Supplementary material

Impact of donor-recipient phylogenetic distance on bacterial genome transplantation

Fabien Labroussaa[1,2†], Anne Lebaudy[1,2†], Vincent Baby[3], Géraldine Gourgues[1,2], Dominick Matteau[3], Sanjay Vashee[4], Pascal Sirand-Pugnet[1,2], Sébastien Rodrigue[3] and Carole Lartigue[1,2§]

† these authors contributed equally to this work

[1] INRA, UMR 1332 de Biologie du Fruit et Pathologie, F-33140 Villenave d'Ornon, France

[2] Univ. Bordeaux, UMR 1332 de Biologie du Fruit et Pathologie, F-33140 Villenave d'Ornon, France

[3]Université de Sherbrooke. Département de biologie, 2500 boul. Université Sherbrooke (Québec), Canada, J1K 2R1

[4]J. Craig Venter Institute, Rockville, MD 20850 USA

§ to whom correspondence should be addressed at INRA, Centre de Recherche de Bordeaux, Institut de Biologie Végétale Moléculaire, 71, avenue Edouard Bourlaux, BP 81, 33883 Villenave d'Ornon Cedex, France.

Tel: +33 5 57 12 23 59; Fax: +33 5 57 12 23 69; clartigu@bordeaux.inra.fr

This file includes

MATERIAL AND METHODS

Construction of pMT85tetM-PSlacZ-PRS313 plasmid. Three overlapping DNA fragments were first produced by PCR amplification (Advantage HF 2 PCR Kit from Clontech), purified and then combined at 50°C using a homemade master mix made of three enzymes (T5 exonuclease, DNA polymerase and T4 DNA ligase) along with other buffer components (1). The first 3.57kb DNA fragment composed of the β-galactosidase gene under the control of the spiraline promotor (PSlacZ) was produced using the pMCO3LacZ plasmid as template (2) and the primers, pMT85tetM/PSlacZF1 (5´-GATCCACCCGCAATTACTGTGAGAGCAAGAATGGCTACTAAAGG) and pRS313/PSlacZR2 (5´-CTGGCTTCAGCAGAGCGCAGATACCGGCCAGTGAATTGTAATACGAC). The second DNA fragment corresponding to the whole transposon based-plasmid pMT85-tetM (4.73 kbp) (3–5) was PCR amplified using the primers, His-pMT85PStetMF2 (5´-TATACGTGTCATTCTGAACGAGGCGCAAGAAAACAGTGAAGCACCAG) andPSlacZ-pMT85PStetMR1 (5´-CCTTTAGTAGCCATTCTTGCTCTCACAGTAATTGCGGGTGGATC). The third DNA fragment (3.22 kbp) carrying the yeast elements (ARSH4, CEN6 and the histidine auxotrophic marker) was amplified using the primers LacZ/pRS313-F2 (5´-GTCGTATTACAATTCACTGGCCGGTATCTGCGCTCTGCTGAAGCCAG) and PS/pRS313-R2 (5´- CTGGTGCTTCACTGTTTTCTTGCGCCTCGTTCAGAATGACACGTATAG) and subsequent amplification using the plasmid pRS313 (ATCC 77142) as template. Cassettes were designed to contain overlaps between 44 to 48bp. Genbank accession number for the whole plasmid sequence is KX011460.

Quantitation of the number of *Mollicutes* genomes included into agarose plugs. Quantitative real-time polymerase chain reaction (qPCR) was used to determine bacterial

culture titers for all species studied before mixing bacterial cells into agarose plugs for genome isolation purposes. One ml of bacterial culture was centrifuged for 2 min at 14,000 x g at room temperature (RT). Cells were washed using 1 ml of phosphate buffer saline (137 mM NaCl; 2,7 mM KCl; 10 mM Na_2HPO_4; 1.76 mM KH_2PO_4; pH 7.4), centrifuged as above and then resuspended into 500 µl of cell lysis buffer (Wizard genomic extraction kit). Cells were incubated for 10 min at 95°C and the lysate was centrifuged for 2 min at 8,000 x g at RT. One µl of the undiluted supernatant was used as DNA template for subsequent qPCR measurements, as well as 5 additional dilutions of the initial supernatant, ranging from 10^{-1} to 10^{-5}. qPCR experiments were conducted into 96-well plates using the SsoFast™ Evagreen® Supermix (Bio-Rad) following the manufacturer's instructions. Fluorescence was recorded during every cycle on the LightCycler® 480 Real-Time PCR System (Roche) and a melt genotyping curve was performed to assess the specificity of the amplification recorded. Amplifications were performed on a 223 bp fragment of 16S ribosomal RNA conserved among *Mollicutes* using the oligonucleotides, qMyc16S-F (5'-TTAAACCACATGCTCCACCA-3') and qMyc16S-R (5'-AGCGGCTTACTGGCTTGTTA-3'). Standard curves used to quantify numbers of *Mollicutes* genomes were independently for each assay using purified genomic DNA (Wizard Genomic DNA Purification Kit, Promega) using samples containing 10^1 to 10^8 estimated genomes per well. Measurements were done in triplicates for all samples.

