrobial resistance and other functional properties in the strains with and without
488 different genes of interest highlighted here. Additionally, the significance of gene gains/losses, and
489 the presence of the newly identified phage, in relation to niche adaptation and host specificity

requires targeted studies focusing on collecting sufficient *M. felis* samples from clinical equine and feline cases.

Conflicts of interest:

The author(s) declare that there are no conflicts of interest.

Funding information:

This study was supported by the Feline Health Research Fund. P.K.V. was supported by a University of Melbourne, Melbourne Postdoctoral Fellowship.

Author contributions:

R.N.B. performed the bacterial isolation and initial classification; A.R.L., S.M.K. and P.K.V. prepared the DNA samples, performed the sequencing and analysis; P.K.V. isolated the phage for TEM; P.K.V., S.M.K., G.F.B. and A.R.L. conceptualised the study design; P.K.V., S.M.K. and A.R.L. acquired the funding; S.M.K., A.R.L. and P.K.V. drafted the manuscript, and R.N.B. and G.F.B. revised the manuscript.

Acknowledgments:

The authors would like to gratefully acknowledge the small animal medicine team of the U-Vet Werribee Animal Hospital at The University of Melbourne for providing the samples taken from clinical cases, and Eric Hanssen at the Ian Holmes Imaging Centre for the cryoTEM imaging of the vB Mfe PM329 phage.

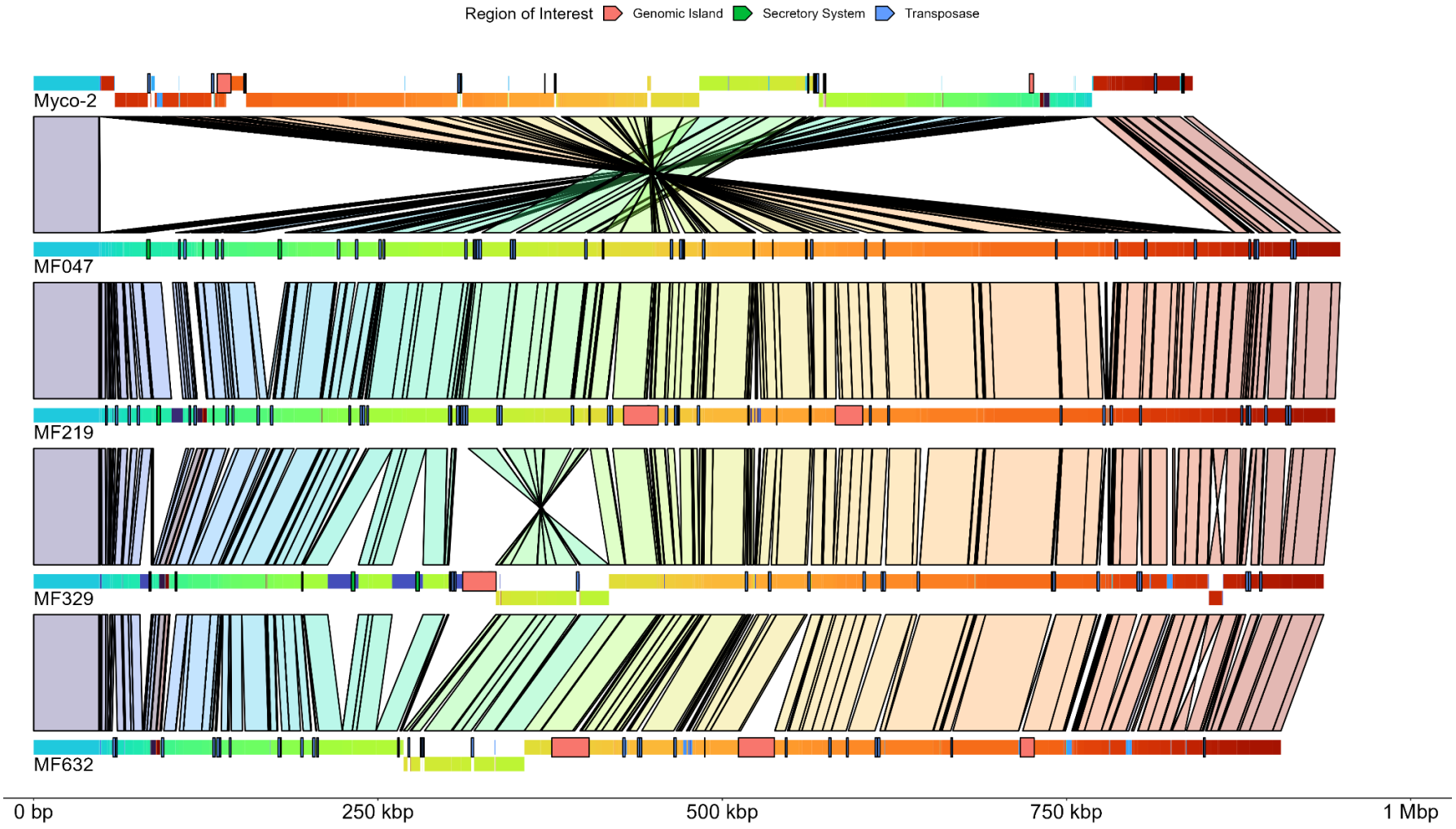
513 **Table 1.** *Mycoplasma felis* isolates sequenced in this study.

Isolate	Month/Year	Animal ID	Patient information	Clinical case	Isolation site	Other notes	GenBank #
047	3/2015	Cat A	Male, 5 y 9 m	Stertor, open-mouth breathing & dyspnoea; ocular & nasal discharge; anorexia; dehydration	Respiratory, BAL	Heavy growth from BAL, treated with doxycycline	CP114890
219	5/2015	Cat A	Male, 5 y 11 m	Repeat presentation of above	Respiratory, BAL	Follow-up isolation from Cat A	CP114889
329	5/2014	Cat B	Male, 8 y 11 m	Stertor & increased inspiratory effort; vomiting; chronic conjunctivitis; dehydration	Respiratory, BAL	Chronic bronchitis. PCR negative for herpesvirus, calicivirus and chlamydia	CP115656 OR753717 (phage)
632	12/2010	Cat C	Male, 4 y 6 m	Non-healing inguinal drain wound; adjacent fat pad swollen	Fat tissue	Isolated from non-healing surgical wound	CP114888

514 BAL, bronchioalveolar lavage; PCR, polymerase chain reaction

515

516



517

518 **Figure 1.** A progressive Mauve alignment of the *Mycoplasma felis* reference genome (Myco-2) and the genomes sequenced in this project. Locally co-linear blocks (LCBs)
519 are separated by colour within genomes, and are the same colour in each genome, with links to identify placement and inversion differences. Solid blocks overlaying the
520 genomes represent the presence of transposases, type IV secretion systems, and genomic islands, whilst solid blocks of blue that break up the colour gradient within the
521 genomes are representative of LCBs not present in linked genomes

