

1 **Minimalistic mycoplasmas harbor different functional toxin-antitoxin systems**

2 Virginia Hill^{1,2}, Hatice Akarsu¹, Rubén Sánchez Barbarroja¹, Valentina Cippà¹, Martin Heller³,
3 Laurent Falquet⁴, Manfred Heller⁵, Michael H. Stoffel⁶, Fabien Labroussaa^{1#}, Joerg Jores^{1#*}

4

5 ¹ Institute of Veterinary Bacteriology, University of Bern, Bern, Switzerland

6 ² Graduate School for Biomedical Science, University of Bern, Bern, Switzerland

7 ³ Friedrich-Loeffler-Institute - Federal Research Institute for Animal Health, Jena, Germany

8 ⁴ Biochemistry Unit, University of Fribourg and Swiss Institute of Bioinformatics, Fribourg,
9 Switzerland

10 ⁵ Proteomics and Mass Spectrometry Core Facility, Department for BioMedical Research
11 (DBMR), University of Bern, Bern, Switzerland.

12 ⁶ Division of Veterinary Anatomy, University of Bern, Department of Clinical Research and
13 Veterinary Public Health, Switzerland

14

15 # equal contribution

16 *Corresponding author: Email: joerg.jores@vetsuisse.unibe.ch, Tel.: +41 31 631 2414

17

18 **Short title: Toxin-antitoxin systems in mycoplasmas**

19 **Keywords:** *Mycoplasma mycoides* subsp. *capri*; toxin-antitoxin system, TA-system, genome,
20 GM12, cloning

21 **Abstract**

22 Mycoplasmas are minute bacteria controlled by very small genomes ranging from 0.6 to 1.4
23 Mbp. They lack a cell wall and have been suggested to have progressed through reductive
24 evolution from phylogenetically closely related Clostridia. They are known to colonize the
25 respiratory tract or the urogenital tract among other organs and can cause chronic and
26 subclinical diseases associated with long persistence of the causative agent. Toxin-antitoxin
27 systems (TAS) are genetic elements that have been described for several respiratory and
28 urogenital pathogens as well as for Clostridia, but never for pathogenic mycoplasmas. Here we
29 describe for the first-time different types of TAS in a *Mycoplasma* pathogen, namely *M.*
30 *mycoides* subsp. *capri*. We identified candidate TAS *in silico* via TASmania database. Two
31 candidate TAS identified *in silico* and another candidate TAS suggested in a minimal cell based
32 on transposon mutagenesis were systematically tested for their functionality in hosts with
33 different phylogenetic distance using heterologous expression. Phylogenetic distance of the
34 host used for heterologous expression influenced the outcome of the functional testing. We
35 corroborated functionality of the three candidate TAS in *Mycoplasma capricolum* subsp.
36 *capricolum*. Moreover, we confirmed transcription and translation of molecules of the TAS
37 investigated during *in vitro* growth. We sequence analyzed 15 genomes of *M. mycoides* subsp.
38 *capri* and revealed an unequal distribution of the TAS studied pointing towards dynamic gain
39 and loss of TAS within the species.

40 **Author summary**

41 Mycoplasmas have a minimal genome and have never been shown to possess TAS. In this
42 work we showed the presence of different functional TAS systems in *Mycoplasma mycoides*
43 subsp. *capri*, a caprine pathogen for the first time. Sequence analysis of a number of
44 *Mycoplasma mycoides* subsp. *capri* strains revealed a plasticity of the genome with respect to
45 TAS carriage. This work paves the way to investigate the biological role of TAS (e.g.

46 persistence, stress tolerance) during infection using mycoplasmas as a simple model organism.
47 Since most mycoplasmas lack classical virulence factors such as exotoxins and go into a kind
48 of stealth mode to evade the immune system, TAS are likely to contribute to the parasitic
49 lifestyle of mycoplasmas and should be investigated in that respect. The availability of
50 synthetic genomics tools to modify a range of *Mycoplasma* pathogens and well-established
51 challenge models for the latter mycoplasmas will foster future research on TAS in
52 mycoplasmas.

53 **Introduction**

54 Mycoplasmas are the smallest bacteria that can replicate in axenic media reported so far. Their
55 cell size and genome size are minute compared to most other bacteria. The genomes range from
56 0.6 to 1.4 Mbp, a feature that attracted researchers to use mycoplasmas as model organisms for
57 synthetic cells [1] and minimal cells [2]. The small genome size has been attributed to a
58 reductive evolution of mycoplasmas, which are lacking a cell wall and are phylogenetically
59 closely related to spore-forming GRAM-positive bacteria, the Clostridia [3]. A characteristic
60 feature of the mycoplasmas is their parasitic lifestyle, which is reflected by their inability to
61 synthesize essential building blocks of life such as amino acids and hitherto their dependence
62 on a host to provide these building blocks to fuel the anabolic pathways.

63 Mycoplasmas colonize different niches of their hosts such as mucous membranes of the
64 respiratory tract, the ear canal or the urogenital system. The genus *Mycoplasma* encompasses
65 important human pathogens such as *M. genitalium* [4] and *M. pneumoniae* [5] as well as
66 pathogens of utmost veterinary importance such as *M. mycoides* subsp. *mycoides* [6], *M.*
67 *capricolum* subsp. *capripneumoniae* [7] and *M. hyopneumoniae* [8]. Although a number of
68 mycoplasmas cause diseases with a relatively short incubation time and high lethality as known
69 for contagious caprine pleuropneumonia [9], many *Mycoplasma* infections are chronic and
70 associated with a long persistence of the causative agent [10, 11].

71 Toxin-antitoxin systems (TAS) are chromosomal or extrachromosomal genetic elements,
72 which can be grouped into six different types according to their mode of action [12]. All TAS
73 encode a toxin (T), which is able to interfere with vital processes of the bacterial cell related to
74 transcription, translation, replication and membrane integrity, as well as an antitoxin (A) that
75 inhibits the activity of the toxin, so that it cannot interfere with the vital cellular processes. TAS
76 have been described for a wide range of GRAM-positive and GRAM-negative bacteria
77 including pathogens of the gastrointestinal and respiratory tract such as Clostridia and

78 Mycobacteria, respectively. The role of TAS has been attributed to reduce metabolism during
79 stress, hinder bacteriophages, stabilize genetic elements, and modulate the build-up of biofilms
80 [13].

81 Pathogenic mycoplasmas are not known for classical exotoxins, the only exception is an ADP-
82 ribosyltransferase (ART) activity-conferring toxin identified in *M. pneumoniae* [14]. So far,
83 TAS have not been documented in mycoplasmas. The presence of a TAS in an engineered
84 strain of *Mycoplasma mycoides* subsp. *capri* (*Mmc*) with a heavily reduced genome has been
85 suggested based on polar effects of transposon insertions [15], but was never confirmed using
86 heterologous expression or shown in its parental wild type strain GM12.

87 The goal of this study was to shed light on the presence of TAS in mycoplasmas, which are
88 known for their minute genome size. Therefore, we chose *Mmc* GM12, a caprine pathogen [16]
89 used to study virulence traits [17-19] and used as model organism for synthetic genomics. We
90 performed an *in silico* analysis to identify candidate TAS, which were subsequently
91 systematically tested for their functionality in hosts with different phylogenetic distance using
92 heterologous expression. Thereafter, we examined the presence of the candidate TAS at the
93 species and genus level and noticed an unequal distribution pointing towards horizontal gene
94 transfer to be involved in dissemination of the TAS over different mycoplasmas.

95 **Results**

96 **Candidate TAS identified *in silico* by TASmania database search**

97 The database TASmania [20], which is a discovery-oriented database, was used here to identify
98 candidate TAS in *Mmc* GM12. Our *in silico* search revealed 38 genes encoding candidate
99 toxins or antitoxins (**Table S1**). With respect to our cut-off settings, the following five
100 genes/gene pairs qualified as candidate TAS: MMCAP1_0160/0161, MMCAP1_0890,
101 MMCAP1_0731, MMCAP1_0525, MMCAP1_0753 (**Table 1**). Four of the latter TAS
102 candidates were without a proposed counterpart, the corresponding toxins or antitoxins were
103 manually chosen using ‘guilt by association’, considering the cognate loci of the respective
104 gene as the partner of a two gene operon. A transposon library of JCVI-syn2.0, a synthetic
105 mycoplasma cell, controlled by a reduced genome of *Mmc* strain GM12, suggested
106 JCVISYN2_0132 (referred here as A₁₃₂) and JCVISYN2_0133 (referred to as T₁₃₃) to act as
107 putative TAS [15]. Therefore, we included this candidate TAS into our study.

108 The six candidate TAS were distributed throughout the *Mmc* GM12 genome (**Figure 1**).
109 Inspecting the gene loci, we observed that TAS_{133/2} and TAS_{731/0} were arranged in a classical
110 TAS operon structure, with the antitoxin upstream of the toxin and one promotor. TAS_{525/4},
111 TAS_{160/1} and TAS_{889/0} were arranged in a classical operon structure but in antisense orientation.
112 TAS_{752/3} had a non-classical TAS operon structure with the toxin upstream of the antitoxin and
113 an intergenic region of 207 bp. Analysis of this intergenic region using BPROM identified one
114 possible promotor with the putative -10 and -35 boxes, indicating that the antitoxin has its own
115 promotor (**Figure 1**).

116 Gene description on TASmania database denominated TAS_{160/1} as abortive infection proteins,
117 AbiGII (T₁₆₀) and AbiGI (A₁₆₁), which belong to a well-described protein family (Abi),
118 described to act as toxin-antitoxin systems in e.g. *Streptococcus agalactiae* [21]. T₇₅₂ was
119 denominated as cell filamentation protein (fic) and A₇₅₃ as DNA-damage-inducible protein J.

120 A₇₅₃ encoded a protein which shares identity with RelB family of TASs, described, amongst
121 others, in *E. coli* [22]. The genes T₁₃₃ and A₁₃₂ were reported to encode an AAA-ATPase and
122 a subtilisin-like serine protease, respectively. A functional TAS (ietAS) with compelling
123 protein similarity exists in *Agrobacterium tumefaciens* [23]. A₈₉₀ and T₇₃₁ were also
124 denominated as belonging to the Abi protein family, according to TASmania and A₅₂₅ was
125 assigned to belong to MraZ protein family, which is similar to the TAS MazEF of *E. coli* [24,
126 25]. Different types of TAS have been described based on their mode of action and the six TAS
127 identified in our study represent type II and type IV TAS (**Table 1**).

128

129 **Detection of transcripts and proteins encoded by candidate TAS**

130 We tested the presence of transcripts covering the genes encoding the six candidate TAS.
131 Therefore, reverse transcription PCR was performed on *Mmc* GM12 RNA with primers
132 targeting the genes encoding candidate toxins, antitoxins and the whole putative two-gene
133 operons. We isolated 24 - 35 µg total RNA from 20 mL cultures. A PCR with primers targeting
134 16S rRNA did not reveal any amplicons confirming absence of DNA in the RNA preparation.
135 Transcripts of all 6 TAS investigated except for A₈₉₀ were detected via reverse transcription
136 PCR (**Figure S1**).

137 Afterwards we investigated the presence of candidate toxins and antitoxins using a shot gun
138 proteomics approach, in which we harvested GM12 at one time point, specifically during the
139 log growth phase. Mass spectrometry showed that in *Mmc* GM12 protein abundance of A₅₂₄ is
140 medium, whereas A₁₃₂, A₇₅₃ and A₁₆₁ as well as T₁₃₃ and T₇₃₁ are available in low abundance.
141 We arbitrarily assigned relative abundance thresholds based on the quartile of the log₁₀ (mean
142 dNSAF) values distribution: value -3.8 of the lower quartile (Q1) defined as low/medium
143 boundary and value -2.8 of the upper quartile (Q3) defined as medium/high boundary. For
144 reference, house-keeping enzymes dihydrolipoamide S-acetyltransferase, adenylate kinase,

145 glucose-6-phosphate isomerase, DNA-directed RNA polymerase beta chain, guanylate kinase,
146 DNA gyrase subunit B and recombination protein present \log_{10} (mean dNSAF) values of -2.0,
147 -2.5, -2.5, -2.6, -2.8, -3.0 and -4.4, respectively. The other candidate toxins and antitoxins were
148 below the detection limit of the mass spectrometry analysis (**Figure 2, Table S2**). In summary
149 only the candidate toxin and antitoxin of TAS_{133/2} have been detected using the shot gun
150 approach, for the TAS.

151

152 **High conservation in the organization of the *Mmc* genomes**

153 Each of the 15 genomes was assembled into one circularized chromosome and the
154 chromosomes were rotated based on the gene *dnaA*. Key features of each genome are displayed
155 in **Table S3**. The size of the 15 genomes ranged from 1,019,884 to 1,172,410 bp and a G+C
156 content of 23.7 %. All genomes contained 908 (\pm 93) CDS, 30 tRNAs, two rRNA operons, and
157 1 tmRNA each. The functional annotation, based on BLASTP of the CDS against a UniProtKB
158 snapshot, revealed an average of 392 hypothetical proteins (~45% of the total annotated CDS)
159 for each genome. In terms of genome organization, we observed a high level of synteny
160 (**Figure S2**). Only strains 152/93 and C260/4 had a large inversion of more than 100.000 bp,
161 which was confirmed by PCRs. Strain 152/93 contained a plasmid with a size of 1.875 bp,
162 which was highly similar to pKMK1 [26] and p*Mmc*-95010 [27] detected in strains GM12 and
163 95010, respectively.

164

165 **Distribution of the six candidate TAS in other mycoplasmas**

166 Next, we were interested to investigate the distribution of the six candidate TAS in different
167 strains of *Mmc* and to extend this analysis to members of the so called '*M. mycoides* cluster',
168 which comprises four additional closely-related ruminant pathogens besides *Mmc*. The
169 presence of genes was investigated by TBLASTN employing complete genome sequences.

170 Different *Mmc* strains showed that the presence of candidate TAS was not clonal, meaning
171 phylogenetically closely related strains had not necessarily the same set of candidate TAS.
172 Specifically, TAS_{752/3} was present in most strains, except C260/4 and Wi8079, which lacked
173 A₇₅₃. TAS_{133/2} and TAS_{160/1} were only present in strains 7730, My-I, My-325, My-5 and My-
174 18 (**Figure 3A, Table S4**). We revealed that our candidate TAS were also not present in all
175 members of the '*M. mycoides* cluster' (**Figure 3B, Table S4**).