522
523

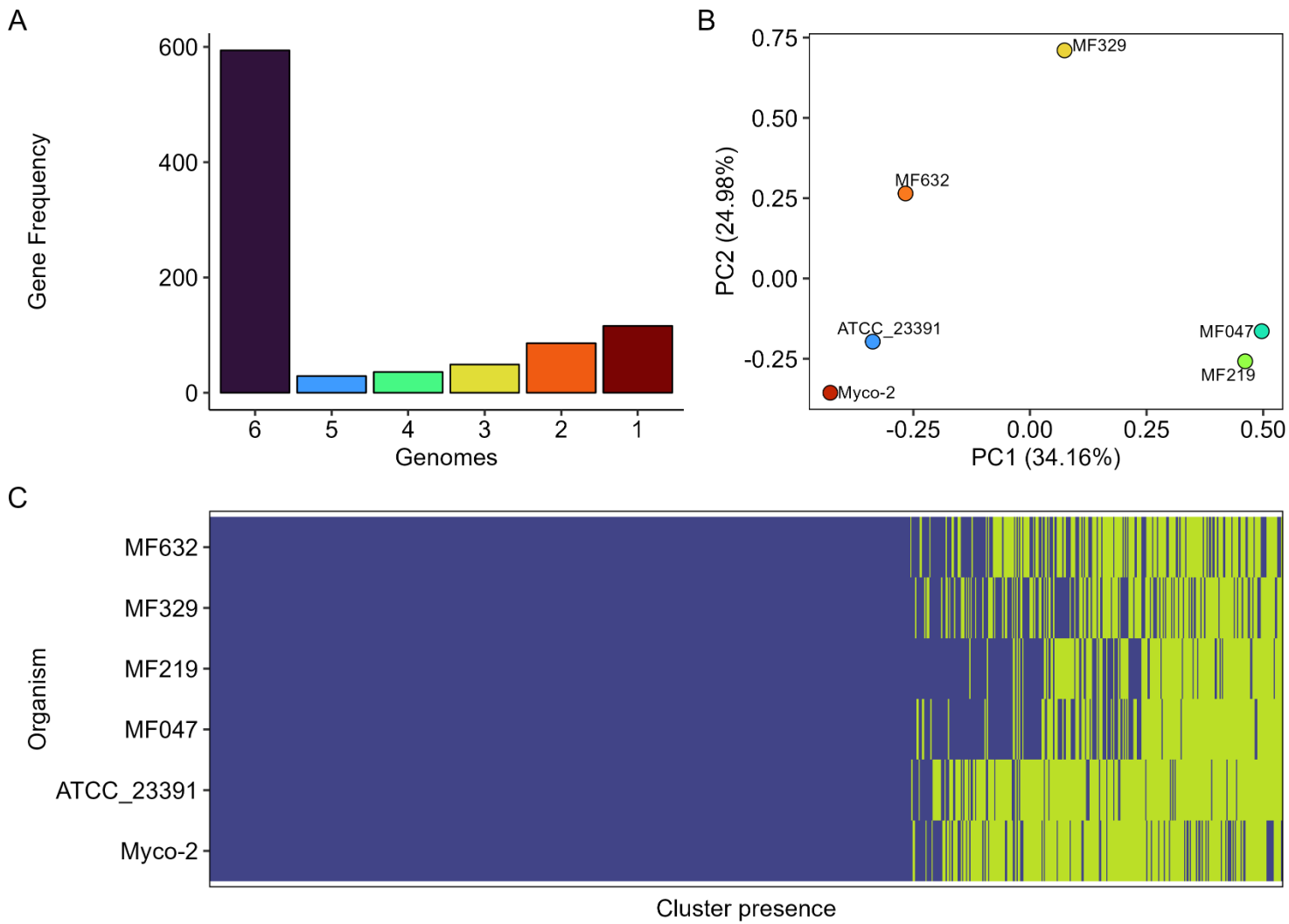


Figure 2. Analysis of the *Mycoplasma felis* core genome and gene presence/absence using Panaroo and plotted using Pagoo. A) the frequency of clustered genes within the six genomes used in the analysis; B) principal component analysis of gene presence/absence using Bray Curtis dissimilarity identifies clustering of the previously published genome sequences; C) clusters of genes present (blue) or absent (yellow) in the six genomes included in analysis.

528 **Table 2.** Coding sequences detected in the genomes of Australian *M. felis* isolates but absent in the equid *M. felis* reference strain (Myco2), and their possible species of
529 origin based on amino acid sequence identity.

Coding sequence	Possible species of origin	% Amino acid pairwise sequence identity	Predominant host
Type IV secretory system conjugative DNA transfer family protein	*M. cynos*	43.2	Dog
IS*256* family transposase	*M. penetrans*	50.9	Human
Type 2 glycerol-3-phosphate oxidase (GlpO)	*M. canis*	87.2	Dog
PTS fructose transporter subunit IIA	*M. flocculare*	52.6	Swine
DHHA1 domain-containing protein	*M. canis*	84.8	Dog
M13 family peptidase	*M. canis*	65.4	Dog
1-phosphofructokinase family	*M. canis*	52.6	Dog
Anti-termination protein NusB	*M. mustelae*	61.6	Mink
Hsp70 family protein	*M. flocculare*	85.4	Pig
C5 methylase (MAV1virus-like) protein	*M. columboralis*	70.2	Pigeon
16S rRNA methyltransferase (RsmG)	*M. mustelae*	63.3	Mink
HNH endonuclease	*M. columbina*	38.0	Pigeon
Eco47II family restriction endonuclease	*M. mycoides* subsp. *mycoides* SC	49.6	Cattle
Sau3AI family type II restriction endonuclease	*M. crocodyli*	52.7	Crocodile

530

531

532 **Table 3.** The location of predicted genomic islands in the genomes of *M. felis* isolates characterised in this study, as well as the reference genome Myco-2.

Strain	Genomic island	Start – End (relative to genome)	Length	Total number of genes (non-hypothetical)
Myco-2	1[b,d,e]	133250 - 143277	10027	11 (5)
	2[b,c,d,e]	723039 - 726091	3052	8 (6)
MF047	-	-	-	-
MF219	1[a,b,d,*]	428340 - 453515	25175	15 (7)
	2[a,b,d,e]	582071 - 602042	19971	11 (6)
MF329	1[b,c,#]	311409 - 335810	24401	19 (15)
MF632	1[a,b,d,*]	376254 - 403316	27062	19 (8)
	2[b,#]	511813 - 537933	26120	22 (15)
	3[a,b,c,d]	716455 - 726594	10139	8 (1)

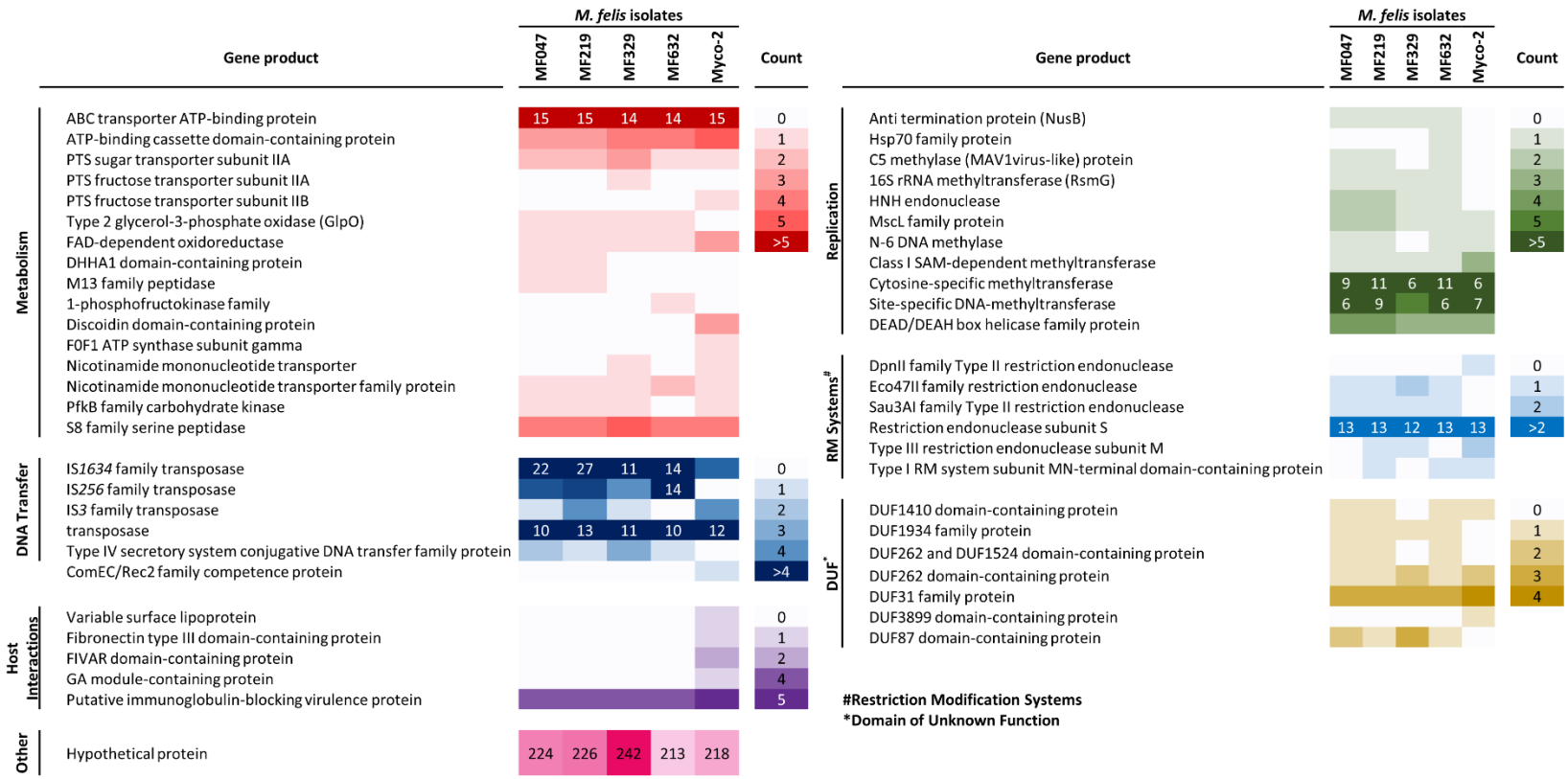
533 [a, b, c, d, e] – Predicted genomic island is present with some variation, such as the presence/absence of a transposase, in (a) Myco-2, (b) MF047, (c) MF219, (d) MF329, or (e)
534 MF632, but not predicted to be a genomic island in those cases.
535 [*, #] - Predicted genomic islands with structural similarity across different genomes
536 Genomic islands determined with the Island Finder 4 web tool. None were detected in the MF047 isolate.
537

538
539 **Table 4.** Quantity and range of amino acid sequence identities and similarities of putative virulence and antimicrobial resistance genes (AMR) based on significant BLASTP
540 hits (bit scores >50) compared to the virulence factor database and CARD database, respectively.

Genome	Putative virulence genes			Putative AMR genes		
	Total	Median identity (min – max)	Median similarity (min – max)	Total	Median identity (min – max)	Median similarity (min – max)
Myco-2	48	31.7 (24.4 - 90.1)	52.6 (44.4 - 95.9)	24	29.8 (22.5 - 40.1)	51.7 (39.1 - 69.6)
MF047	49	31.6 (24.6 - 90.1)	52.7 (41.8 - 95.9)	25	29.4 (22.6 - 40.3)	50.6 (39.2 - 69.6)
MF219	49	31.6 (24.6 - 90.1)	52.4 (41.8 - 95.9)	25	29.6 (22.6 - 40.3)	50.6 (39.2 - 69.6)
MF329	50	32 (24.4 - 90.1)	52.2 (41.5 - 95.9)	25	29.7 (22.5 - 40.2)	50.6 (39.1 - 69.6)
MF632	51	31.6 (24.4 - 90.1)	52.4 (41.5 - 95.9)	26	29.7 (22.4 - 40.2)	51 (38.9 - 69.6)

541
542

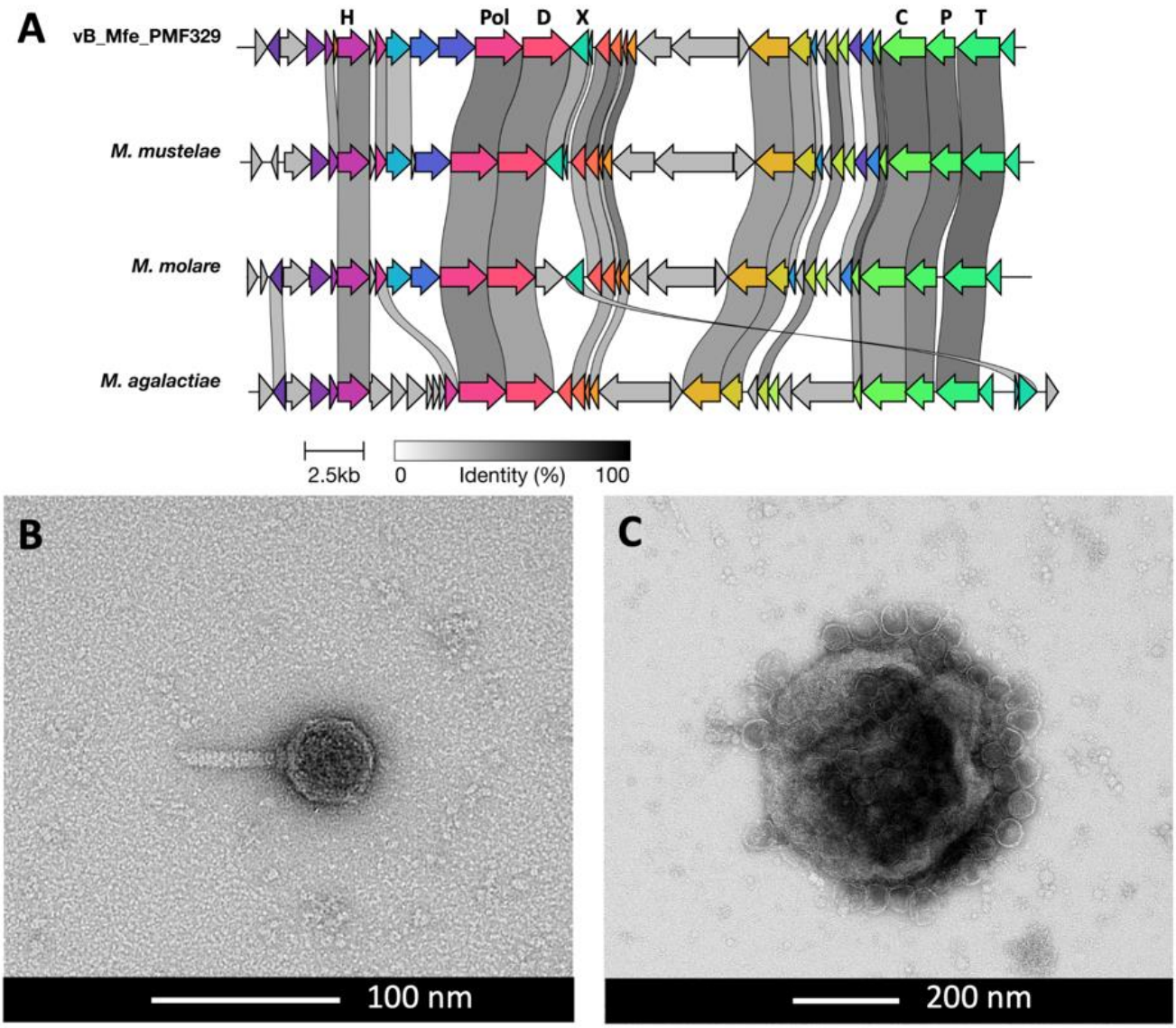
543



544

545 **Figure 3**. Presence and absence of genes within the genomes of Australian felid *M. felis* isolates compared to the genome of the Myco-2 equid *M. felis* reference strain,
546 clustered by putative protein functions. Only genes that differ between isolates or differ from the reference are listed. Counts indicate the number of gene copies detected
547 within each isolate.