176

177 **Functionality testing of candidate TAS using heterologous expression**

178 Three of the six candidate TAS were chosen for further characterization. These ones included
179 two candidate TAS identified via TASmania, namely TAS_{160/1} and TAS_{752/3}, whereby TAS_{160/1}
180 represented the candidate system with the overall highest score (e-value of 1.2e-53 and 3.8e-
181 49) and TAS_{752/3} represented the candidate system with the lowest score among the five pairs
182 (e-value of 3.5e-9) identified. Moreover, we included the candidate TAS_{133/2} in our studies,
183 since it was suspected to be a TAS because of polar effects of transposon insertions [15]. We
184 aimed to test toxicity of candidate toxins using heterologous expression. Since mycoplasmas
185 lack a cell wall and represent minimal organisms, we opted for testing heterologous expression
186 in different recipients such as *E. coli*, *B. subtilis* and *M. capricolum*. Consequently we started
187 with *E. coli*, which has been used to confirm TAS in both GRAM-negative and GRAM-positive
188 bacteria [28]. In this bacterial species a bigger toolbox was available than for *Mycoplasma*, like
189 expression vectors with inducible promoters and fast and efficient transformation protocols.
190 Subsequently, genes encoding candidate toxins and antitoxins were cloned into *E. coli* using
191 the established and robust pBAD/His expression system. Heterologous expression confirmed
192 only a negative effect on bacterial growth for the toxin T₁₃₃, while the other two candidate
193 toxins and all three antitoxins tested did not affect the growth of *E. coli* after induction. As
194 expected, growth was only affected when expression of T₁₃₃ was induced, but not when it was

195 repressed. Immunoblotting verified expression of recombinant proteins after induction for all
196 cloned genes (**Figure S3A**). Specifically, four hours post induction, the OD₆₀₀ of *E. coli*
197 transformed with pBAD7His-T₁₃₃ decreased drastically, compared to the various controls
198 (**Figure 4A and Figure S3B**). Interestingly, we observed a phenotype characterized of
199 clumping of *E. coli* cells later in the growth phase in induced *E. coli* cells harboring the vector
200 with T₁₃₃. Then we visualized *E. coli* morphology of those cells using scanning electron
201 microscopy. Visualization revealed that the cell shape started to alter four hours post induction,
202 specifically a fraction of the induced cells was elongated, others were Z-shaped. After seven
203 hours post induction, many cells seem to have burst with only the cell wall remaining visible
204 and a large fraction of cells displayed blebbing at their poles as well as elongation (**Figure 4B**).
205 Empty cell walls clumped together, leading to the clumping observed in liquid cultures by the
206 naked eye. In contrast, cell morphology of all other constructs remained normal (**Figure S3C**).
207 We tested the candidate TAS subsequently in another model organism, namely *B. subtilis*. The
208 rationale behind testing heterologous expression in a GRAM-positive model bacterium was
209 rooted in the closer phylogenetic distance of *B. subtilis* to mycoplasmas that are closely related
210 to the GRAM-positive bacteria and specifically to Clostridia. For these heterologous
211 expression experiments we used the commercial expression system based on pHT01 vector
212 cloned into *B. subtilis* strain 168 Marburg. Construction of the different shuttle plasmids in *E.*
213 *coli* resulted in six plasmids, each containing a gene encoding either a candidate toxin or
214 antitoxin belonging to the three candidate TAS to be functionally tested. All plasmids except
215 the one containing T₁₃₃ revealed transformants in *B. subtilis*. Since all other plasmids were
216 transformable, toxicity associated with T₁₃₃ is likely to have resulted in absence of viable
217 transformants. Even without IPTG-induction, the promotor of pHT01 causes a low background
218 expression level, which might account for toxic effects of T₁₃₃. Out of the two toxins
219 transformed successfully and tested subsequently in a growth assay, only the toxin T₇₅₂ affected

220 bacterial growth after induction with IPTG, while the other toxin and antitoxins tested had
221 growth curves comparable to *B. subtilis* carrying the empty vector pHT01. Specifically, two
222 hours after IPTG induction, cell density of *B. subtilis* harboring T₇₅₂ started to drop, compared
223 to the other induced toxins and antitoxins, indicating lysis (**Figure 5A**, **Figure S4A**). Laser
224 scanning micrographs of *B. subtilis* harboring T₇₅₂, showed an altered cell morphology
225 compared to non-induced cells or cells harboring the empty vector (induced or non-induced)
226 (**Figure 5B**), (**Figure S4B**). A large fraction of cells harboring T₇₅₂ burst, shrunk or had holes
227 on the cell surface. Many cells observed during binary fission had one half burst and the other
228 half intact.

229 Finally, we took advantage of the availability of stably replicating plasmids available for
230 selected members of the ‘*M. mycoides* cluster’. Since our genome comparisons showed that
231 *Mcap* did not harbor any of the three candidate TAS to be investigated for functionality, *Mcap*
232 represented a “natural knock-out” and was an ideal candidate for heterologous expression. We
233 constructed a set of 4 plasmids for each TAS including the entire TAS under the control of the
234 natural promotor, the antitoxin under the control of the natural promotor and the toxin under
235 the control of the natural promotor as well as the strong spiralin promotor. The cloning of the
236 different plasmids was done via Gibson assembly in *E. coli*. Sequence-verified recombinant
237 plasmids (**Supplementary file S1**) were transformed into *Mcap* and transformation efficacies
238 were monitored to assess toxicity. Transformation of only the antitoxins yielded same amount
239 of transformants as transformation of the empty vector (between 10⁵ and 10⁶ transformants
240 per µg plasmid). When only the toxins were transformed, 2 logs less transformants were
241 observed compared to the empty vector, which was a significant result. Number of
242 transformants harboring the toxin with the natural promotor compared to the ones containing
243 the spiralin promotor did not differ significantly, except for T₁₃₃ under the control of the spiralin
244 promotor, which did not result in any transformants at all. Transformation rates of the entire

245 TAS did not differ to the empty vector or the individual antitoxins indicating neutralization of
246 the toxic effects by the antitoxin (**Figure 6**).

247 Assessment of vector and insert presence was tested in 3-5 transformants per transformation
248 experiment. It confirmed the presence of antitoxins and entire TAS as expected. *Mcap* cells
249 transformed with the toxin T₁₃₃ under the control of the natural promotor, were found to contain
250 at most partial toxin sequences. We confirmed the presence of T₇₅₂ sequence with its natural
251 promotor in a fraction of the transformants investigated, but transformants of the toxin under
252 the control with the spiralin promotor did not contain T₇₅₂ sequences. The sequence of T₁₆₀
253 were found in all the examined clones with natural and spiralin promotor.

254 **Donor-recipient phylogenetic distance impacts functional confirmation of TAS**

255 Toxicity testing has been carried out in the three recipient species *E. coli*, *B. subtilis* and *Mcap*
256 listed from most to least phylogenetic distance to the TAS donor species *Mmc*. We started
257 testing the two components of the TAS individually in *E. coli* using the established expression
258 system based on pBAD/His. Only T₁₃₃ conferred toxic activity to its recipient upon induction
259 of expression, indicating TAS_{133/2} to be functional. Next, we proceeded to test the systems in
260 *B. subtilis*, which is phylogenetically much closer to *Mmc* than *E. coli*, but likewise consists of
261 a much bigger proteome and a cell wall, factors that are likely to interfere with TAS
262 functionality. The pHT01 vector used has some basal expression, since its promotor is not tight.
263 It was not possible to transform the vector containing T₁₃₃ into *B. subtilis*, indicating toxicity
264 affecting transformation. Additionally, the growth of the clone carrying T₇₅₂ was impacted after
265 induction of expression. Therefore, two out of three TAS tested functional in *B. subtilis*. In
266 *Mcap* all TAS tested functional and moreover the antitoxins neutralized the toxins. In summary,
267 heterologous expression systems confirmed functionality influenced by donor-recipient
268 phylogenetic distance.

269 Discussion

270 This study aimed to identify and to produce proof of the functionality of toxin antitoxin systems
271 (TAS) in minimalistic mycoplasmas for the first time. Therefore, we employed our model
272 organism, the highly pathogenic *Mycoplasma mycoides* subsp. *capri* strain GM12. Many
273 pathogenic mycoplasmas are associated with chronic infections, which theoretically suggests
274 a role for TAS to support existence in a stealth mode that protects the mycoplasmas from
275 immune responses. In fact, live *Mycoplasma mycoides* subsp. *mycoides*, the causative agent of
276 contagious bovine pleuropneumonia, have been isolated from lung sequestra >6 months after
277 experimental infection [10] and the same mycoplasmas survived even in the presence of high
278 antibody titers against the latter. The *in silico* analysis of strain GM12 revealed 18
279 chromosomally encoded putative candidate TAS elements of which five candidate TAS with
280 the lowest TASmania output e-values were short-listed for our study. In *Mycobacterium*
281 *tuberculosis*, a deadly human respiratory pathogen with a much larger genome, at least 79
282 TAS have been reported to be involved in numerous mechanisms regarding cellular processes
283 such as cell death, persistence and virulence [29-31]. In contrast, *Campylobacter jejuni*, a
284 human pathogen causing diarrhea, has been reported to contain no more than three TAS, which
285 underlines a great variety in the abundance of TAS in different prokaryotic genera [29]. So far,
286 functional TAS in mycoplasmas have never been convincingly shown but only indicated *in*
287 *silico* based on sequence similarity to other TAS [32]. In this study, we focused on the
288 characterization of the TAS with the lowest e-values, since they provided the biggest odds of
289 having a functional TAS. We added the candidate TAS_{133/2}, which was proposed recently [15]
290 based on the analysis of a transposon insertion library of JCVI-syn2.0, a synthetic cell
291 controlled by a drastically reduced genome [2]. These six candidate TAS were grouped into
292 types II and IV (**Table 1**) systems. In type II systems the proteinaceous antitoxin is binding to
293 the toxin to prevent its action, in type IV systems the antitoxin competitively binds the toxin

294 cell target. First, we aimed to identify transcripts and protein signatures of the six systems
295 during in vitro growth indicating a biological function of the genes investigated. We identified
296 transcripts of all TAS elements except for A₈₉₀ supporting a role of transcribed genes in the
297 physiology of GM12. The absence of a transcript for A₈₉₀ does not exclude its transcription,
298 since low levels of transcription and a high turnaround time of transcripts can cause transcripts
299 to be below the detection limit. Protein signatures encoded by the six TAS elements were
300 detected for the entire TAS_{133/2} and the individual elements of four TAS, namely A₁₆₁, A₅₂₄,
301 T₇₃₁ and A₇₅₃. The only system we had no proteomic signatures was the TAS_{889/0} for which we
302 did not detect a transcript of A₈₉₀. The detection of proteomic signatures in five out of six
303 candidate TAS using a generic proteomics approach indicates a high probability of
304 functionality especially for the TAS_{133/2}. The complete genomes of 16 *Mmc* strains revealed a
305 high level of synteny, only two strains had a large inversion in contrast to the other strains.
306 Interestingly, three out of six TAS were present in all 16 *Mmc* genomes investigated, while
307 phylogenetic clades sometimes also had different TAS fingerprints (**Figure 3A**) making
308 horizontal gene transfer likely to have acted on TAS acquisition or loss. Investigation of
309 genomes of the five members of the *M. mycooides* cluster, which consists of three species (two
310 species contain two subspecies) [33], also confirmed the presence of TAS_{752/3}, TAS_{731/0} and
311 TAS_{525/4} in all species investigated (**Figure 3B**) pointing towards old rather conserved
312 acquisitions and a certain plasticity of the genome with respect to TAS. The '*M. mycooides*
313 cluster' is evolutionary young, since its common ancestor dates back 10,000 years [33]
314 coinciding with the start of ruminants' domestication. Therefore, these different TAS
315 fingerprints likely formed within the last 10,000 years and horizontal gene transfer [34-36] is
316 likely to have contributed to it.

317 Next, we tested functionality and shortlisted three candidate TAS for this purpose, based on at
318 least one of the following arbitrary criteria i) a low e-value (<10e-7), ii) TAS-related name for

319 the nearest Pfam annotation of the TASmania HMM profile that hit the locus, iii) TAS-related
320 gene description and iv) presence of candidate toxin and cognate antitoxin in our proteomic
321 analysis. We selected TAS_{133/2}, since protein signatures were reported for its toxin and cognate
322 antitoxin, TAS_{160/1}, since it was the type IV system with the highest TASmania score and
323 TAS_{752/3}, and since it was suggested to be a type II system and its toxin and cognate antitoxin
324 encoding gene had gene annotations. For functional testing we established a model organism-
325 based approach by expressing recombinant toxins and antitoxins in the different recipients *E.*
326 *coli* and *B. subtilis* employing expression plasmids with inducible promoters as well as
327 transforming the candidates in the closely related *Mycoplasma capricolum* subsp. *capricolum*.
328 We started to test toxicity of candidate toxins in *E. coli*, which has been done for other
329 candidate TAS of GRAM-positive bacteria, e.g. Streptomyces [37] or Mycobacteria [28].
330 Unexpectedly, although being high score candidates, two out of three candidate TAS did not
331 test positive in *E. coli*. Only recombinant T₁₃₃, annotated as subtilisin-like serine protease, had
332 a toxic effect on the growth of *E. coli* LMG194. Serine proteases cleave peptides or proteins.
333 A role of serine proteases as toxins in TAS has not been reported except for IetS [23] in the
334 GRAM-negative plant pathogen *Agrobacterium tumefaciens*, a toxin which was shown to be
335 involved in plasmid stability. Heterologous expression of T₁₃₃ in *E. coli* resulted in cell
336 clumping in liquid culture and elongated and filamented cells. The observed morphological
337 changes can be attributed to either direct toxic effects or a stress response of *E. coli* towards
338 the recombinant T₁₃₃, which should be investigated in future studies. Cell filamentation has
339 been reported to occur as an SOS response upon DNA damage in other bacteria [38, 39].
340 Heterologous expression in *B. subtilis* revealed toxicity of recombinant T₁₃₃ (plasmid carrying
341 T₁₃₃ was not transformable) and of T₇₅₂, which had a toxic effect on the growth. The latter
342 encodes for a fic protein, which are known in bacteria to disrupt the DNA topology via
343 adenylation and causes growth arrest [40]. Finally, functionality of all three TAS investigated

344 was shown in the transformation experiments carried out with *Mcap*. These experiments clearly
345 also showed the capacity of the antitoxins to neutralize the effects of the toxins. Interestingly,
346 the type IV system TAS_{160/1} only showed functionality in *Mcap*, probably due to the fact that
347 in such systems toxin and antitoxin bind to the same target, which was probably only available
348 in a mycoplasma.

349 In conclusion, this study identified six candidate TAS and tested functionality of three out of
350 the six chromosomally encoded TAS in *Mmc* GM12. Functionality of the three candidate TAS
351 was shown, as expression of only the toxin results in perturbations of cell vitality or in cell
352 death whereas the cognate antitoxin compensated the toxin-induced effects completely.
353 Further, we evaluated different model organisms as *E. coli*, *B. subtilis* and *Mycoplasma*
354 *capricolum* for heterologous expression of TAS. According to our results, toxin action seems
355 to be dependent on phylogenetic proximity towards the bacterial species from which candidate
356 TAS originate. TAS patterns were different in different *Mycoplasma* species and subspecies
357 indicating a level of plasticity of the mycoplasmas' genomes with respect to TAS. Since
358 mycoplasmas are minimal organisms [11] amendable to synthetic genomics techniques [41],
359 their host-pathogen interactions can be investigated in the native host [17, 18, 42] and as they
360 contain only a limited number of TAS, they represent an ideal model organism to study the
361 role of TAS in pathogenicity.

362 **Materials and methods**

363 ***In silico* identification of candidate TAS**

364 Identification of candidate TAS in *Mmc* strain GM12 [16] was performed by employing the
365 TASmania discovery pipeline [20], which uses data mining on the Ensembl Bacteria database
366 [43] to identify candidate toxin and antitoxin proteins. The e-value was set to the default value
367 of 1e-05. TAS candidates to be characterized in the study had the following minimal
368 requirements: an e-value lower than 1e-07, a protein family (pfam) describing known TAS and
369 gene description indicating possible candidates. If TAS candidates were without a counterpart
370 (toxin or antitoxin), the counterparts were manually selected using ‘guilt by association’,
371 considering the cognate loci of the respective gene as the partner of a two gene operon.
372 Intergenic regions close to toxin and antitoxin genes on the *Mmc* GM12 genome were analyzed
373 using BPROM (Softberry) [44] to identify putative promotor regions.

374

375 **Bacterial strains and plasmids used for cloning in this study**

376 Plasmid based on pBAD/His (Invitrogen) or pHT01 (MoBiTech) and pMYCO1 were
377 constructed in *E. coli* DH5 α (Takara) and StellarTM (Takara), respectively. Heterologous
378 expression based on pBAD/His and pHT01 was done in *E. coli* LMG194 (Invitrogen) and *B.*
379 *subtilis* 168 Marburg (MoBiTech), respectively. The restriction-free *M. capricolum* subsp.
380 *capricolum* mutant strain *M. capricolum* RE(-) [41] was used as recipient for transformation
381 experiments involving pMYCO1-derived plasmids [45].

382

383 **Detection of transcripts for candidate TAS**

384 *Mycoplasma mycoides* subsp. *capri* (*Mmc*) strain GM12 was grown in 10 ml SP5 medium at
385 37°C until the culture reached pH 6.5. Subsequently, the culture was centrifuged at 4,255 x g
386 at 10°C for 15 min. The supernatant was discarded and the bacterial cell pellet was used for

387 RNA isolation. RNA was isolated using the Zymo Research Quick-RNA Fungal/Bacterial
388 Miniprep™ kit. Briefly, 800 µl of RNA Lysis Buffer were added to the pellet and the protocol
389 was followed starting at step 4 of the manual. After isolation, a DNase treatment was performed
390 using the Clean & Concentrator™ -5 kit (ZymoResearch), omitting the first step of the protocol
391 (addition of Binding Buffer) and using the IICR columns provided in the kit. Reverse
392 transcription PCR was performed using SuperScript™ IV RT Mix (Invitrogen) and 2x
393 Platinum™ SuperFi™ RT-PCR Master Mix (Invitrogen) using primers No. 59-82 listed in
394 **Table S4**. In brief, 12.5 µl of 2x Platinum™ SuperFi™ RT-PCR Master Mix, 0.25 µl
395 SuperScript™ IV RT Mix, 1.25 µl per primer (10 µM), 1µl DNase treated RNA (~ 32 ng) and
396 water up to a volume of 25 µl were mixed per sample. Amplicons were visualized on a 1%
397 agarose gel.

398

399 **Proteomic analysis**

400 Strain GM12 was grown in 50 mL of SP5 media until mid to late logarithmic phase (pH of
401 media at harvest was 6.5) and centrifuged at 3,000 x g for 20 min. Cells were washed three
402 times with PBS and pellets were kept at -80 for subsequent analysis. The protein concentration
403 was measured using the Pierce BCA Protein Assay Kit (ThermoScientific) according to
404 vendor's instructions. Samples for mass spectrometry analysis were prepared following
405 standard protocols [46]. Briefly, cell pellets were thawed on ice and lysed in 8M urea/100mM
406 Tris-HCl; proteins were precipitated in cold acetone over-night and denatured with 8M
407 urea/50mM Tris-HCl, then digested with trypsin over-night at room-temperature in 1.6M urea
408 /20mM Tris-Hcl /2mM CaCl₂. Enzymatic digestion was stopped by adding 1/20-volume of
409 20% (v/v) tri-fluoroacetic (TFA). The digests were analyzed by liquid chromatography (LC)-
410 MS/MS (PROXEON coupled to a QExactive HF mass spectrometer, ThermoFisher Scientific).
411 Samples were further processed against custom databases by Transproteomics pipeline (TPP)

412 tools [47]. Four database search engines were used: Comet [48], X tandem [49], MSGF [50]
413 and MyriMatch [51]. Each search was followed by the application of the PeptideProphet tool
414 [52]; the iProphet [53] tool was then used to combine the search results, which were filtered at
415 the false discovery rate of 0.01; furthermore, the identification was only accepted if at least two
416 of the search engines agreed on the identification. The decoy approach was used for such
417 custom databases containing standard entries. Protein inference was performed with
418 ProteinProphet. For those protein groups accepted by a false discovery rate filter of 0.01, a
419 Normalized Spectral Abundance Factor (NSAF) [54] was calculated based on the peptide to
420 spectrum match count; shared peptides were accounted for by the method published elsewhere
421 [55].

422

423 **Cloning of genes encoding candidate toxins and antitoxins in *Escherichia coli***

424 *E. coli*-codon optimized synthetic genes encoding candidate toxins or antitoxins were
425 synthesized and cloned into pUC57 and sequence-verified by GenScript (**Supplementary File**
426 **S1**). For ligation into the expression vector pBAD/His, genes to be cloned, were PCR-amplified
427 using primers that contained restriction sites for *EcoRI* and *SacI* (**Table S5**, primers No. 1-6)
428 and the high-fidelity polymerase Q5 (New England Biolabs) according to vendor's manual.
429 Amplicons were purified using the High Pure PCR Product Purification Kit (Roche) and
430 afterwards restricted with *EcoRI* (New England Biolabs) and *SacI* (New England Biolabs)
431 followed by another purification using the above-mentioned kit. Toxin genes to be cloned were
432 retrieved from the plasmids that contained the synthetic genes via restriction followed by gel
433 purification. Restricted fragments to be cloned were ligated into the *EcoRI* and *SacI* restricted
434 pBAD/His using T4 Ligase (Promega) before being heat shock-transformed into chemically
435 competent *E. coli* DH5 α (Takara) or StellarTM (Takara) using standard protocols [56].
436 Recombinant plasmids from transformants were isolated using the QIAprep[®] Spin Miniprep

437 Kit (Qiagen) and sequence-verified via Sanger sequencing (Microsynth) with specific primers
438 (**Table S5**, primers No. 7-8). Sequence-verified plasmids were heat shock-transformed into
439 chemically competent *E. coli* LMG194 for subsequent expression. Transformants were verified
440 as described above and stored at -80 °C for subsequent experiments.

441

442 **Functional testing of candidate *Mycoplasma* TAS using heterologous expression in *E. coli***

443 Growth curves of *E. coli* LMG194 containing inducible expression vector pBAD/His carrying
444 genes encoding individual candidate toxins and antitoxins were performed in 96-well plates
445 (Tissue culture test plates, 92096, TPP) and monitored over a period of seven hours. All clones
446 were cultured in LB medium supplemented with 50 µg/mL ampicillin. Twenty µL of overnight
447 cultures of clones to be tested were mixed with 180 µL LB supplemented with either 0.2%
448 (w/v) arabinose or 0.2% (w/v) glucose to induce or repress protein expression, respectively.

449 Cultures in 96-well plates were incubated at 37 °C under agitation (500 rpm) using a
450 ThermoMixerC (Eppendorf). Optical density was measured at 600nm using the VERSAmax
451 plate reader (Bucher Biotec, Basel, Switzerland). Final datasets were obtained from three
452 biological replicates (each biological replicate consisted of three technical replicates). To
453 assess the significance of the heterologous expression on growth of the clones at every time
454 point, we applied a three-way (time, treatment, construct) repeated-measures ANOVA and
455 decomposed into post hoc tests based on pairwise comparisons with a Bonferroni adjustment
456 of the p-value. The statistical analysis was done using R version 3.6.2.

457

458 **Detection of recombinant proteins using immunoblots**

459 Heterologous expression in *E. coli* was verified by immunoblot using anti-HIS antibodies
460 tagging the recombinant fusion proteins. Briefly, 200 µL of overnight culture were mixed with
461 20 ml of LB supplemented with 50 µg/mL ampicillin and arabinose (final concentration 0.2%).

462 Cultures were grown on a shaking incubator at 220 rpm (Lab-Shaker) at 37 °C and samples of
463 1 mL were sequentially removed at several time points for subsequent analysis. As controls
464 served the same clones grown in the same media except that arabinose was exchanged with
465 glucose (final concentration 0.2%). Samples were spun down at 16,100 x g at room temperature
466 for 3 min. The supernatants were discarded, the pellets were resuspended in PBS (Merck) and
467 the protein concentration was determined using the Pierce™ BCA Protein Assay Kit
468 (ThermoScientific). Immunoblotting was basically carried out as described recently [18].
469 Briefly, 0.1 mg of proteins in the cell pellet were separated onto a 12% SDS-PAGE gel using
470 standard procedures [56]. Size separated proteins were transferred onto a 0.2 µm pore-size
471 nitrocellulose membrane (BIO RAD) using the Trans-Blot® Turbo Transfer System (BIO
472 RAD) at 25 volts, 1.0 A for 30 min. PBS supplemented with 0.1% Tween-20 (Merck) and 2%
473 BSA (Sigma) served as blocking buffer and monoclonal mouse-derived anti-HIS antibody
474 (LS-C57341, LsBio) and HRP conjugated goat anti-mouse IgG antibody (AP308P, Sigma)
475 served as primary and secondary antibodies, respectively. Primary antibodies were diluted in
476 blocking buffer at 1:1,000 and incubated with the membrane for 1 hour, afterwards the
477 membrane was washed 3 times in PBS supplemented with 0.1% Tween-20 for 10 min. Then
478 the membrane was incubated with the secondary antibody diluted at 1:70,000 in blocking
479 buffer for 1 hour. After three washes as done before the membrane was incubated with
480 SuperSignal™ West Pico PLUS Chemiluminescent Substrate (ThermoScientific) according to
481 manufacturer's recommendations. Chemiluminescence was visualized with the CCD-camera
482 Fusion FX (Vilber) and analyzed by the software Evolution.

483

484 **Cloning of genes encoding candidate toxins and antitoxins in *Bacillus subtilis***

485 Synthetic genes to be cloned into *B. subtilis* were first cloned into *E. coli* using the shuttle
486 vector pHT01 (MoBiTec). Therefore, genes were PCR-amplified using synthetic genes cloned

487 into pUC57 (see above) as template, primers that contained restriction sites for *BamHI* and
488 *XmaI* or *XbaI* (**Table S5**, primers No. 9-16 and 95-98) and the high-fidelity polymerase Q5
489 (New England Biolabs) according to vendor's manual. Amplicons were purified using the High
490 Pure PCR Product Purification Kit (Roche) and afterwards restricted with *BamHI*
491 (ThermoScientific) and *XmaI* (ThermoScientific) or *XbaI* (ThermoScientific) followed by
492 another purification step using the above-mentioned kit. The restricted amplicons were ligated
493 into the *BamHI* and *XmaI* (*BamHI* and *XbaI*) restricted pHT01 using T4 Ligase (Promega)
494 before being heat shock-transformed into chemically competent *E. coli* Stellar™ (Takara)
495 following standard protocols [56]. *E. coli* Stellar™ was grown on LB media supplemented with
496 ampicillin at 50 µl/mL. Colony PCR was performed with primers No. 17-24 and 99-102 (**Table**
497 **S5**). Recombinant plasmids from transformants were isolated using the QIAprep® Spin
498 Miniprep Kit (Qiagen) and sequence verified via Sanger sequencing (Microsynth) with specific
499 primers (**Table S5**, No. 25-26). Sequence-verified plasmids were transformed into *Bacillus*
500 *subtilis* 168 Marburg (MoBiTec) using vendor's protocols. Therefore, we followed the protocol
501 for making naturally competent cells. Transformants were selected on agar plates containing 5
502 µg/mL chloramphenicol. Transformants were lysed as described elsewhere [57] and presence
503 of plasmids was confirmed via PCR targeting cloned genes. Clones were stored at -80 °C for
504 subsequent experiments.

505

506 **Functional testing of candidate TAS using heterologous expression in *B. subtilis***

507 Growth curves of *B. subtilis* 168 Marburg containing inducible expression vector pHT01
508 carrying genes encoding individual candidate toxins and antitoxins were performed in 96-well
509 plates (Tissue culture test plates, 92096, TPP) and monitored over a period of seven hours. All
510 clones were cultured in LB medium supplemented with 5 µg/mL chloramphenicol (Serva).
511 Overnight cultures of *B. subtilis* 168 Marburg were diluted to an OD_{600nm} of 0.15 in a total

512 volume of 5 mL, which was subsequently incubated at 37 °C. When cultures reached an OD₆₀₀
513 of 0.7-0.8, IPTG (Roche Diagnostics) was added to a final concentration of 1 mM for induction
514 of heterologous protein expression. Cultures in 96-well plates were incubated at 37 °C under
515 agitation (500 rpm) using a ThermoMixerC (Eppendorf). Optical density was measured at 600
516 nm using the VERSAmax plate reader (Bucher Biotec). Final datasets were obtained from three
517 biological replicates (each biological replicate consisted of three technical replicates).
518 To assess the significance of the heterologous expression on growth of the clones at every time
519 point, we applied a three-way (time, treatment, construct) repeated-measures ANOVA and
520 decomposed into post hoc tests based on pairwise comparisons with a Bonferroni adjustment
521 of the *p*-value. The statistical analysis was done using R version 3.6.2.

522

523 **Cloning of genes encoding candidate toxins and antitoxins into *Mycoplasma capricolum***
524 **subsp. *capricolum***

525 The three candidate TAS operons, as well as each of the genes encoding candidate toxins and
526 antitoxins, were individually cloned into pMYCO1. Candidate genes or operons were PCR
527 amplified using Q5 high fidelity DNA polymerase (New England Biolabs) according to
528 manufacturer's instructions using primers (No. 27-49) listed in **Table S5**. DNA assembly was
529 performed in *E. coli* using the NEBuilder HiFi DNA Assembly kit (New England Biolabs)
530 following manufacturers' instructions. When necessary, the natural promoters of each of the
531 TAS were amplified independently and positioned in frame with their respective candidate
532 toxins. Similarly, the spiralin promoter was amplified from the plasmid pMT85tetM-PSlacZ-
533 pRS313 (GenBank accession number KX011460) and assembled upstream of each toxin. All
534 the constructions were sequenced using primers No. 57 and 94 (**Table S5**).

535

536 **Functional testing of candidate TAS using heterologous expression in *M. capricolum***
537 **subsp. *capricolum* (RE-)**

538 *M. capricolum* subsp. *capricolum* (*Mcap*) (RE-) was selected to probe functionality of the three
539 candidate TAS as none of the latter appeared to be present in its genome (**Figure 3B**). All 12
540 pMYCO1-based constructs were transformed into *Mcap* (RE-) as previously described [58]
541 using 1 µg of plasmid DNA. Three biological replicates including three technical replicates per
542 biological replicate were carried out. Transformants were passaged for three consecutive
543 rounds in 1ml SP5 medium supplemented with 5 µg/mL tetracycline and incubated at 37 °C to
544 confirm the resistant phenotype. Transformants were screened using primer pairs specific for
545 the inserted TAS elements (**Table S5**, No. 50-58). Numbers of transformants were used to
546 analyse transformation efficacies. An ordinary one-way ANOVA test with Tukey's multiple
547 comparison test was applied (GraphPad Prism version 8.0.0, San Diego, California USA).