548



549

550 **Figure 4.** A) Genome of the bacteriophage identified in *Mycoplasma felis* strain MF329 compared to MAgV1-like
551 prophage sequences found in *M. mustelae*, *M. molare* and *M. agalactiae*. Genome annotations depicting conserved
552 phage/prophage gene sequences are indicated in black. H = helicase; Pol = DNA polymerase; D = DNA primase; X = Xer
553 recombinase; C = prohead protein; P = portal protein; T = terminase; HP = hypothetical protein. B) and C) cryo-
554 transmission electron photomicrographs of bacteriophages detected in a cell lysate of *M. felis* isolate MF329, showing
555 that the phage structure consists of a contractile tail and polyhedral head.

556

557 ## References:

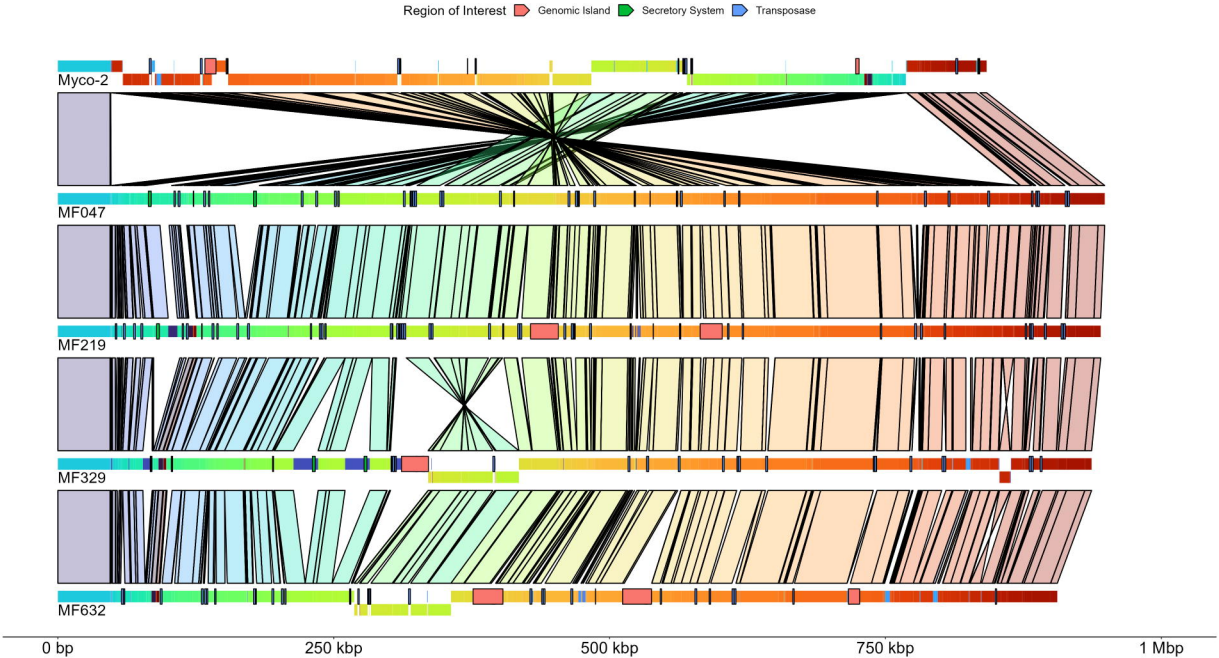
558 1. **Glass JI, Assad-Garcia N, Alperovich N, Yooseph S, Lewis MR et al.** Essential genes of a
559 minimal bacterium. *Proceedings of the National Academy of Sciences of the United States of America*
560 2006;103(2):425-430.
561 2. **Woese CR, Stackebrandt E, Ludwig W**. What are mycoplasmas: The relationship of tempo
562 and mode in bacterial evolution. *Journal of Molecular Evolution* 1985;21(4):305-316.
563 3. **Fadiel A, Eichenbaum KD, El Semary N, Epperson B**. Mycoplasma genomics: tailoring the
564 genome for minimal life requirements through reductive evolution. *Frontiers in Bioscience-Landmark*
565 2007;12(6):2020-2028.
566 4. **Rocha EPC, Blanchard A**. Genomic repeats, genome plasticity and the dynamics of
567 Mycoplasma evolution. *Nucleic Acids Research* 2002;30(9):2031-2042.
568 5. **Shil PK, Wawegama NK, Browning GF, Noormohammadi AH, Marenda MS**. Mycoplasmas.
569 In: J.F. Prescott, A.N. Rycroft, J.D. Boyce, J.I. MacInnes, Immerseel FV et al. (editors). *Pathogenesis of*
570 *Bacterial Infections in Animals* 2022. pp. 667-700.
571 6. **Krause DC, Leith DK, Wilson RM, Baseman JB**. Identification of *Mycoplasma pneumoniae*
572 proteins associated with hemadsorption and virulence. *Infect Immun* 1982;35(3):809-817.
573 7. **Delaney NF, Balenger S, Bonneaud C, Marx CJ, Hill GE et al.** Ultrafast evolution and loss of
574 CRISPRs following a host shift in a novel wildlife pathogen, *Mycoplasma gallisepticum*. *PLOS Genetics*
575 2012;8(2):e1002511.
576 8. **Momynaliev KT, Govorun VM**. Mechanisms of genetic instability in Mollicutes
577 (Mycoplasmas). *Russian Journal of Genetics* 2001;37(9):979-992.
578 9. **Citti C, Dordet-Frisoni E, Nouvel LX, Kuo CH, Baranowski E**. Horizontal gene transfers in
579 mycoplasmas (mollicutes). *Current Issues in Molecular Biology* 2018;29(1):3-22.
580 10. **Klose SM, Olaogun OM, Disint JF, Shil P, Gyuranecz M et al.** Genomic diversity of a globally
581 used, live attenuated mycoplasma vaccine. *Microbiology Spectrum* 2022;10(6):e02845-02822.
582 11. **Klose SM, Shil P, Underwood GJ, Morrow CJ, Marenda MS et al.** Obg plays a significant role
583 in temperature sensitivity of *Mycoplasma synoviae* live attenuated vaccine strain MS-H. *Veterinary*
584 *Microbiology* 2023:109818.
585 12. **Klose SM, Omotainse OS, Zare S, Vaz PK, Armat P et al.** Virulence factors of *Mycoplasma*
586 *synoviae*: Three genes influencing colonization, immunogenicity, and transmissibility. *Frontiers in*
587 *Microbiology* 2022;13:1042212.
588 13. **Klose SM, De Souza DP, Disint JF, Andrews DM, Underwood GJ et al.** Reversion of mutations
589 in a live mycoplasma vaccine alters its metabolism. *Vaccine* 2023;41(21):3358-3366.
590 14. **Noormohammadi A, Markham P, Duffy M, Whithear K, Browning G**. Multigene families
591 encoding the major hemagglutinins in phylogenetically distinct mycoplasmas. *Infection and Immunity*
592 1998;66(7):3470-3475.
593 15. **Sirand-Pugnet P, Lartigue C, Marenda M, Jacob D, Barré A et al.** Being pathogenic, plastic,
594 and sexual while living with a nearly minimal bacterial genome. *PLOS Genetics* 2007;3(5):e75.
595 16. **Randolph J, Moise N, Scarlett J, Shin S, Blue J et al.** Prevalence of mycoplasmal and
596 ureaplasmal recovery from tracheobronchial lavages and of mycoplasmal recovery from pharyngeal
597 swab specimens in cats with or without pulmonary disease. *American journal of veterinary research*
598 1993;54(6):897-900.
599 17. **Foster S, Barrs V, Martin P, Malik R**. Pneumonia associated with *Mycoplasma* spp in three
600 cats. *Australian veterinary journal* 1998;76(7):460-464.
601 18. **Sykes JE**. Feline hemotropic mycoplasmas. *Journal of Veterinary Emergency and Critical Care*
602 2010;20(1):62-69.

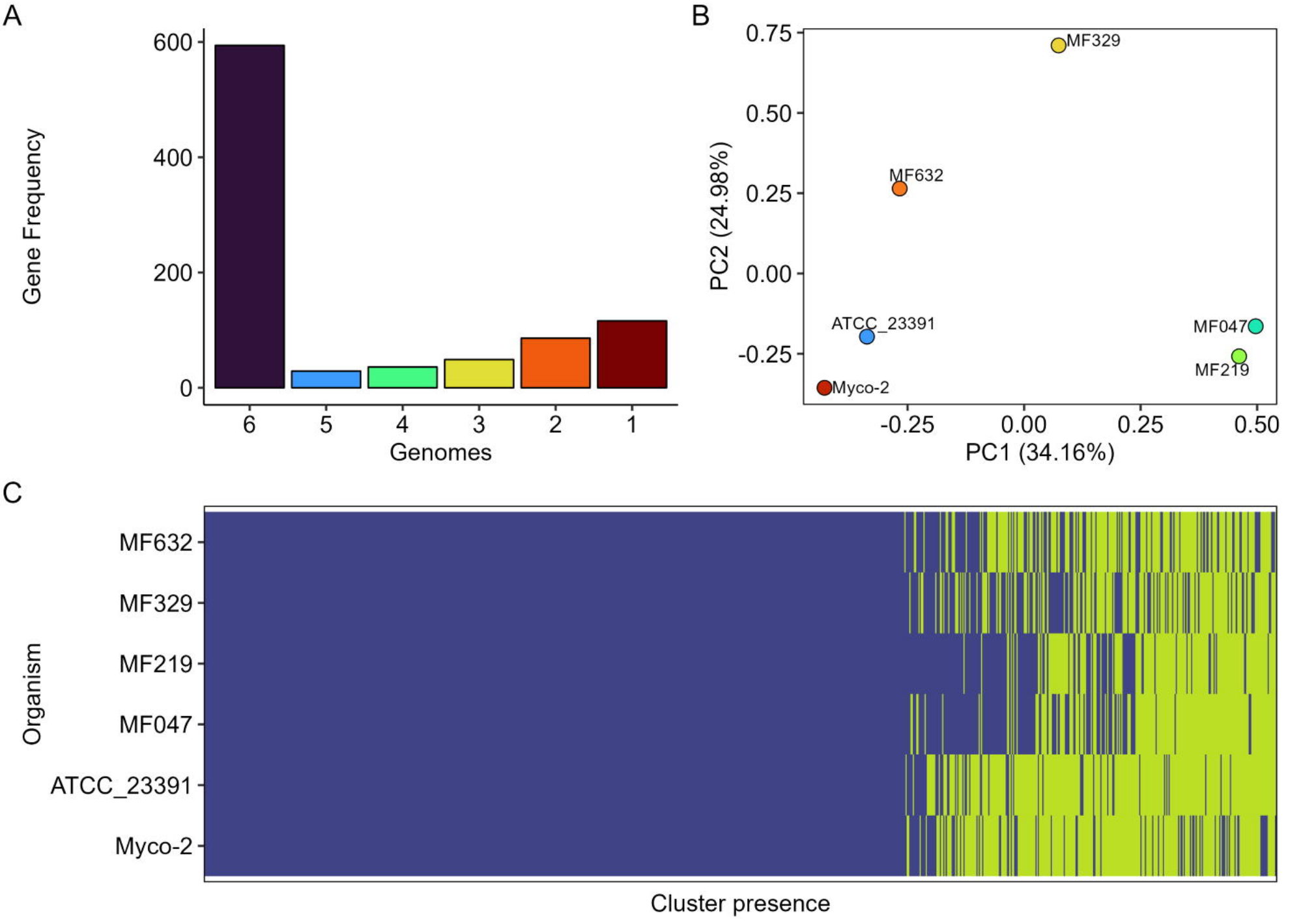
603 19. **Le Boedec K**. A systematic review and meta-analysis of the association between *Mycoplasma*
604 spp and upper and lower respiratory tract disease in cats. *Journal of the American Veterinary Medical*
605 *Association* 2017;250(4):397-407.
606 20. **Wood JLN, Chanter N, Newton JR, Burrell MH, Dugdale D et al.** An outbreak of respiratory
607 disease in horses associated with *Mycoplasma felis* infection. *Veterinary Record* 1997;140(15):388-
608 391.
609 21. **Cohn LA**. Feline respiratory disease complex. *Veterinary Clinics: Small Animal Practice*
610 2011;41(6):1273-1289.
611 22. **Nguyen D, Barrs VR, Kelman M, Ward MP**. Feline upper respiratory tract infection and
612 disease in Australia. *Journal of Feline Medicine and Surgery* 2019;21(10):973-978.
613 23. **Fernandez M, Manzanilla EG, Lloret A, León M, Thibault J-C**. Prevalence of feline herpesvirus-
614 1, feline calicivirus, *Chlamydophila felis* and *Mycoplasma felis* DNA and associated risk factors in cats
615 in Spain with upper respiratory tract disease, conjunctivitis and/or gingivostomatitis. *Journal of Feline*
616 *Medicine and Surgery* 2017;19(4):461-469.
617 24. **Lee-Fowler T**. Feline respiratory disease: What is the role of Mycoplasma species? *Journal of*
618 *Feline Medicine and Surgery* 2014;16(7):563-571.
619 25. **Tonozzi CC**. *Feline Respiratory Disease Complex*. MSD Veterinary Manual 2022.
620 26. **Kinoshita Y, Niwa H, Uchida-Fujii E, Nukada T**. Complete genome sequence of mycoplasma
621 felis Strain Myco-2, isolated from an equine tracheal wash sample in Japan. *Microbiology Resource*
622 *Announcements* 2020;9(9):10.1128.
623 27. **Framst I, D`Andrea C, Baquero M, Maboni G**. Development of a long-read next generation
624 sequencing workflow for improved characterization of fastidious respiratory mycoplasmas.
625 *Microbiology* 2022;168(11).
626 28. **Frey M**. A medium for the isolation of avian mycoplasmas. *American journal of veterinary*
627 *research* 1968;29:2163-2171.
628 29. **Chen S, Zhou Y, Chen Y, Gu J**. fastp: an ultra-fast all-in-one FASTQ preprocessor.
629 *Bioinformatics* 2018;34(17):i884-i890.
630 30. **Wick R**. *Filtlong v. 0.2. 0*. 2021.
631 31. **Wick RR, Judd LM, Gorrie CL, Holt KE**. Unicycler: Resolving bacterial genome assemblies from
632 short and long sequencing reads. *PLOS Computational Biology* 2017;13(6):e1005595.
633 32. **Wick RR, Holt KE**. Polypolish: Short-read polishing of long-read bacterial genome assemblies.
634 *PLOS Computational Biology* 2022;18(1):e1009802.
635 33. **Li H**. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*
636 *preprint arXiv:13033997* 2013.
637 34. **Darling AE, Mau B, Perna NT**. Progressive Mauve: Multiple alignment of genomes with gene
638 flux and rearrangement. *arXiv preprint arXiv:09105780* 2009.
639 35. **Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP et al.** NCBI prokaryotic
640 genome annotation pipeline. *Nucleic Acids Research* 2016;44(14):6614-6624.
641 36. **Buchfink B, Reuter K, Drost H-G**. Sensitive protein alignments at tree-of-life scale using
642 DIAMOND. *Nature Methods* 2021;18(4):366-368.
643 37. **Tonkin-Hill G, MacAlasdair N, Ruis C, Weimann A, Horesh G et al.** Producing polished
644 prokaryotic pangenomes with the Panaroo pipeline. *Genome Biology* 2020;21(1):180.
645 38. **Ferrés I, Iraola G**. An object-oriented framework for evolutionary pangenome analysis. *Cell*
646 *Reports Methods* 2021;1(5):100085.
647 39. **Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P et al.** DnaSP 6: DNA
648 Sequence Polymorphism Analysis of Large Data Sets. *Molecular Biology and Evolution*
649 2017;34(12):3299-3302.