548

549 **Scanning electron microscopy (SEM) of *E. coli* and *B. subtilis* recombinant clones**

550 The morphology of *E. coli* and *B. subtilis* cells upon toxin induction was investigated by
551 scanning electron microscopy. Clones to be visualized were cultured and expression of the
552 candidate toxins was induced as described above. Samples collected during expression (1 ml)
553 were centrifuged at 1,500 x g at room temperature for 3 min. The pellets were washed 3 times
554 with 1 ml PBS and finally resuspended in 250 µl PBS. Subsequently, 250 µl of 5%
555 glutaraldehyde (Merck) in 0.2M cacodylate buffer, pH 7.4, was slowly added to the samples
556 with gentle agitation. The tubes were stored at 4 °C until further processing. Fixed cells were
557 cytospun onto platinum-sputtered and PLL-coated coverslips (40 µl at 125 x g, 5 min). The
558 coverslips were then washed 3 times with 0.1M cacodylate buffer, pH 7.4, postfixed in 1%
559 OsO4 (Polysciences) in 0.1M cacodylate buffer for 30 min, washed 3 times again and
560 dehydrated in an ascending series of ethanol (70%, 80%, 94%, 100%, 100%, 100% ethanol at

561 room temperature for 15 min each). Samples were dried by evaporation of
562 Hexamethyldisilazane (Merck). Samples were coated with 15 nm of platinum in a high vacuum
563 coating unit (CCU-010; Safematic) and examined with a scanning electron microscope DSM
564 982 Gemini (Zeiss) at an accelerating voltage of 5 kV at a working distance of 4 mm.

565

566 **Isolation of genomic DNA from for next generation sequencing**

567 Mycoplasmas grown in 8 ml of SP5 medium overnight were pelleted at 4,255 x g at 10 °C for
568 15 min. The supernatant was discarded and gDNA was extracted using the Wizard Genomic
569 DNA purification kit (Promega) according to vendor's protocol. Plasmid DNA was extracted
570 from 10 ml overnight culture using QIAprep® Spin Miniprep Kit (QIAGEN) with the
571 following adaptations: additional washing with buffer PB and using the double of the
572 recommended volume of buffers P1, P2 and P3. DNA concentration and purity were
573 determined by agarose gel separation, Nanodrop (ND-1000 Spectrophotometer, Witec AG) and
574 Qubit Fluorometric quantification (Invitrogen).

575

576 **Next generation sequencing of *Mycoplasma mycoides* subsp. *capri* strains**

577 Sequencing was carried out at the Lausanne Genomic Technologies Facility using the PacBio.
578 Briefly, DNA was sheared in a Covaris g-TUBE (Covaris, Woburn, MA, USA) to obtain 10
579 kbp fragments. After shearing, the DNA size distribution was checked on a Fragment Analyzer
580 (Advanced Analytical Technologies, Ames, IA, USA). A barcoded SMRTbell library was
581 prepared using 480 ng of gDNA using the PacBio SMRTbell Template Express Prep Kit 2.0
582 (Pacific Biosciences, Menlo Park, CA, USA) according to the manufacturer's
583 recommendations. Libraries were pooled and sequenced with v3.0/v3.0 chemistry on a PacBio
584 Sequel instrument (Pacific Biosciences, Menlo Park, CA, USA) at 10 hours movie time, pre-
585 extension time of 2 hours, using one SMRT cell v3.

586

587 **Genome assembly, annotation and alignments**

588 Genomes were assembled from PacBio reads using the software Flye, version 2.6 [59].
589 Circularized genomes were polished with three rounds with the software Arrow [single-
590 molecule real-time (SMRT) Link version 8 package]. Plasmids were Sanger sequenced after
591 cloning into *E. coli* using pUC19. Genomes were rotated to the first nucleotide of the start
592 codon of the *dnaA* gene. Sequences were then annotated using Prokka, version 1.13 [60]. The
593 program Mauve (version 20150226) [61] was used for the construction of the genome
594 alignment of multiple *Mycoplasma* strains.

595

596 **Generation of a phylogenetic tree of members of the ‘*M. mycoides* cluster’**

597 The phylogenetic tree was generated using 16S rRNA sequences retrieved from GenBank. The
598 following strains were included in the analysis: *Mycoplasma mycoides* subsp. *capri* strain
599 GM12 (GenBank accession number NZ_CP001668.1), *Mycoplasma mycoides* subsp. *mycoides*
600 strain PG1^T (GenBank accession number BX293980.2), *Mycoplasma capricolum* subsp.
601 *capricolum* strain California kid^T (GenBank accession number NC_007633.1), *Mycoplasma*
602 *capricolum* subsp. *capripneumoniae* strain F38^T (GenBank accession number LN515398.1),
603 *Mycoplasma leachii* strain PG50^T (GenBank accession number NC_014751.1) and
604 *Mycoplasma putrefaciens* strain KS1^T (GenBank accession number NC_015946.1). The
605 software MEGA, version 10.0.5 [62] was used for construction of phylogenetic trees. For the
606 phylogenetic tree displaying the *M. mycoides* cluster, evolutionary history was inferred using
607 the Neighbor-Joining method and evolutionary distance was computed with Kimura 2-
608 parameter method. For statistical support of the phylogeny, 100 replicates were used for the
609 bootstrap test.

610

611 **Generation of phylogenetic trees of *Mmc* strains using whole genome sequences**

612 For tree generation, whole genome sequences of *Mycoplasma mycoides* subsp. *capri* strains
613 were used and *Mycoplasma mycoides* subsp. *mycoides* PG1^T was included as an outgroup.
614 Mapping of genomes was performed by the Reference sequence Alignment based Phylogeny
615 builder (REALPHY) [63]. For tree construction, the software MEGA (version 10.0.5) [62] was
616 used. Evolutionary history was inferred using the Neighbor-joining method. Evolutionary
617 distances were computed with Maximum Composite Likelihood method. Rate variation among
618 sites was modelled with a gamma distribution.

619

620 **Graphic representation of toxin and antitoxin presence on genomes of mycoplasmas**

621 A graphic representation of gene presence was generated based on TBLASTN similarity of
622 candidate toxin and antitoxin amino acid sequences and genomes of different mycoplasmas.
623 The presence and conservation of each toxin and antitoxin candidate was assessed using
624 TBLASTN [64] with *Mmc* GM12 amino acid sequences of candidate toxins and antitoxins as
625 query against whole genome sequences of different *Mycoplasmas* and *Mmc* strains. All
626 sequences with query coverage >99% and percent identity >80% were considered as being
627 present on the respective genome. If Query coverage was <55%, we did assume the candidate
628 toxin/antitoxin does not exist in the respective genome. In case the query coverage was >70%
629 but percent identity was <61%, a BLASTN was additionally performed to check for nucleotide
630 sequence similarity. If the nucleotide sequence identity query coverage of BLASTN was >80%
631 and percent identity >70%, we assumed the candidate toxin/antitoxin to be present on the
632 respective genome.

633

634 **Author contributions:** JJ designed the research. VH, FL, RSB, HAE, MHS performed the
635 research. VH, FL, HAE, LF, ManH and JJ analyzed the data. MarH provided *Mycoplasma*

636 *mycoides* subsp. *capri* strains, ManH the proteomics infrastructure. VH, FL and JJ drafted the
637 manuscript. All authors read and approved the final manuscript.

638

639 **Acknowledgements**

640 We thank Helga Mogel, Sybille Schwendener, Simon Feyer and Isabelle Brodard (University
641 of Bern, Switzerland) for their technical help. We thank Joerg Stülke for provision of the
642 *Bacillus subtilis* strain 168 Marburg and the plasmid pHT01. We thank Patrick Viollier, Sanjay
643 Vashee and Volker Thiel for valuable discussions.

644 **References**

645

646 1. Gibson DG, Glass JI, Lartigue C, Noskov VN, Chuang RY, Algire MA, et al.
647 Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome. *Science*.
648 2010;329(52):52-6. Epub 2010/05/22. doi: science.1190719 [pii]

649 10.1126/science.1190719. PubMed PMID: 20488990.

650 2. Hutchison CA, 3rd, Chuang RY, Noskov VN, Assad-Garcia N, Deerinck TJ, Ellisman
651 MH, et al. Design and synthesis of a minimal bacterial genome. *Science*.
652 2016;351(6280):aad6253. doi: 10.1126/science.aad6253. PubMed PMID: 27013737.

653 3. Rogers MJ, Simmons J, Walker RT, Weisburg WG, Woese CR, Tanner RS, et al.
654 Construction of the mycoplasma evolutionary tree from 5S rRNA sequence data. *Proc Natl*
655 *Acad Sci U S A*. 1985;82(4):1160-4. Epub 1985/02/01. doi: 10.1073/pnas.82.4.1160.
656 PubMed PMID: 2579388; PubMed Central PMCID: PMC397214.

657 4. McGowin CL, Totten PA. The Unique Microbiology and Molecular Pathogenesis of
658 *Mycoplasma genitalium*. *J Infect Dis*. 2017;216(suppl_2):S382-S8. Epub 2017/08/26. doi:
659 10.1093/infdis/jix172. PubMed PMID: 28838077; PubMed Central PMCID:
660 PMC5853509.

661 5. Waites KB, Xiao L, Liu Y, Balish MF, Atkinson TP. *Mycoplasma pneumoniae* from
662 the Respiratory Tract and Beyond. *Clin Microbiol Rev*. 2017;30(3):747-809. Epub
663 2017/05/26. doi: 10.1128/CMR.00114-16. PubMed PMID: 28539503; PubMed Central
664 PMCID: PMC5475226.

665 6. Di Teodoro G, Marruchella G, Di Provvido A, D'Angelo AR, Orsini G, Di Giuseppe
666 P, et al. Contagious Bovine Pleuropneumonia: A Comprehensive Overview. *Vet Pathol*.

- 667 2020:300985820921818. Epub 2020/05/12. doi: 10.1177/0300985820921818. PubMed
668 PMID: 32390522.
- 669 7. Yattoo IM, Raffiq Parray O, Tauseef Bashir S, Ahmed Bhat R, Gopalakrishnan A,
670 Karthik K, et al. Contagious caprine pleuropneumonia - a comprehensive review. *Vet Q.*
671 2019;39(1):1-25. Epub 2019/04/02. doi: 10.1080/01652176.2019.1580826. PubMed PMID:
672 30929577.
- 673 8. Maes D, Sibila M, Kuhnert P, Segales J, Haesebrouck F, Pieters M. Update on
674 *Mycoplasma hyopneumoniae* infections in pigs: Knowledge gaps for improved disease
675 control. *Transbound Emerg Dis.* 2018;65 Suppl 1:110-24. Epub 2017/08/24. doi:
676 10.1111/tbed.12677. PubMed PMID: 28834294.
- 677 9. Liljander A, Sacchini F, Stoffel MH, Schieck E, Stokar-Regenscheit N, Labroussaa F,
678 et al. Reproduction of contagious caprine pleuropneumonia reveals the ability of convalescent
679 sera to reduce hydrogen peroxide production in vitro. *Vet Res.* 2019;50(1):10. Epub
680 2019/02/10. doi: 10.1186/s13567-019-0628-0. PubMed PMID: 30736863; PubMed Central
681 PMCID: PMC6368817.
- 682 10. Schieck E, Liljander A, Hamsten C, Gicheru N, Scacchia M, Sacchini F, et al. High
683 antibody titres against predicted *Mycoplasma* surface proteins do not prevent sequestration in
684 infected lung tissue in the course of experimental contagious bovine pleuropneumonia.
685 *Veterinary Microbiology.* 2014;172(1-2):285-93. Epub 2014/06/02. doi:
686 10.1016/j.vetmic.2014.04.018. PubMed PMID: 24880898.
- 687 11. Citti C, Blanchard A. Mycoplasmas and their host: emerging and re-emerging
688 minimal pathogens. *Trends Microbiol.* 2013;21(4):196-203. Epub 2013/02/20. doi:
689 10.1016/j.tim.2013.01.003. PubMed PMID: 23419218.

- 690 12. Lobato-Marquez D, Diaz-Orejas R, Garcia-Del Portillo F. Toxin-antitoxins and
691 bacterial virulence. FEMS Microbiol Rev. 2016;40(5):592-609. Epub 2016/08/01. doi:
692 10.1093/femsre/fuw022. PubMed PMID: 27476076.
- 693 13. Song S, Wood TK. Toxin/Antitoxin System Paradigms: Toxins Bound to Antitoxins
694 Are Not Likely Activated by Preferential Antitoxin Degradation. Adv Biosyst.
695 2020;4(3):e1900290. Epub 2020/04/16. doi: 10.1002/adbi.201900290. PubMed PMID:
696 32293143.
- 697 14. Kannan TR, Baseman JB. ADP-ribosylating and vacuolating cytotoxin of
698 *Mycoplasma pneumoniae* represents unique virulence determinant among bacterial
699 pathogens. Proc Natl Acad Sci U S A. 2006;103(17):6724-9. Epub 2006/04/18. doi:
700 0510644103 [pii]
701 10.1073/pnas.0510644103. PubMed PMID: 16617115; PubMed Central PMCID:
702 PMC1458948.
- 703 15. Hutchison CA, 3rd, Merryman C, Sun L, Assad-Garcia N, Richter RA, Smith HO, et
704 al. Polar Effects of Transposon Insertion into a Minimal Bacterial Genome. J Bacteriol.
705 2019;201(19). Epub 2019/07/03. doi: 10.1128/JB.00185-19. PubMed PMID: 31262838;
706 PubMed Central PMCID: PMC6755753.
- 707 16. DaMassa AJ, Brooks DL, Adler HE. Caprine mycoplasmosis: widespread infection in
708 goats with *Mycoplasma mycoides* subsp *mycoides* (large-colony type). Am J Vet Res.
709 1983;44(2):322-5. Epub 1983/02/01. PubMed PMID: 6338774.
- 710 17. Jores J, Schieck E, Liljander A, Sacchini F, Posthaus H, Lartigue C, et al. In vivo role
711 of capsular polysaccharide in *Mycoplasma mycoides*. The Journal of Infectious Diseases.

712 2019;219(10):1559-63. Epub 2018/12/13. doi: 10.1093/infdis/jiy713. PubMed PMID:
713 30541131.

714 18. Jores J, Ma L, Ssajjakambwe P, Schieck E, Liljander A, Chandran S, et al. Removal
715 of a Subset of Non-essential Genes Fully Attenuates a Highly Virulent *Mycoplasma* Strain.
716 Front Microbiol. 2019;10:664. Epub 2019/04/20. doi: 10.3389/fmicb.2019.00664. PubMed
717 PMID: 31001234; PubMed Central PMCID: PMC6456743.

718 19. Arfi Y, Minder L, Di Primo C, Le Roy A, Ebel C, Coquet L, et al. MIB-MIP is a
719 mycoplasma system that captures and cleaves immunoglobulin G. Proc Natl Acad Sci U S A.
720 2016;113(19):5406-11. doi: 10.1073/pnas.1600546113. PubMed PMID: 27114507; PubMed
721 Central PMCID: PMC4868467.