650 40. **Dhillon BK, Chiu TA, Laird MR, Langille MG, Brinkman FS**. IslandViewer update: improved
651 genomic island discovery and visualization. *Nucleic Acids Research* 2013;41(W1):W129-W132.
652 41. **Liu M, Li X, Xie Y, Bi D, Sun J et al.** ICEberg 2.0: an updated database of bacterial integrative
653 and conjugative elements. *Nucleic Acids Research* 2018;47(D1):D660-D665.
654 42. **Alcock BP, Huynh W, Chalil R, Smith KW, Raphenya AR et al.** CARD 2023: expanded curation,
655 support for machine learning, and resistome prediction at the Comprehensive Antibiotic Resistance
656 Database. *Nucleic Acids Research* 2023;51(D1):D690-D699.
657 43. **Liu B, Zheng D, Zhou S, Chen L, Yang J**. VFDB 2022: a general classification scheme for
658 bacterial virulence factors. *Nucleic Acids Research* 2022;50(D1):D912-D917.
659 44. **Li H**. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*
660 2018;34(18):3094-3100.
661 45. **Garneau JR, Depardieu F, Fortier L-C, Bikard D, Monot M**. PhageTerm: a tool for fast and
662 accurate determination of phage termini and packaging mechanism using next-generation
663 sequencing data. *Scientific Reports* 2017;7(1):8292.
664 46. **Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW**. CheckM: assessing the
665 quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome*
666 *Research* 2015;25(7):1043-1055.
667 47. **Nishimura Y, Yoshida T, Kuronishi M, Uehara H, Ogata H et al.** ViPTree: the viral proteomic
668 tree server. *Bioinformatics* 2017;33(15):2379-2380.
669 48. **Bin Jang H, Bolduc B, Zablocki O, Kuhn JH, Roux S et al.** Taxonomic assignment of
670 uncultivated prokaryotic virus genomes is enabled by gene-sharing networks. *Nature Biotechnology*
671 2019;37(6):632-639.
672 49. **Citti C, Baranowski E, Dordet-Frisoni E, Faucher M, Nouvel L-X**. Genomic islands in
673 mycoplasmas. *Genes* 2020;11(8):836.
674 50. **Gourlay RN, Wyld SG, Garwes DJ**. Some properties of mycoplasma virus Br1. *Archives of*
675 *Virology* 1983;75(1):1-15.
676 51. **Gourlay RN, Wyld SG, Poulton ME**. Some characteristics of Mycoplasma virus Hr1, isolated
677 from and infectingMycoplasma hyorhinis. *Archives of Virology* 1983;77(1):81-85.
678 52. **Mahillon J, Chandler M**. Insertion sequences. *Microbiology and Molecular Biology Reviews*
679 1998;62(3):725-774.
680 53. **Frey J**. Insertion sequence analysis. In: Miles R, Nicholas R (editors). *Mycoplasma Protocols*.
681 Totowa, NJ: Humana Press; 1998. pp. 197-205.
682 54. **Lysnyansky I, Calcutt MJ, Ben-Barak I, Ron Y, Levisohn S et al.** Molecular characterization of
683 newly identified IS 3, IS 4 and IS 30 insertion sequence-like elements in *Mycoplasma bovis* and their
684 possible roles in genome plasticity. *FEMS Microbiology Letters* 2009;294(2):172-182.
685 55. **Thomas A, Linden A, Mainil J, Bischof DF, Frey J et al.** *Mycoplasma bovis* shares insertion
686 sequences with *Mycoplasma agalactiae* and *Mycoplasma mycoides* subsp. *mycoides* SC: evolutionary
687 and developmental aspects. *FEMS Microbiology Letters* 2005;245(2):249-255.
688 56. **Loreto ELS, Ortiz MF, Porto JIR**. Insertion sequences as variability generators in the
689 *Mycoplasma hyopneumoniae* and *M. synoviae* genomes. *Genetics and Molecular Biology*
690 2007;30:283-289.
691 57. **Vilei EM, Nicolet J, Frey J**. IS*1634*, a novel insertion element creating long, variable-length
692 direct repeats which is specific for *Mycoplasma mycoides* subsp. *mycoides* small-colony type. *Journal*
693 *of Bacteriology* 1999;181(4):1319-1323.
694 58. **Baranowski E, Dordet-Frisoni E, Sagné E, Hygonenq M-C, Pretre G et al.** The integrative
695 conjugative element (ICE) of *Mycoplasma agalactiae*: key elements involved in horizontal
696 dissemination and influence of coresident ices. *mBio* 2018;9(4):10.1128.

697 59. **Alvarez-Martinez CE, Christie PJ**. Biological diversity of prokaryotic type IV secretion systems.
698 *Microbiology and Molecular Biology Reviews* 2009;73(4):775-808.
699 60. **Dordet-Frisoni E, Sagné E, Baranowski E, Breton M, Nouvel LX et al.** Chromosomal transfers
700 in mycoplasmas: when minimal genomes go mobile. *MBio* 2014;5(6):10.1128/mbio. 01958-01914.
701 61. **Hartmann AD, Hawley J, Werckenthin C, Lappin MR, Hartmann K**. Detection of bacterial and
702 viral organisms from the conjunctiva of cats with conjunctivitis and upper respiratory tract disease.
703 *Journal of Feline Medicine & Surgery* 2010;12(10):775-782.
704 62. **Ahmer BMM, Tran M, Heffron F**. The virulence plasmid of *Salmonella typhimurium* is self-
705 transmissible. *Journal of Bacteriology* 1999;181(4):1364-1368.
706 63. **Szczepanek SM, Boccaccio M, Pflaum K, Liao X, Geary SJ**. Hydrogen peroxide production from
707 glycerol metabolism is dispensable for virulence of *Mycoplasma gallisepticum* in the tracheas of
708 chickens. *Infection and Immunity* 2014;82(12):4915-4920.
709 64. **Hames C, Halbedel S, Hoppert M, Frey J, Stülke J**. Glycerol metabolism is important for
710 cytotoxicity of *Mycoplasma pneumoniae*. *Journal of Bacteriology* 2009;191(3):747-753.
711 65. **Chang T-H, Lo W-S, Ku C, Chen L-L, Kuo C-H**. Molecular evolution of the substrate utilization
712 strategies and putative virulence factors in mosquito-associated Spiroplasma species. *Genome*
713 *Biology and Evolution* 2014;6(3):500-509.
714 66. **Pilo P, Vilei EM, Peterhans E, Bonvin-Klotz L, Stoffel MH et al.** A metabolic enzyme as a
715 primary virulence factor of *Mycoplasma mycoides* subsp. *mycoides* small colony. *Journal of*
716 *Bacteriology* 2005;187(19):6824-6831.
717 67. **Khan LA, Miles RJ, Nicholas RAJ**. Hydrogen peroxide production by *Mycoplasma bovis* and
718 *Mycoplasma agalactiae* and effect of *in vitro* passage on a *Mycoplasma bovis* strain producing high
719 levels of H_2O_2. *Veterinary Research Communications* 2005;29(3):181-188.
720 68. **Szczepanek S, Tulman E, Gorton T, Liao X, Lu Z et al.** Comparative genomic analyses of
721 attenuated strains of *Mycoplasma gallisepticum*. *Infection and immunity* 2010;78(4):1760-1771.
722 69. **Tseng C-W, Chiu C-J, Kanci A, Citti C, Rosengarten R et al.** The *oppD* gene and putative
723 peptidase genes may be required for virulence in *Mycoplasma gallisepticum*. *Infection and Immunity*
724 2017;85(6):10.1128.
725 70. **Masukagami Y, Nijagal B, Mahdizadeh S, Tseng C-W, Dayalan S et al.** A combined
726 metabolomic and bioinformatic approach to investigate the function of transport proteins of the
727 important pathogen *Mycoplasma bovis*. *Veterinary Microbiology* 2019;234:8-16.
728 71. **Masukagami Y, De Souza D, Dayalan S, Bowen C, O'callaghan S et al.** Comparative
729 metabolomics of *Mycoplasma bovis* and *Mycoplasma gallisepticum* reveals fundamental differences
730 in active metabolic pathways and suggests novel gene annotations. *MSystems*
731 2017;2(5):10.1128/msystems. 00055-00017.
732 72. **Mahdizadeh S, Masukagami Y, Tseng C-W, Markham PF, De Souza DP et al.** A *Mycoplasma*
733 *gallisepticum* glycerol ABC transporter involved in pathogenicity. *Applied and Environmental*
734 *Microbiology* 2021;87(11):e03112-03120.
735 73. **Chen H, Yu S, Shen X, Chen D, Qiu X et al.** The *Mycoplasma gallisepticum* α-enolase is cell
736 surface-exposed and mediates adherence by binding to chicken plasminogen. *Microbial*
737 *Pathogenesis* 2011;51(4):285-290.
738 74. **Vilei EM, Frey J**. Genetic and biochemical characterization of glycerol uptake in *Mycoplasma*
739 *mycoides* subsp. *mycoides* SC: Its impact on H_2O_2 production and virulence. *Clinical Diagnostic*
740 *Laboratory Immunology* 2001;8(1):85-92.
741 75. **Shterzer N, Mizrahi I**. The animal gut as a melting pot for horizontal gene transfer. *Canadian*
742 *Journal of Microbiology* 2015;61(9):603-605.

743 76. **Norman A, Hansen LH, Sørensen SJ**. Conjugative plasmids: vessels of the communal gene
744 pool. *Philosophical Transactions of the Royal Society B: Biological Sciences* 2009;364(1527):2275-
745 2289.
746 77. **Oliveira PH, Touchon M, Rocha EP**. The interplay of restriction-modification systems with
747 mobile genetic elements and their prokaryotic hosts. *Nucleic Acids Research* 2014;42(16):10618-
748 10631.
749 78. **Dordet-Frisoni E, Vandecasteele C, Contarin R, Sagné E, Baranowski E et al.** Impacts of
750 *Mycoplasma agalactiae* restriction-modification systems on pan-epigenome dynamics and genome
751 plasticity. *Microbial Genomics* 2022;8(5).
752 79. **Oliveira PH, Touchon M, Rocha EPC**. Regulation of genetic flux between bacteria by
753 restriction–modification systems. *Proceedings of the National Academy of Sciences*
754 2016;113(20):5658-5663.
755 80. **Chen I, Dubnau D**. DNA uptake during bacterial transformation. *Nature Reviews Microbiology*
756 2004;2(3):241-249.
757 81. **Kapteijn R, Shitut S, Aschmann D, Zhang L, de Beer M et al.** Endocytosis-like DNA uptake by
758 cell wall-deficient bacteria. *Nature Communications* 2022;13(1):5524.
759 82. **Pimentel ZT, Zhang Y**. Evolution of the natural transformation protein, ComEC, in bacteria.
760 *Frontiers in Microbiology* 2018:2980.
761 83. **Wang Y, Huang JM, Wang SL, Gao ZM, Zhang AQ et al.** Genomic characterization of symbiotic
762 mycoplasmas from the stomach of deep-sea isopod *Bathynomus* sp. *Environmental Microbiology*
763 2016;18(8):2646-2659.
764 84. **Christner M, Franke GC, Schommer NN, Wendt U, Wegert K et al.** The giant extracellular
765 matrix-binding protein of *Staphylococcus epidermidis* mediates biofilm accumulation and attachment
766 to fibronectin. *Molecular Microbiology* 2010;75(1):187-207.
767 85. **Speziale P, Pietrocola G, Foster TJ, Geoghegan JA**. Protein-based biofilm matrices in
768 staphylococci. *Frontiers in cellular and infection microbiology* 2014;4:171.
769 86. **Ledger L, Eidt J, Cai HY**. Identification of antimicrobial resistance-associated genes through
770 whole genome sequencing of *Mycoplasma bovis* isolates with different antimicrobial resistances.
771 *Pathogens* 2020;9(7):588.
772 87. **Mahdizadeh S, Sansom FM, Lee S-W, Browning GF, Marenda MS**. Targeted mutagenesis of
773 *Mycoplasma gallisepticum* using its endogenous CRISPR/Cas system. *Veterinary Microbiology*
774 2020;250:108868.
775 88. **Pereyre S, Goret J, Bébéar C**. *Mycoplasma pneumoniae*: current knowledge on macrolide
776 resistance and treatment. *Frontiers in Microbiology* 2016;7:974.
777 89. **Dégrange S, Renaudin H, Charron A, Pereyre S, Bébéar C et al.** Reduced susceptibility to
778 tetracyclines is associated *in vitro* with the presence of 16S rRNA mutations in *Mycoplasma hominis*
779 and *Mycoplasma pneumoniae*. *Journal of Antimicrobial Chemotherapy* 2008;61(6):1390-1392.
780 90. **Shimada Y, Deguchi T, Nakane K, Yasuda M, Yokoi S et al.** Macrolide resistance–associated
781 23S rRNA mutation in *Mycoplasma genitalium*, Japan. *Emerging Infectious Diseases* 2011;17(6):1148.
782 91. **Kovács ÁB, Wehmann E, Sváb D, Bekő K, Grózner D et al.** Novel prophage-like sequences in
783 *Mycoplasma anserisalpingitidis*. *Infection, Genetics and Evolution* 2021;92:104886.
784 92. **Turner D, Shkoporov AN, Lood C, Millard AD, Dutilh BE et al.** Abolishment of morphology-
785 based taxa and change to binomial species names: 2022 taxonomy update of the ICTV bacterial
786 viruses subcommittee. *Archives of Virology* 2023;168(2):74.
787 93. **Röske K, Calcutt MJ, Wise KS**. The *Mycoplasma fermentan*s prophage ϕMFV1: genome
788 organization, mobility and variable expression of an encoded surface protein. *Molecular*
789 *Microbiology* 2004;52(6):1703-1720.