722 20. Akarsu H, Bordes P, Mansour M, Bigot DJ, Genevaux P, Falquet L. TASmania: A
723 bacterial Toxin-Antitoxin Systems database. PLoS Comput Biol. 2019;15(4):e1006946. Epub
724 2019/04/26. doi: 10.1371/journal.pcbi.1006946. PubMed PMID: 31022176; PubMed Central
725 PMCID: PMC6504116.

726 21. Hampton HG, Jackson SA, Fagerlund RD, Vogel AIM, Dy RL, Blower TR, et al.
727 AbiEi Binds Cooperatively to the Type IV abiE Toxin-Antitoxin Operator Via a Positively-
728 Charged Surface and Causes DNA Bending and Negative Autoregulation. J Mol Biol.
729 2018;430(8):1141-56. doi: 10.1016/j.jmb.2018.02.022. PubMed PMID: 29518409.

730 22. Gotfredsen M, Gerdes K. The Escherichia coli relBE genes belong to a new toxin-
731 antitoxin gene family. Molecular Microbiology. 1998;29(4):1065-76. doi: DOI
732 10.1046/j.1365-2958.1998.00993.x. PubMed PMID: WOS:000075451700013.

733 23. Yamamoto S, Kiyokawa K, Tanaka K, Moriguchi K, Suzuki K. Novel toxin-antitoxin
734 system composed of serine protease and AAA-ATPase homologues determines the high level

- 735 of stability and incompatibility of the tumor-inducing plasmid pTiC58. *J Bacteriol.*
736 2009;191(14):4656-66. Epub 2009/05/19. doi: 10.1128/JB.00124-09. PubMed PMID:
737 19447904; PubMed Central PMCID: PMCPMC2704731.
- 738 24. Zhang Y, Zhang J, Hoeflich KP, Ikura M, Qing G, Inouye M. MazF cleaves cellular
739 mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. *Mol Cell.*
740 2003;12(4):913-23. doi: 10.1016/s1097-2765(03)00402-7. PubMed PMID: 14580342.
- 741 25. UniProt C. UniProt: a worldwide hub of protein knowledge. *Nucleic acids research.*
742 2019;47(D1):D506-D15. doi: 10.1093/nar/gky1049. PubMed PMID: 30395287; PubMed
743 Central PMCID: PMCPMC6323992.
- 744 26. King KW, Dybvig K. Nucleotide sequence of *Mycoplasma mycoides* subspecies
745 *mycoides* plasmid pKMK1. *Plasmid.* 1992;28(1):86-91. Epub 1992/07/01. doi: 10.1016/0147-
746 619x(92)90039-d. PubMed PMID: 1518915.
- 747 27. Breton M, Tardy F, Dordet-Frisoni E, Sagne E, Mick V, Renaudin J, et al.
748 Distribution and diversity of mycoplasma plasmids: lessons from cryptic genetic elements.
749 *BMC Microbiol.* 2012;12:257. Epub 2012/11/14. doi: 10.1186/1471-2180-12-257. PubMed
750 PMID: 23145790; PubMed Central PMCID: PMCPMC3541243.
- 751 28. Lepplae R, Geeraerts D, Hallez R, Guglielmini J, Dreze P, Van Melderen L. Diversity
752 of bacterial type II toxin-antitoxin systems: a comprehensive search and functional analysis
753 of novel families. *Nucleic acids research.* 2011;39(13):5513-25. Epub 2011/03/23. doi:
754 10.1093/nar/gkr131. PubMed PMID: 21422074; PubMed Central PMCID:
755 PMCPMC3141249.

- 756 29. Van Melder L, Saavedra De Bast M. Bacterial toxin-antitoxin systems: more than
757 selfish entities? PLoS Genet. 2009;5(3):e1000437. doi: 10.1371/journal.pgen.1000437.
758 PubMed PMID: 19325885; PubMed Central PMCID: PMC2654758.
- 759 30. Wen Y, Behiels E, Devreese B. Toxin-Antitoxin systems: their role in persistence,
760 biofilm formation, and pathogenicity. Pathog Dis. 2014;70(3):240-9. doi: 10.1111/2049-
761 632X.12145. PubMed PMID: 24478112.
- 762 31. Tandon H, Sharma A, Sandhya S, Srinivasan N, Singh R. *Mycobacterium*
763 *tuberculosis* Rv0366c-Rv0367c encodes a non-canonical PezAT-like toxin-antitoxin pair. Sci
764 Rep. 2019;9(1):1163. Epub 2019/02/06. doi: 10.1038/s41598-018-37473-y. PubMed PMID:
765 30718534; PubMed Central PMCID: PMC6362051.
- 766 32. Akarsu H, Bordes P, Mansour M, Bigot D-J, Genevaux P, Falquet L. TASmania: A
767 bacterial Toxin-Antitoxin Systems database. PLOS Computational Biology.
768 2019;15(4):e1006946. doi: 10.1371/journal.pcbi.1006946.
- 769 33. Fischer A, Shapiro B, Muriuki C, Heller M, Schnee C, Bongcam-Rudloff E, et al. The
770 Origin of the '*Mycoplasma mycoides* Cluster' Coincides with Domestication of Ruminants.
771 PLoS One. 2012;7(4):e36150. Epub 2012/05/05. doi: 10.1371/journal.pone.0036150.
772 PubMed PMID: 22558362.
- 773 34. Thiaucourt F, Manso-Silvan L, Salah W, Barbe V, Vacherie B, Jacob D, et al.
774 *Mycoplasma mycoides*, from "*mycoides* Small Colony" to "*capri*". A microevolutionary
775 perspective. BMC Genomics. 2011;12:114. Epub 2011/02/18. doi: 10.1186/1471-2164-12-
776 114. PubMed PMID: 21324191; PubMed Central PMCID: PMC3053259.
- 777 35. Sirand-Pugnet P, Lartigue C, Marena M, Jacob D, Barre A, Barbe V, et al. Being
778 pathogenic, plastic, and sexual while living with a nearly minimal bacterial genome. PLoS

- 779 Genet. 2007;3(5):e75. Epub 2007/05/22. doi: 10.1371/journal.pgen.0030075. PubMed PMID:
780 17511520; PubMed Central PMCID: PMCPMC1868952.
- 781 36. Sirand-Pugnet P, Citti C, Barre A, Blanchard A. Evolution of mollicutes: down a
782 bumpy road with twists and turns. Res Microbiol. 2007;158(10):754-66. Epub 2007/11/21.
783 doi: 10.1016/j.resmic.2007.09.007. PubMed PMID: 18023150.
- 784 37. Sevillano L, Diaz M, Yamaguchi Y, Inouye M, Santamaria RI. Identification of the
785 first functional toxin-antitoxin system in Streptomyces. PLoS One. 2012;7(3):e32977. doi:
786 10.1371/journal.pone.0032977. PubMed PMID: 22431991; PubMed Central PMCID:
787 PMCPMC3303803.
- 788 38. Huisman O, D'Ari R. An inducible DNA replication-cell division coupling
789 mechanism in E. coli. Nature. 1981;290(5809):797-9. doi: 10.1038/290797a0. PubMed
790 PMID: 7012641.
- 791 39. Justice SS, Hunstad DA, Cegelski L, Hultgren SJ. Morphological plasticity as a
792 bacterial survival strategy. Nat Rev Microbiol. 2008;6(2):162-8. doi: 10.1038/nrmicro1820.
793 PubMed PMID: WOS:000252426700015.
- 794 40. Harms A, Stanger FV, Scheu PD, de Jong IG, Goepfert A, Glatter T, et al.
795 Adenylation of Gyrase and Topo IV by FicT Toxins Disrupts Bacterial DNA Topology.
796 Cell Rep. 2015;12(9):1497-507. Epub 2015/08/25. doi: 10.1016/j.celrep.2015.07.056.
797 PubMed PMID: 26299961.
- 798 41. Lartigue C, Vashee S, Algire MA, Chuang RY, Benders GA, Ma L, et al. Creating
799 bacterial strains from genomes that have been cloned and engineered in yeast. Science.
800 2009;325(5948):1693-6. Epub 2009/08/22. doi: 10.1126/science.1173759. PubMed PMID:
801 19696314.

- 802 42. Lartigue C, Valverde Timana Y, Labroussaa F, Schieck E, Liljander A, Sacchini F, et
803 al. Attenuation of a Pathogenic *Mycoplasma* Strain by Modification of the *obg* Gene by
804 Using Synthetic Biology Approaches. *mSphere*. 2019;4(3). doi: 10.1128/mSphere.00030-19.
805 PubMed PMID: 31118296.
- 806 43. Howe KL, Contreras-Moreira B, De Silva N, Maslen G, Akanni W, Allen J, et al.
807 Ensembl Genomes 2020-enabling non-vertebrate genomic research. *Nucleic acids research*.
808 2020;48(D1):D689-D95. Epub 2019/10/11. doi: 10.1093/nar/gkz890. PubMed PMID:
809 31598706; PubMed Central PMCID: PMC6943047.
- 810 44. Salamov VSA, Solovyevand A. Automatic annotation of microbial genomes and
811 metagenomic sequences. *Metagenomics and its applications in agriculture* Nova Science
812 Publishers, Hauppauge. 2011:61-78.
- 813 45. Lartigue C, Blanchard A, Renaudin J, Thiaucourt F, Sirand-Pugnet P. Host specificity
814 of mollicutes oriC plasmids: functional analysis of replication origin. *Nucleic acids research*.
815 2003;31(22):6610-8. Epub 2003/11/07. doi: 10.1093/nar/gkg848. PubMed PMID: 14602921;
816 PubMed Central PMCID: PMC6943047.
- 817 46. Braga-Lagache S, Buchs N, Iacovache MI, Zuber B, Jackson CB, Heller M. Robust
818 Label-free, Quantitative Profiling of Circulating Plasma Microparticle (MP) Associated
819 Proteins. *Mol Cell Proteomics*. 2016;15(12):3640-52. Epub 2016/10/16. doi:
820 10.1074/mcp.M116.060491. PubMed PMID: 27738094; PubMed Central PMCID:
821 PMC6943047.
- 822 47. Deutsch EW, Mendoza L, Shteynberg D, Farrah T, Lam H, Tasman N, et al. A guided
823 tour of the Trans-Proteomic Pipeline. *Proteomics*. 2010;10(6):1150-9. Epub 2010/01/27. doi:

- 824 10.1002/pmic.200900375. PubMed PMID: 20101611; PubMed Central PMCID:
825 PMCPMC3017125.
- 826 48. Eng JK, Hoopmann MR, Jahan TA, Egertson JD, Noble WS, MacCoss MJ. A deeper
827 look into Comet--implementation and features. *J Am Soc Mass Spectrom.* 2015;26(11):1865-
828 74. Epub 2015/06/28. doi: 10.1007/s13361-015-1179-x. PubMed PMID: 26115965; PubMed
829 Central PMCID: PMCPMC4607604.
- 830 49. Craig R, Beavis RC. A method for reducing the time required to match protein
831 sequences with tandem mass spectra. *Rapid Commun Mass Spectrom.* 2003;17(20):2310-6.
832 Epub 2003/10/15. doi: 10.1002/rcm.1198. PubMed PMID: 14558131.
- 833 50. Kim S, Pevzner PA. MS-GF+ makes progress towards a universal database search
834 tool for proteomics. *Nat Commun.* 2014;5:5277. Epub 2014/11/02. doi:
835 10.1038/ncomms6277. PubMed PMID: 25358478; PubMed Central PMCID:
836 PMCPMC5036525.
- 837 51. Tabb DL, Fernando CG, Chambers MC. MyriMatch: highly accurate tandem mass
838 spectral peptide identification by multivariate hypergeometric analysis. *J Proteome Res.*
839 2007;6(2):654-61. Epub 2007/02/03. doi: 10.1021/pr0604054. PubMed PMID: 17269722;
840 PubMed Central PMCID: PMCPMC2525619.
- 841 52. Choi H, Ghosh D, Nesvizhskii AI. Statistical validation of peptide identifications in
842 large-scale proteomics using the target-decoy database search strategy and flexible mixture
843 modeling. *J Proteome Res.* 2008;7(1):286-92. Epub 2007/12/15. doi: 10.1021/pr7006818.
844 PubMed PMID: 18078310.
- 845 53. Shteynberg D, Deutsch EW, Lam H, Eng JK, Sun Z, Tasman N, et al. iProphet: multi-
846 level integrative analysis of shotgun proteomic data improves peptide and protein

- 847 identification rates and error estimates. *Mol Cell Proteomics*. 2011;10(12):M111 007690.
848 Epub 2011/08/31. doi: 10.1074/mcp.M111.007690. PubMed PMID: 21876204; PubMed
849 Central PMCID: PMCPMC3237071.
- 850 54. Zybailov BL, Florens L, Washburn MP. Quantitative shotgun proteomics using a
851 protease with broad specificity and normalized spectral abundance factors. *Mol Biosyst*.
852 2007;3(5):354-60. Epub 2007/04/27. doi: 10.1039/b701483j. PubMed PMID: 17460794.
- 853 55. Zhang Y, Wen Z, Washburn MP, Florens L. Refinements to label free proteome
854 quantitation: how to deal with peptides shared by multiple proteins. *Anal Chem*.
855 2010;82(6):2272-81. Epub 2010/02/20. doi: 10.1021/ac9023999. PubMed PMID: 20166708.
- 856 56. Green MR, Sambrook J. *Molecular cloning : a laboratory manual*. 4th ed. Cold Spring
857 Harbor, N.Y.: Cold Spring Harbor Laboratory Press; 2012.
- 858 57. Schwendener S, Perreten V. The integrase of the *Macrococcus caseolyticus* resistance
859 island *mecD* (McRI_{mecD}) inserts DNA site-specifically into *Staphylococcus* and *Bacillus*
860 chromosomes. *Mol Microbiol*. 2018;110(3):455-68. doi: 10.1111/mmi.14112. PubMed
861 PMID: 30152907.
- 862 58. Dybvig K, Voelker LL. Molecular biology of mycoplasmas. *Annu Rev Microbiol*.
863 1996;50:25-57. doi: 10.1146/annurev.micro.50.1.25. PubMed PMID: 8905075.
- 864 59. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads
865 using repeat graphs. *Nat Biotechnol*. 2019;37(5):540-6. Epub 2019/04/03. doi:
866 10.1038/s41587-019-0072-8. PubMed PMID: 30936562.
- 867 60. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*.
868 2014;30(14):2068-9. doi: 10.1093/bioinformatics/btu153. PubMed PMID: 24642063.