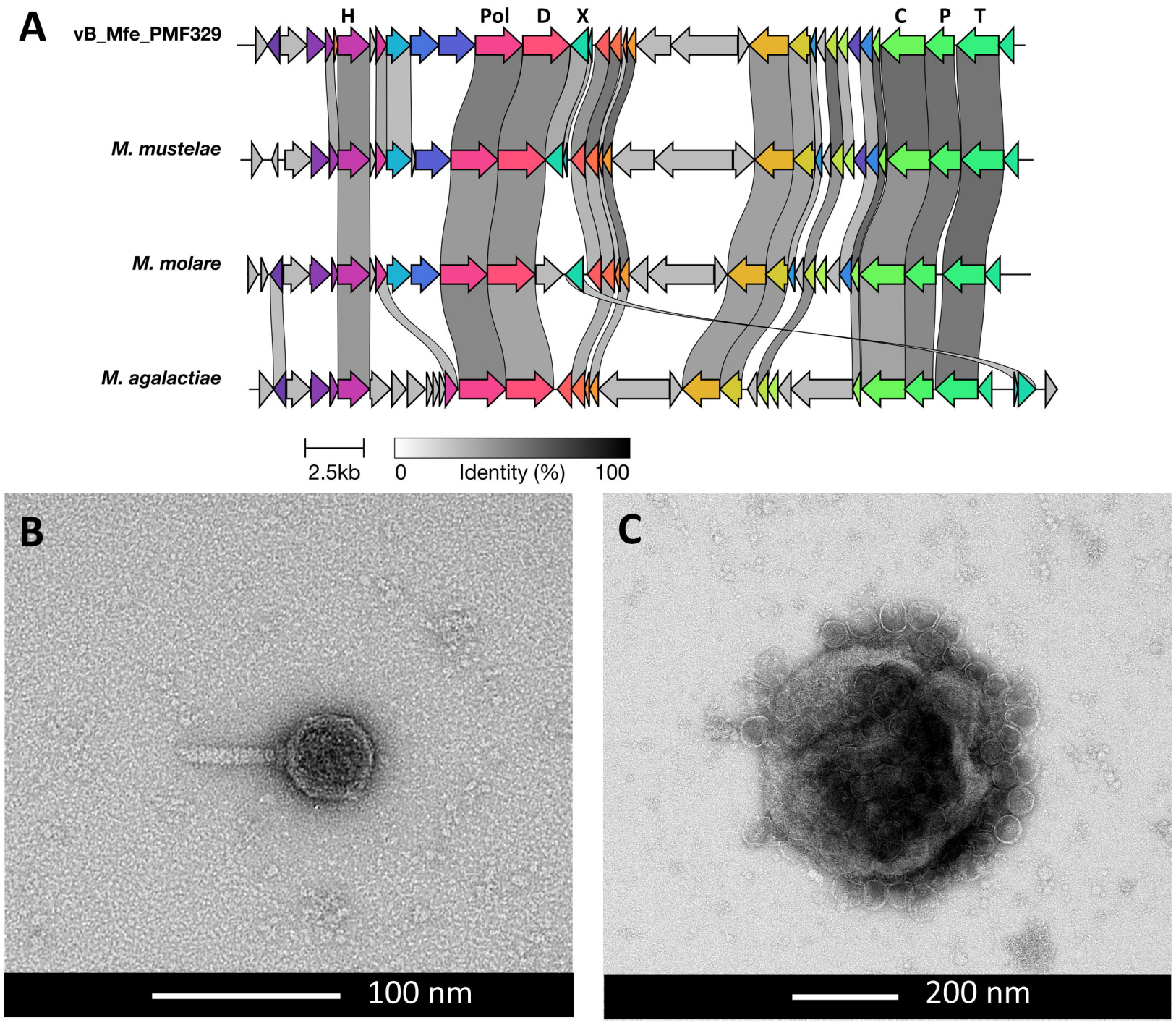

Region of Interest
Genomic Island
Secretory System
Transposase
Myco-2
MF047
MF219
MF329
MF632
0 bp
250 kbp
500 kbp
750 kbp
1 Mbp


A
Gene Frequency
600
400
200
0
6
5
4
3
2
1
Genomes
B
PC2 (24.98%)
0.75
0.50
0.25
0.00
-0.25
MF329
MF632
ATCC_23391
MF047
MF219
Myco-2
-0.25
0.00
0.25
0.50
PC1 (34.16%)
C
Organism
MF632
MF329
MF219
MF047
ATCC_23391
Myco-2
Cluster presence

	Gene product	*M. felis* isolates					Count
		MF047	MF219	MF329	MF632	Myco-2	
Metabolism	ABC transporter ATP-binding protein	15	15	14	14	15	0
	ATP-binding cassette domain-containing protein						1
	PTS sugar transporter subunit IIA						2
	PTS fructose transporter subunit IIA						3
	PTS fructose transporter subunit IIB						4
	Type 2 glycerol-3-phosphate oxidase (GlpO)						5
	FAD-dependent oxidoreductase						>5
	DHHA1 domain-containing protein						
	M13 family peptidase						
	1-phosphofructokinase family						
	Discoidin domain-containing protein						
	F0F1 ATP synthase subunit gamma						
	Nicotinamide mononucleotide transporter						
	Nicotinamide mononucleotide transporter family protein						
	PfkB family carbohydrate kinase						
	S8 family serine peptidase						
DNA Transfer	IS*1634* family transposase	22	27	11	14		0
	IS*256* family transposase				14		1
	IS*3* family transposase						2
	transposase	10	13	11	10	12	3
	Type IV secretory system conjugative DNA transfer family protein						4
	ComEC/Rec2 family competence protein						>4
Host Interactions	Variable surface lipoprotein						0
	Fibronectin type III domain-containing protein						1
	FIVAR domain-containing protein						2
	GA module-containing protein						4
	Putative immunoglobulin-blocking virulence protein						5
Other	Hypothetical protein	224	226	242	213	218	

	Gene product	*M. felis* isolates					Count
		MF047	MF219	MF329	MF632	Myco-2	
Replication	Anti termination protein (NusB)						0
	Hsp70 family protein						1
	C5 methylase (MAV1virus-like) protein						2
	16S rRNA methyltransferase (RsmG)						3
	HNH endonuclease						4
	MscL family protein						5
	N-6 DNA methylase						>5
	Class I SAM-dependent methyltransferase						
	Cytosine-specific methyltransferase	9	11	6	11	6	
	Site-specific DNA-methyltransferase	6	9		6	7	
	DEAD/DEAH box helicase family protein						
RM Systems#	DpnII family Type II restriction endonuclease						0
	Eco47II family restriction endonuclease						1
	Sau3AI family Type II restriction endonuclease						2
	Restriction endonuclease subunit S	13	13	12	13	13	>2
	Type III restriction endonuclease subunit M						
	Type I RM system subunit MN-terminal domain-containing protein						
DUF*	DUF1410 domain-containing protein						0
	DUF1934 family protein						1
	DUF262 and DUF1524 domain-containing protein						2
	DUF262 domain-containing protein						3
	DUF31 family protein						4
	DUF3899 domain-containing protein						
	DUF87 domain-containing protein						

#Restriction Modification Systems
***Domain of Unknown Function**


A
vB_Mfe_PMF329
M. mustelae
M. molare
M. agalactiae
H
Pol
D
X
C
P
T
2.5kb
0
Identity (%)
100
B
100 nm
C
200 nm