- 869 61. Darling AC, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved
870 genomic sequence with rearrangements. *Genome research*. 2004;14(7):1394-403. doi:
871 10.1101/gr.2289704. PubMed PMID: 15231754; PubMed Central PMCID: PMC442156.
- 872 62. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular Evolutionary
873 Genetics Analysis across Computing Platforms. *Mol Biol Evol*. 2018;35(6):1547-9. doi:
874 10.1093/molbev/msy096. PubMed PMID: 29722887; PubMed Central PMCID:
875 PMC5967553.
- 876 63. Bertels F, Silander OK, Pachkov M, Rainey PB, van Nimwegen E. Automated
877 reconstruction of whole-genome phylogenies from short-sequence reads. *Mol Biol Evol*.
878 2014;31(5):1077-88. doi: 10.1093/molbev/msu088. PubMed PMID: 24600054; PubMed
879 Central PMCID: PMC3995342.
- 880 64. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search
881 tool. *J Mol Biol*. 1990;215(3):403-10. doi: 10.1016/S0022-2836(05)80360-2. PubMed PMID:
882 2231712.
- 883

884 **Tables**

885 **Table 1:** Candidate toxins and antitoxins of *M. mycoides* subsp. *capri* GM12 investigated in
 886 this study. The top five candidate TAS were identified *in silico* by using the TASmania
 887 database.

Name	Mnemonic/ designation in manuscript	Size [nt/aa]	Gene annotation	Pfam/ predicted TAS type	E-value in TASmania
TAS _{160/1}	MMCAP2_0160 / T ₁₆₀	753 /250	Abortive infection protein	AbiEii / IV	1.2e-53
	MMCAP2_0161 / A ₁₆₁	597 /198	Abortive infection protein	AbiEi / IV	3.8e-49
TAS _{889/0}	MMCAP2_0889 / T ₈₈₉	835 /278	Uncharacterized protein	AbiEii / IV	Guilty by association
	MMCAP2_0890 / A ₈₉₀	600 /199	Conserved hypothetical protein	AbiEi / IV	2.6e-45
TAS _{731/0}	MMCAP2_0731 / T ₇₃₁	1014 /337	Conserved hypothetical protein	AbiEii / IV	3.8e-29
	MMCAP2_0730 / A ₇₃₀	600 /199	Conserved hypothetical protein	AbiEi / IV	Guilty by association
TAS _{525/4}	MMCAP2_0525 / T ₅₂₅	402 /133	Hypothetical protein	MraZ / II	7e-26
	MMCAP2_0524 / A ₅₂₄	927 /308	S-adenosyl- methyltransferase	MraW / II	Guilty by association
TAS _{752/3}	MMCAP2_0752 / T ₇₅₂	573 /190	Cell filamentation protein	Fic / II	Guilty by association
	MMCAP2_0753 / A ₇₅₃	273 /90	DNA-damage- inducible protein J	RelB / II	3.5e-9
TAS _{133/2}	MMCAP2_0133 / T ₁₃₃	2274 /757	Subtilisin-like serine protease	- / II	N/A
	MMCAP2_0132 / A ₁₃₂	1062 /353	AAA-ATPase	- / II	N/A

888 TAS: toxin-antitoxin system, T: toxin, A: antitoxin

889

890 **Legends to Figures**

891

892 **Figure 1: Cartoon displaying the genomic localization and operon structure of candidate** 893 **TAS of *M. mycoides* subsp. *capri* GM12**

894 The circle displays the genome of GM12 (GenBank accession number: NZ_CP001668.1)
895 including its GC content. Six candidate TAS operons are shown in boxes in relation to their
896 localization in the genome. The promotor region, antitoxin and toxin are marked white, grey
897 and black, respectively. Candidate TAS selected for functionality testing are highlighted with
898 a red box.

899

900 **Figure 2: Detection of proteomic signatures of TAS in *in vitro* grown GM12**

901 The histogram shows relative protein abundance in strain GM12 [frequency over distributed
902 normalized spectral abundance factor (dNSAF)], with candidate toxins (T) and antitoxins (A)
903 marked as red arrows and the respective number of the toxin/antitoxin. Toxins and antitoxins
904 had the following dNSAF values: A₅₂₄ -3.47, T₁₃₃ -3.92, A₁₃₂ -3.92, A₇₅₃ -3.92, A₁₆₁ -4.45, T₇₃₁
905 -4.52. Threshold values of "low, medium and high" expression levels are based on data
906 distribution. Candidate toxins and antitoxins not displayed were below detection limits or
907 absent.

908

909 **Figure 3: Presence of six GM12 candidate TAS identified *in silico* in other mycoplasmas**

910 Dots on the right indicate the presence of genes encoding candidate toxins or antitoxins. A)
911 Phylogenetic tree based on whole genome sequences of different *M. mycoides* subsp. *capri*
912 (*Mmc*) strains, *M. mycoides* subsp. *mycoides* (*Mmm*) strain PG1^T was used as outgroup. The
913 tree was constructed using Neighbor-joining inference method, evolutionary distance was
914 calculated using Maximum Composite Likelihood method with rate variation among site

915 modelled with a gamma distribution. Bootstrap values (100 replicates) are displayed next to
916 tree branches. B) Phylogenetic tree of different members of the '*M. mycoides* cluster' based on
917 16S rRNA sequences, *M. putrefaciens* strain KS1^T was used as outgroup. The tree was
918 constructed using Neighbor-joining inference method. Bootstrap values (100 replicates) are
919 displayed next to tree branches.

920

921 **Figure 4: Toxicity of recombinant T₁₃₃ on *E. coli* LMG194**

922 A) Growth curves of *E. coli* LMG194 in response to induction or repression of heterologous
923 expression of T₁₃₃ and empty vector pBAD/His. Each data point represents the mean of three
924 biological replicates, bars indicate standard deviation. The p-values are displayed (* $p \leq 0.05$,
925 ** $p \leq 0.01$). B) Scanning electron micrograph (magnification 10,000x) displaying
926 morphological changes of *E. coli* LMG194 after heterologous expression of T₁₃₃. Blebbing at
927 the pole is indicated by an arrow, z-shaped cells by asterisk and cell debris by arrowheads.

928

929 **Figure 5: Toxicity of recombinant T₇₅₂ on *B. subtilis* 168 Marburg**

930 A) Growth curves of *B. subtilis* 168 Marburg in response to induction of heterologous
931 expression of T₇₅₂ and strain 168 Marburg. Each data point represents the mean of three
932 biological replicates, bars indicate standard deviation. The p-values are displayed (* $p \leq 0.05$,
933 ** $p \leq 0.01$ and *** $p \leq 0.001$). B) Scanning electron micrograph (magnification 10,000x)
934 displaying morphological changes of *B. subtilis* 168 Marburg after induction of heterologous
935 expression. Indentations of cells are indicated by asterisk, empty/ shrunk cells by arrowhead
936 and shrunk cells upon division by arrows.

937

938 **Figure 6: Effect of cloned toxins, antitoxins and TAS on the transformation rate into *M.***
939 ***capricolum* subsp. *capricolum***

940 *M. capricolum* subsp. *capricolum* ATCC 27343^T was transformed with different plasmid
941 constructs harboring either entire TAS operons (TAS_{133/2}, TAS_{752/3} and TAS_{160/1}), individual
942 toxins or antitoxins. Empty pMYCO1 plasmid was used as a positive control. Each column
943 represents the mean of three independent biological replicates and bars indicate standard
944 deviations. Significance is indicated (** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$).

945 **Supporting information**

946 **Figure S1: Detection of transcripts of candidate TAS**

947 PCR using cDNA as template was used to detect transcripts. Amplicons have been separated
948 using a 1% agarose gel. Gels on the left show amplicons obtained using gDNA as template.
949 The gels on the right show amplicons obtained using cDNA as template and therefore presence
950 of transcripts. Complementary DNA (cDNA) was synthesized by RT-PCR using RNA of
951 GM12 as template. M: 1 kbp GeneRuler DNA ladder.

952

953 **Figure S2: MAUVE alignment of 16 *Mmc* genomes**

954 Multiple sequence alignment of different *M. mycoides* subsp. *capri* genomes using Progressive
955 MAUVE and default parameters. Colored blocks are collinear and homologous regions.
956 Inversions of collinear blocks are displayed below the centre line of the genome. Strain names
957 are marked at the left of each alignment block and positions on the genomes are marked on top.

958

959 **Figure S3: Heterologous expression of different candidate toxins and antitoxins in *E. coli***

960 A) Immunoblot analysis of *E. coli* expressing heterologous candidate toxins and antitoxins
961 cloned into the pBAD/His expression vector. Expression was induced by addition of arabinose.
962 Expression was monitored over 7 h. 100 µg total protein was separated onto a 12% SDS PAGE
963 before being transferred to a nitrocellulose membrane. PageRuler™ was used as size marker.

964 B) Growth curves of *E. coli* harboring candidate toxins T₇₅₂, T₁₆₀ or antitoxins A₇₅₃, A₁₆₁ or
965 A₁₃₂ upon induction (solid line) or repression of expression (dashed line). C) Scanning electron
966 micrographs (magnification 10,000x) displaying *E. coli* harboring pBAD/His constructs with
967 toxin T₁₃₃, antitoxin A₁₃₂ or empty vector pBAD/His either induced with arabinose or repressed
968 with glucose.

969

970 **Figure S4: Heterologous expression of different candidate toxins and antitoxins in *B.***

971 *subtilis*

972 A) Growth curves of *B. subtilis* harboring candidate antitoxin A₁₆₁, toxin T₁₆₀ or 168 Marburg
973 with (solid line) and without (dashed line) induction of expression. B) Scanning electron
974 micrographs (magnification 10,000x) displaying *B. subtilis* harboring pHT01 constructs with
975 antitoxin A₁₆₁, toxin T₁₆₀, or empty vector pHT01 either induced with IPTG or uninduced.

976

977 **Table S1: TASmania database output file for *M. mycoides* subsp. *capri* GM12**

978

979 **Table S2: Output list of proteome analysis of *M. mycoides* subsp. *capri* GM12**

980

981 **Table S3: *M. mycoides* subsp. *capri* strains sequenced in this study**

982

983 **Table S4: TBLASTN results for candidate TAS presence in different *M. mycoides* subsp.
984 *capri* strains and different other mycoplasmas**

985

986 **Table S5: Oligonucleotide primers used in this study**

987

988 **File S1: Plasmids constructed in this study and codon-optimized toxin and antitoxin
989 encoding genes cloned for heterologous expression as well as TAS nucleotide sequences
990 cloned into pMYCO1.**

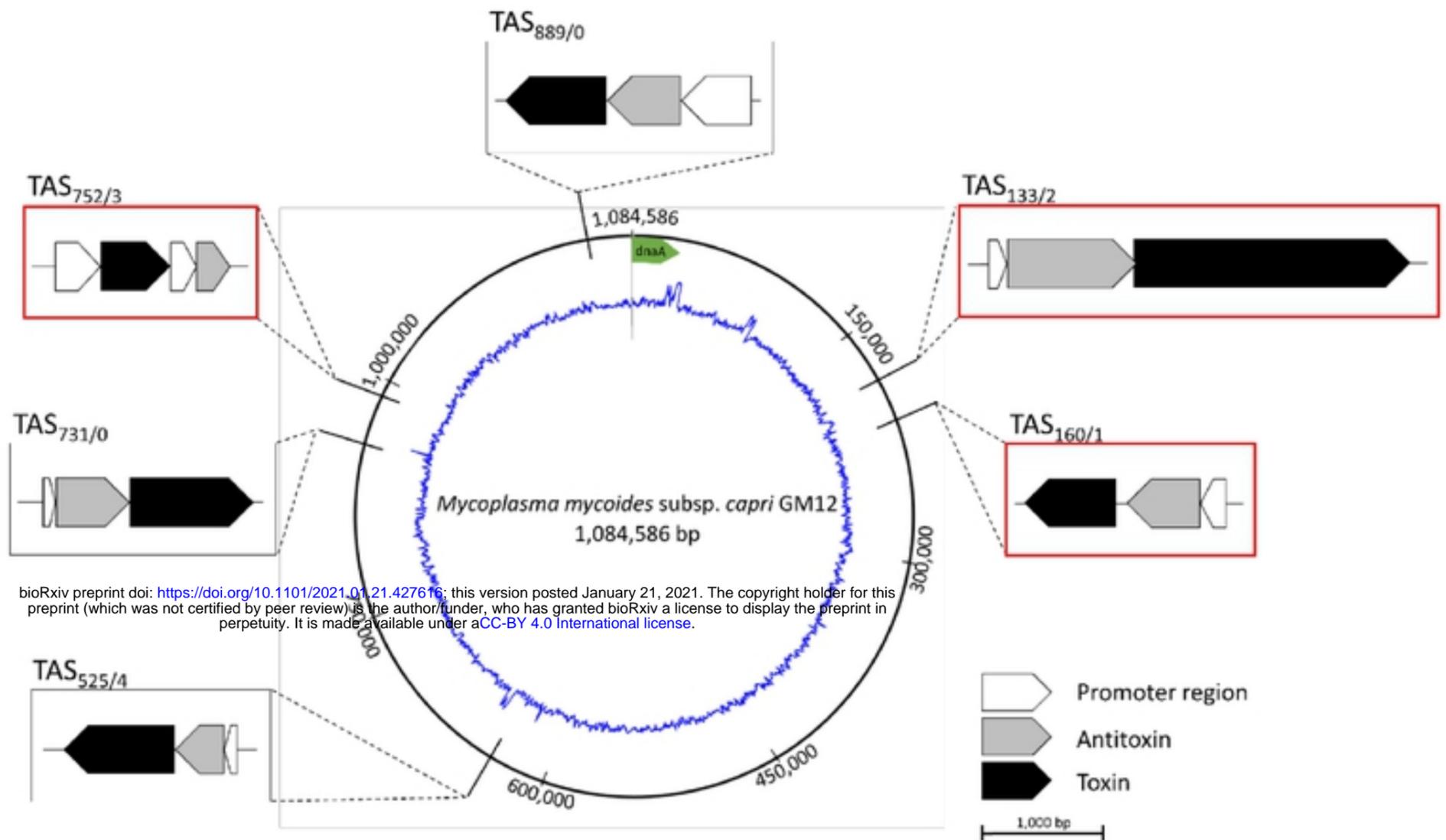


Figure 1: Cartoon displaying the genomic localization and operon structure of candidate TAS of *M. mycoides* subsp. *capri* GM12

The circle displays the genome of GM12 (GenBank accession number: NZ_CP001668.1) including its GC content. Six candidate TAS operons are shown in boxes in relation to their localization in the genome. The promoter region, antitoxin and toxin are marked white, grey and black, respectively. Candidate TAS selected for functionality testing are highlighted with a red box.

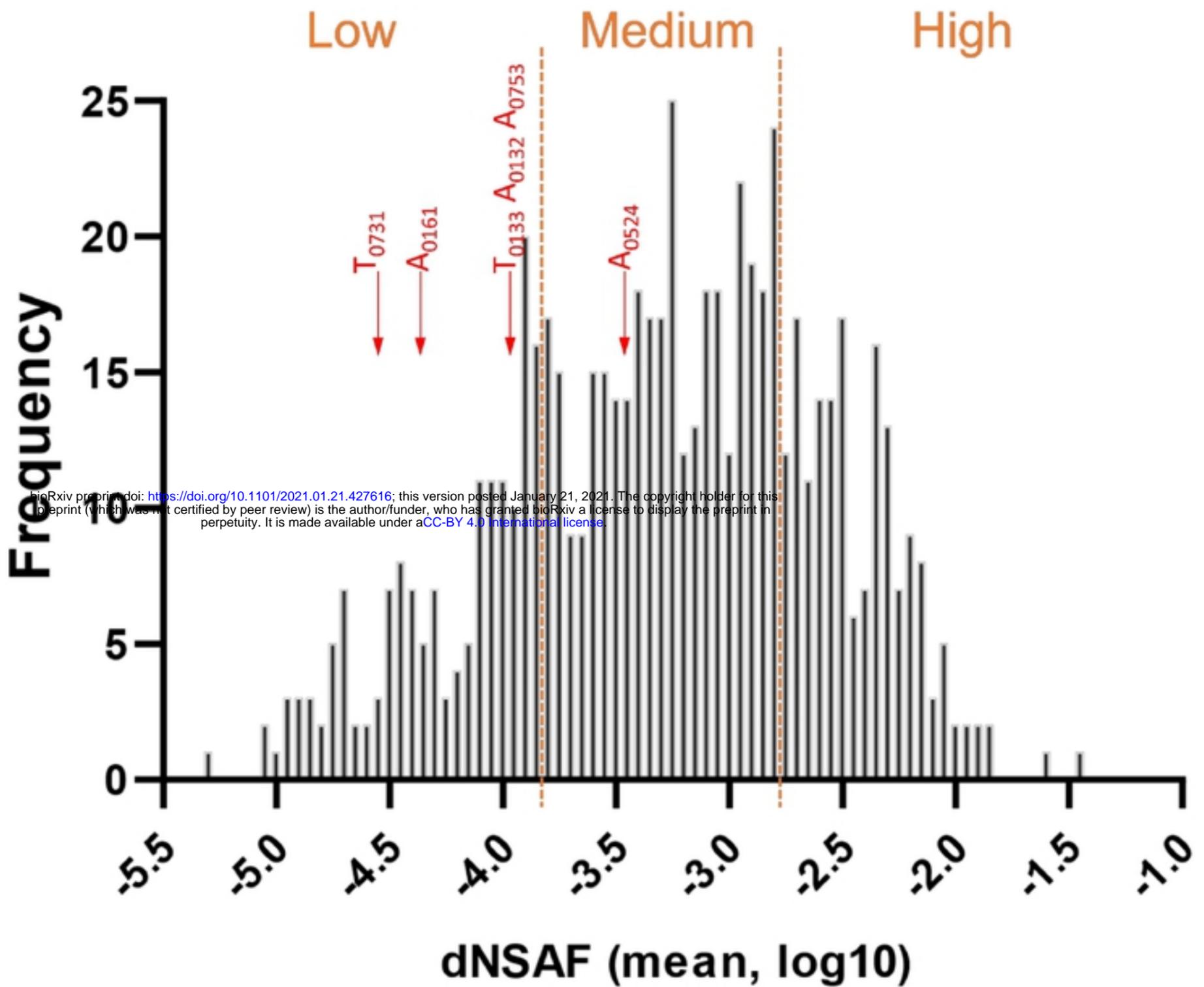


Figure 2: Detection of proteomic signatures of TAS in in vitro grown GM12

The histogram shows relative protein abundance in strain GM12 [frequency over distributed normalized spectral abundance factor (dNSAF)], with candidate toxins (T) and antitoxins (A) marked as red arrows and the respective number of the toxin/antitoxin. Toxins and antitoxins had the following dNSAF values: A₅₂₄ - 3.47, T₁₃₃ -3.92, A₁₃₂ -3.92, A₇₅₃ -3.92, A₁₆₁ -4.45, T₇₃₁ -4.52. Threshold values of "low, medium and high" expression levels are based on data distribution. Candidate toxins and antitoxins not displayed were below detection limits or absent.

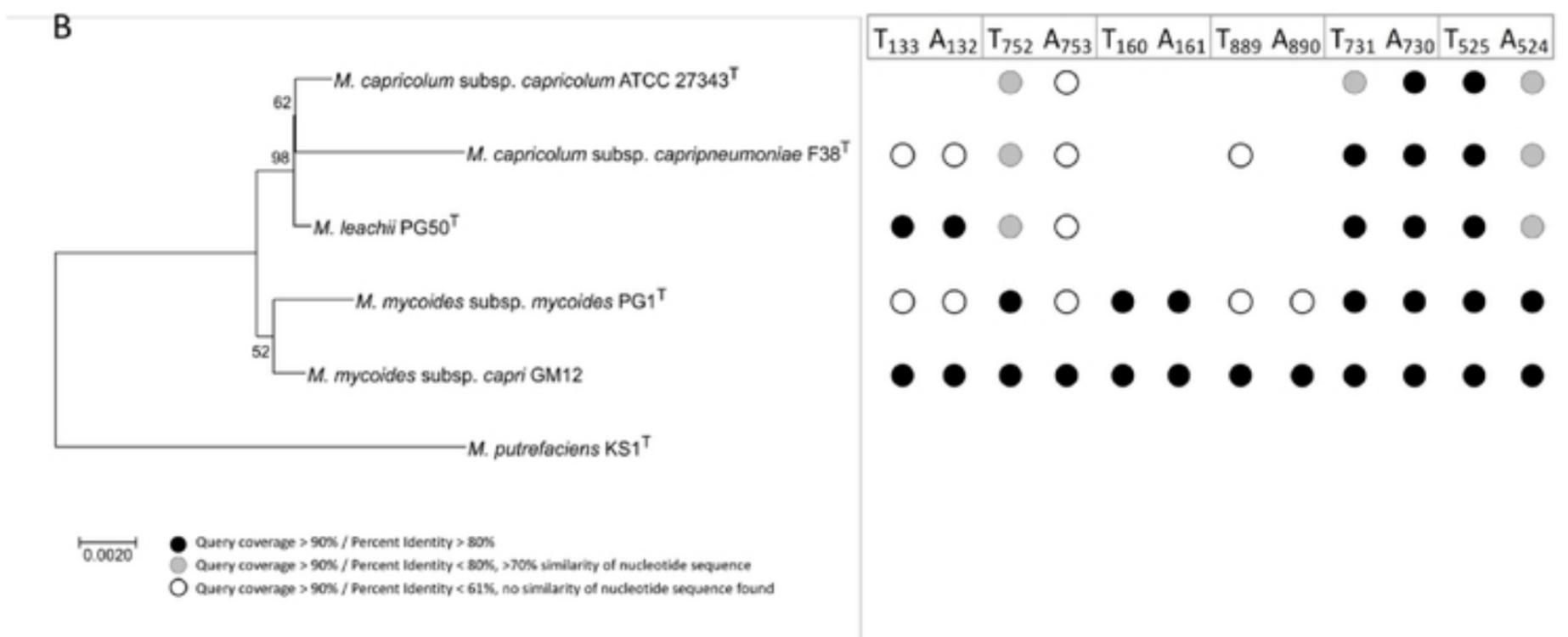
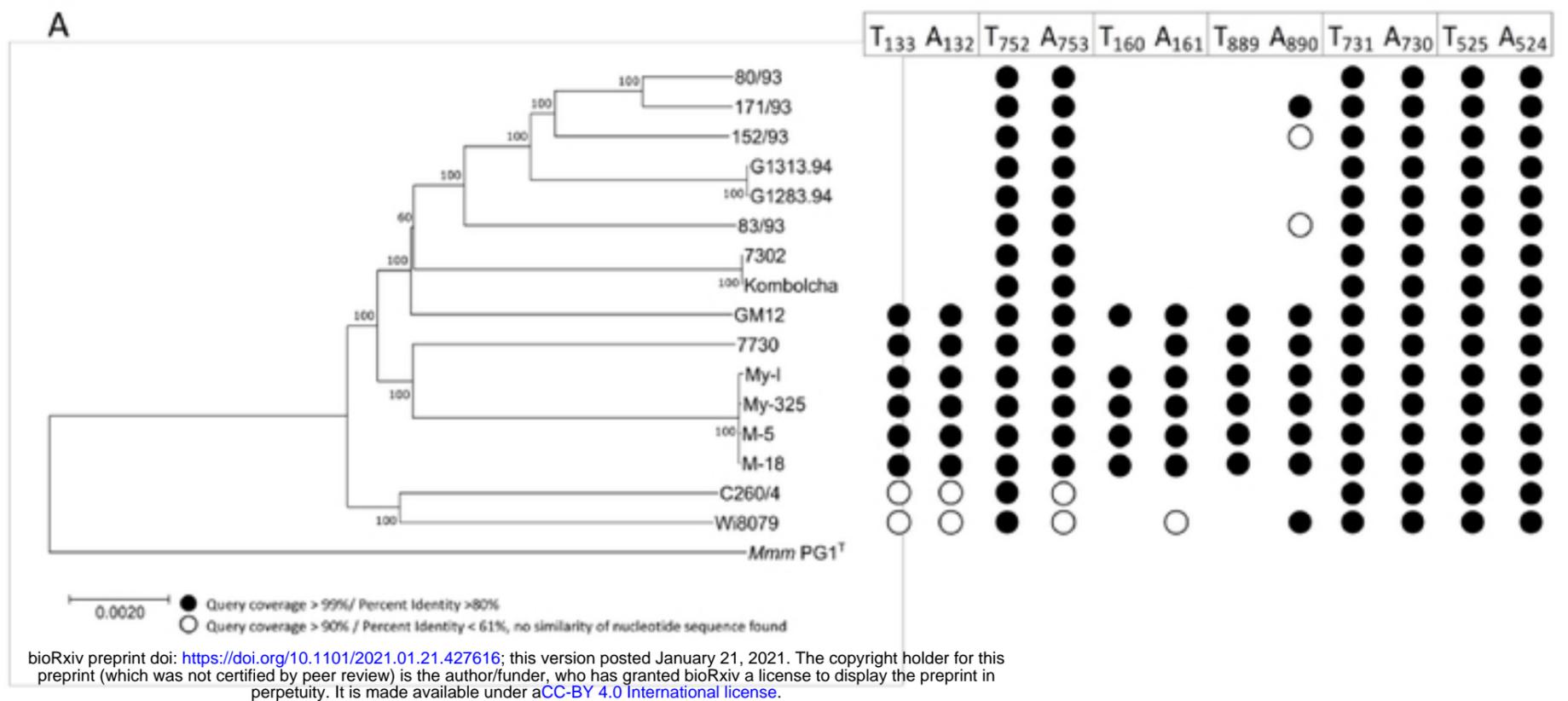


Figure 3: Presence of six GM12 candidate TAS identified in silico in other mycoplasmas

Dots on the right indicate the presence of genes encoding candidate toxins or antitoxins. A) Phylogenetic tree based on whole genome sequences of different *M. mycooides* subsp. *capri* (*Mmc*) strains, *M. mycooides* subsp. *mycooides* (*Mmm*) strain PG1^T was used as outgroup. The tree was constructed using Neighbor-joining inference method, evolutionary distance was calculated using Maximum Composite Likelihood method with rate variation among site modelled with a gamma distribution. Bootstrap values (100 replicates) are displayed next to tree branches. B) Phylogenetic tree based on 16S rRNA sequences of different members of the ‘*M. mycooides* cluster’, *M. putrefaciens* strain KS1 was used as outgroup. The tree was constructed using Neighbor-joining inference method. Bootstrap values (100 replicates) are displayed next to tree branches.

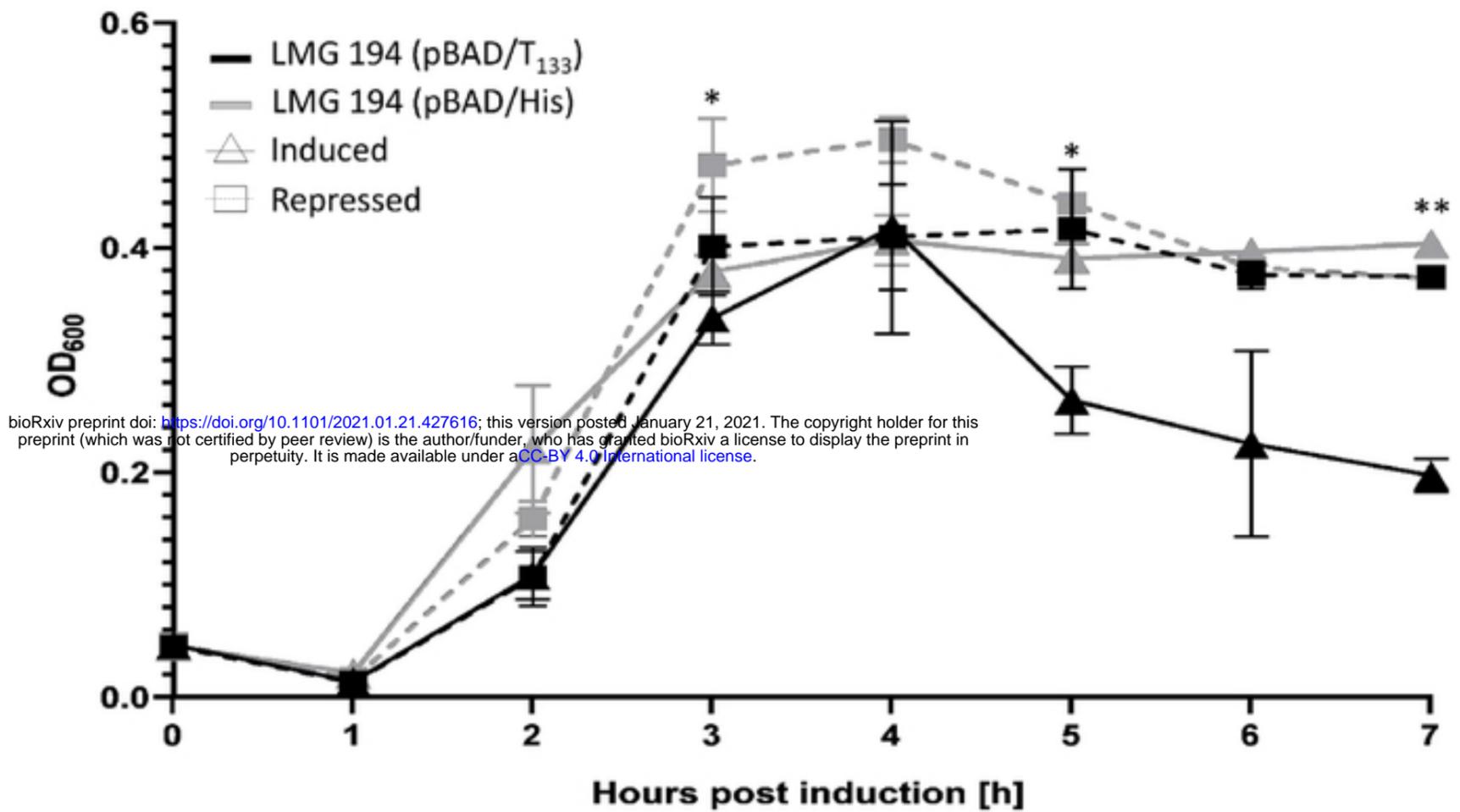
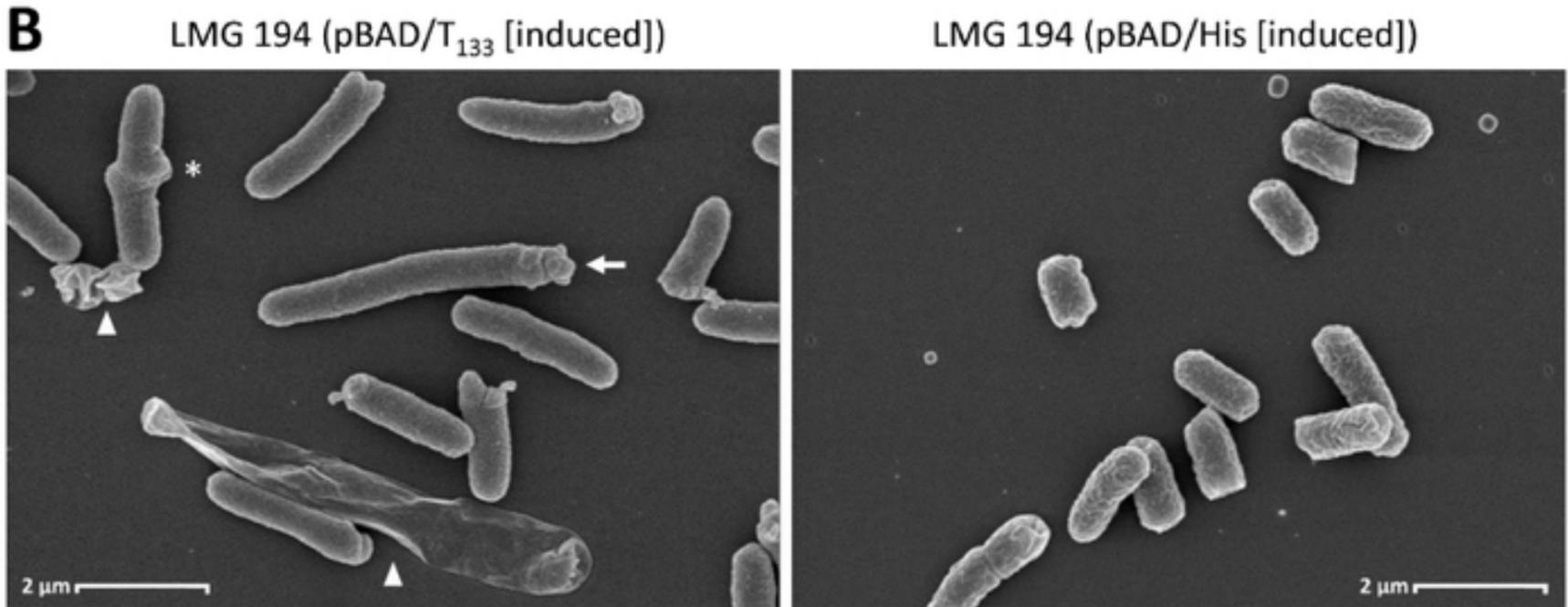
A**B**

Figure 4: Toxicity of recombinant T₁₃₃ on E. coli LMG 194

A) Growth curves of E. coli LMG 194 in response to induction or repression of heterologous expression of T₁₃₃ and empty vector pBAD/His. Each data point represents the mean of three biological replicates, bars indicate standard deviation. The p-values are displayed (* $p \leq 0.05$, ** $p \leq 0.01$). B) Scanning electron micrograph (magnification 10,000x) displaying morphological changes of E. coli LMG 194 after heterologous expression of T₁₃₃. Blebbing at the pole is indicated by an arrow, z-shaped cells by asterisk and cell debris by arrowheads.

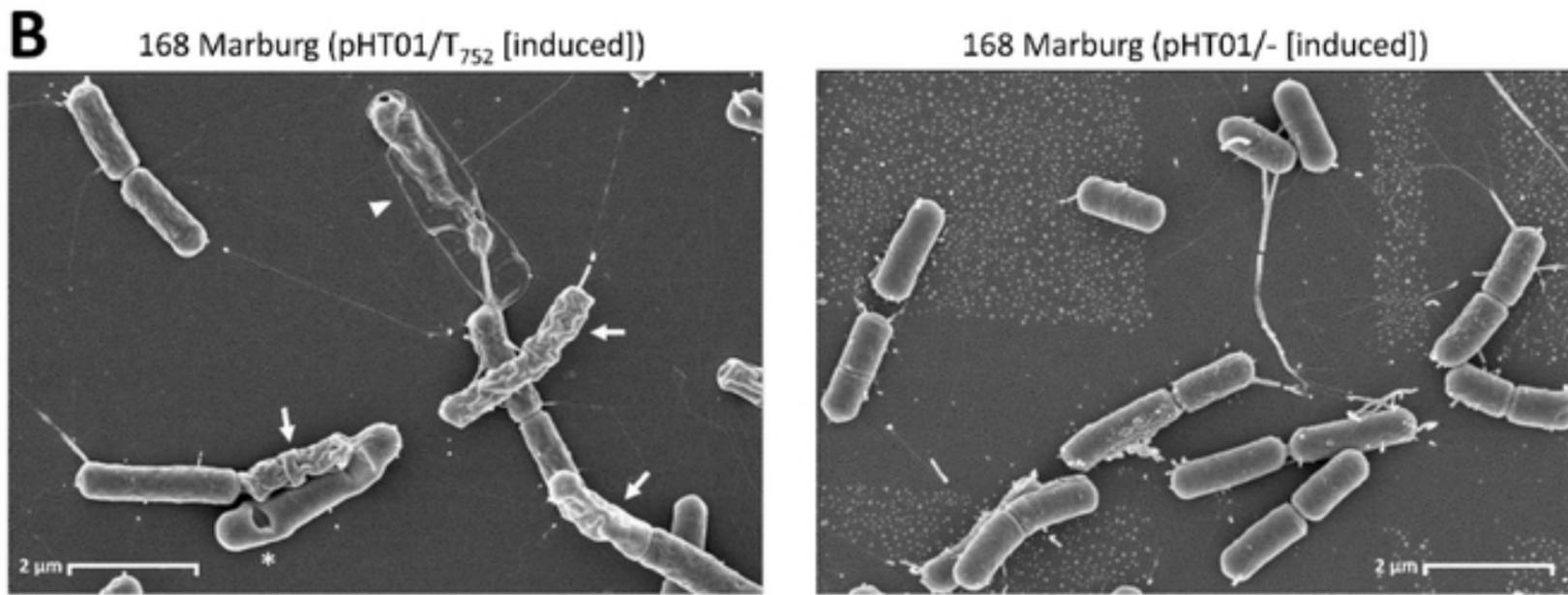
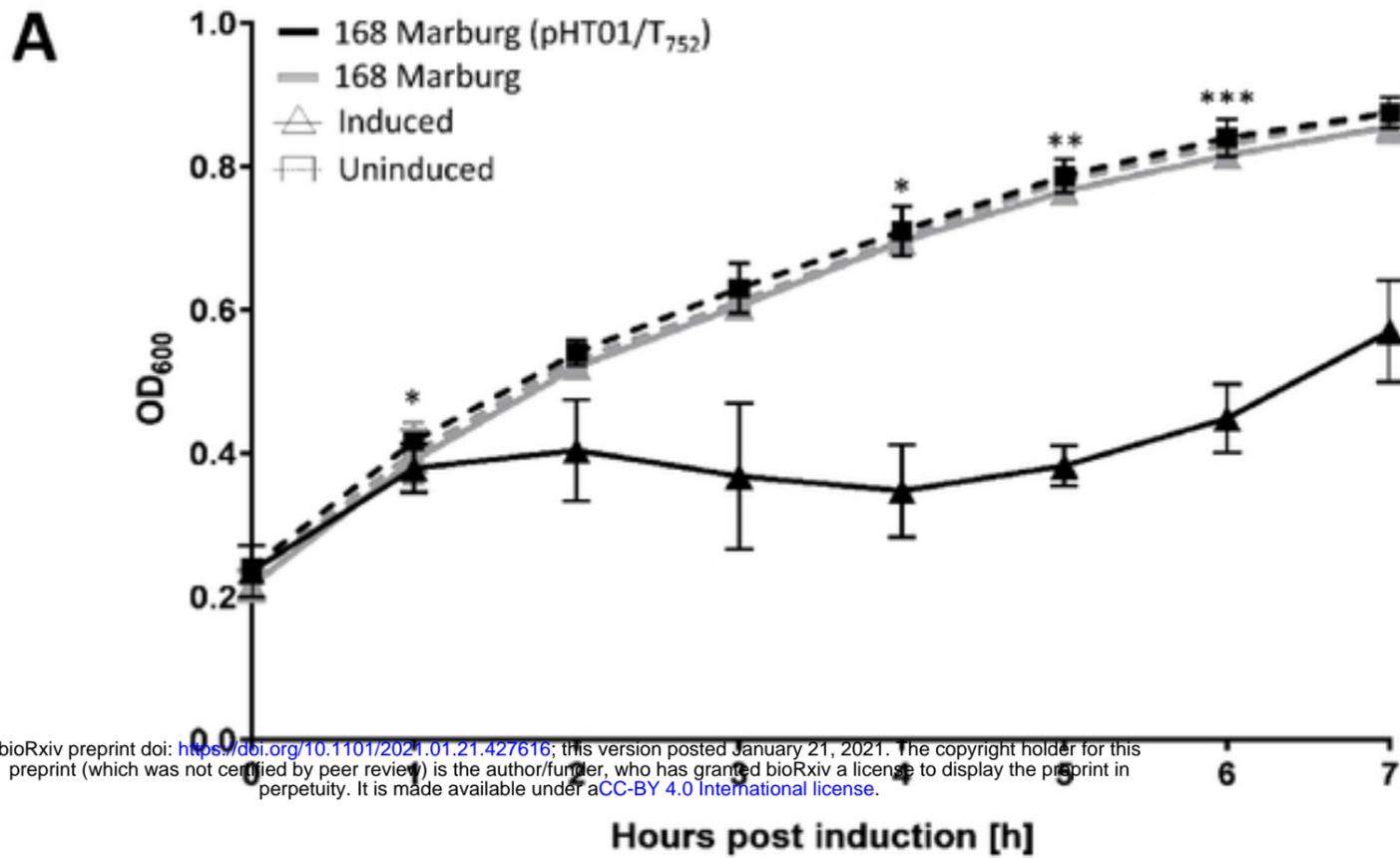


Figure 5: Toxicity of recombinant T₇₅₂ on *B. subtilis* 168 Marburg

A) Growth curves of *B. subtilis* 168 Marburg in response to induction of heterologous expression of T₇₅₂ and strain 168 Marburg. Each data point represents the mean of three biological replicates, bars indicate standard deviation. The p-values are displayed (* $p \leq 0.05$, ** $p \leq 0.01$ and *** $p \leq 0.001$). B) Scanning electron micrograph (magnification 10,000x) displaying morphological changes of *B. subtilis* 168 Marburg after induction of heterologous expression. Indentations of cells are indicated by asterisk, empty/ shrunk cells by arrowhead and shrunk cells upon division by arrows.

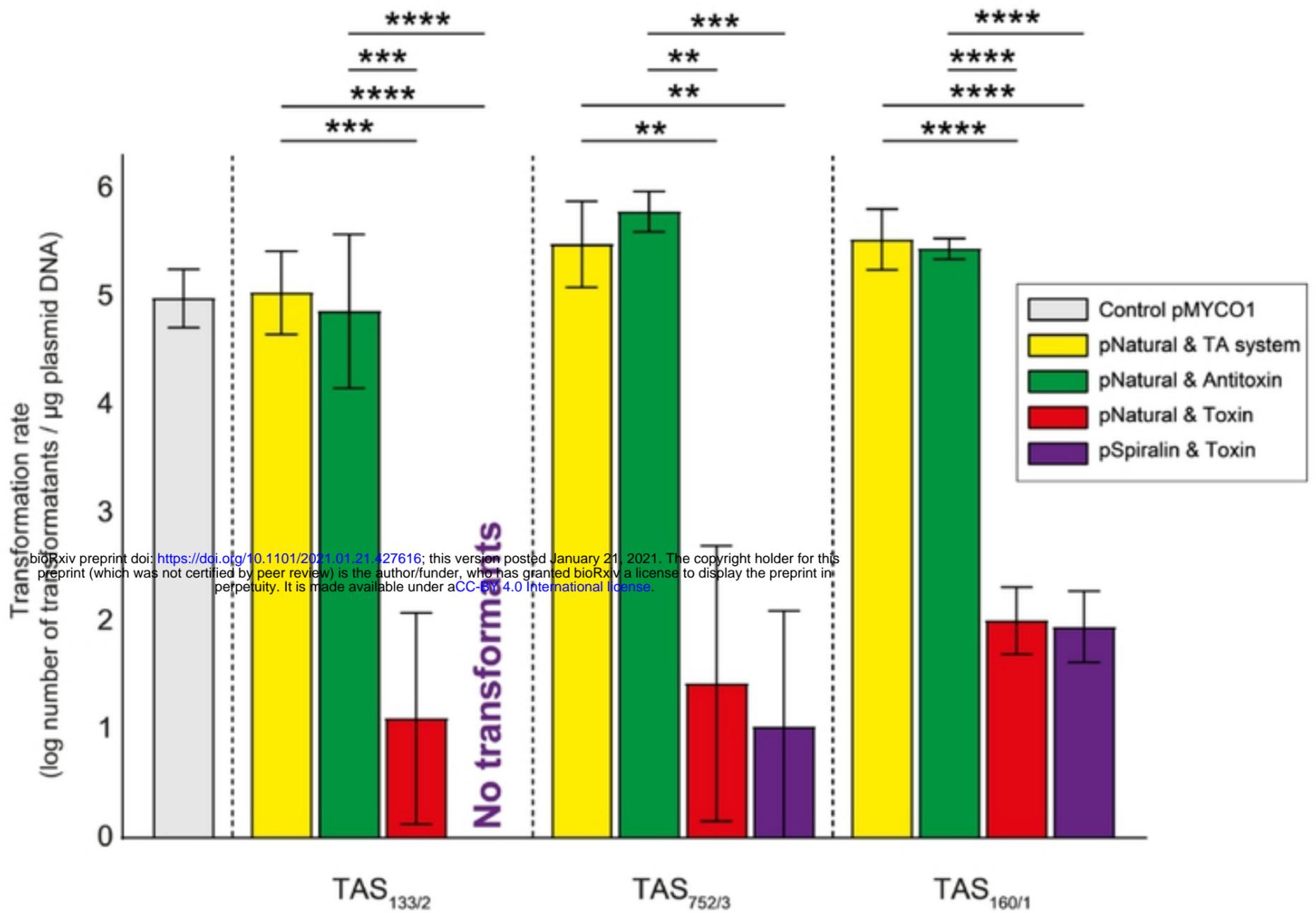


Figure 6: Effect of cloned toxins, antitoxins and TAS on the transformation rate into *M. capricolum* subsp. *capricolum*

M. capricolum subsp. *capricolum* ATCC 27343^T was transformed with different plasmid constructs harboring either entire TAS operons (TAS_{133/2}, TAS_{752/3} and TAS_{160/1}), individual toxins or antitoxins. Empty pMYCO1 plasmid was used as a positive control. Each column represents the mean of three independent biological replicates and bars indicate standard deviations. Significance is indicated (** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$).