Transplant analyses. Transplants were analyzed by PCR, Pulsed Field Gel Electrophoresis (PFGE), and whole genome sequencing. **(i) PCR with species specific primers.** Genomic DNA from transplants was extracted as follows: 200µl of each culture (3rd passage) was centrifuged for 10 min at 14,000 x *g*. Pellets were resuspended into 100 µl of TE 1X buffer and then, each sample was incubated for 10 min at 95°C to lyse the cell membrane. After a

1/10th dilution, 1 µl of each sample was used as template for PCR reactions using species specific primers (Table S6). PCR reactions were performed with the *Taq* DNA polymerase (NEB) following the manufacturer's instructions. Results obtained for each species are summarized in Figure S3. **(ii) PFGE analysis.** Intact genomic DNA from transplants was isolated in agarose plugs as described in Material and Methods. Plugs were washed twice in 1 ml 1X wash buffer (CHEF Genomic DNA Plug Kit, Bio-Rad) supplemented with 1mM phenylmethylsulfonyl fluoride (PMSF) for 1h. The same washing steps were then conducted with 0.1X Wash buffer without the addition of PMSF. Plugs were subsequently equilibrated in the appropriate 1X restriction enzyme buffer for 1hr with gentle agitation. The previous step was repeated for 20 min with fresh buffer before the addition of 40 units of the restriction enzyme (ApaI or XhoI), followed by overnight incubation at the appropriate temperature (i.e. 37°C). Plugs were finally loaded into a 1% pulsed-field agarose gel (Bio-Rad) in 1X TAE and electrophoresed. Pulse times were ramped from 50 to 90 s for 22 hr at 6V.cm-1. After electrophoresis, gels were stained with SYBR Gold and cleaved DNA patterns were visualized using the E-BOX VX2 Gel Documentation System. Sizes of restriction fragments observed (Figure S3) were then compared with the expected fragment sizes calculated based on each donor genomes (Table S7).

RESULTS

Isolation of known quantities of marked *Mollicutes* genomes in agarose plugs. We developed a rapid quantitative real-time polymerase chain reaction (qPCR) assay allowing an enhanced control of donor gDNA quantities added to the reaction during the from-bacteria genome transplantation process. This qPCR assay, targeting the ribosomal 16S DNA conserved among all species tested, was conducted during the chloramphenicol treatment that occurs at the beginning of the gDNA plugs preparation (6). The duration of the chloramphenicol treatment, known to allow compaction of DNA at sub-optimal concentrations (7, 8), is species-dependent and takes its respective generation time into account (comprised between 1 hr and 2 hr, Material and Methods). Overall, gDNA was extracted independently on each bacterial culture and dilutions of each extraction were used for the qPCR reaction. Results presented here were obtained for *Mmc* GM12 (Figure S2) but identical results were obtained with all other species (data not shown). As expected, no bacterial growth was observed during the chloramphenicol treatment (no statistical difference recorded between the beginning (black bars, 0h) and the end (dotted bars) of the treatment) offering us a timeframe to accurately quantify bacterial concentration. To validate the qPCR assay, colony counting was used to estimate cultures titers (Figure S2, dashed line) and results were compared to qPCR quantification. We observed that undiluted samples, 10^{-1} or 10^{-2} dilutions gave inaccurate measurements resulting in an overall over-estimation of the real bacterial concentration. On the contrary, higher dilutions of the DNA extractions (10^{-3} to 10^{-5}) gave more accurate measurements of the bacterial population (Figure S2). After validation, this qPCR assay was systematically integrated to the gDNA preparation process allowing an early quantification of the amount of cells in cultures and a preparation of agarose plugs with known quantities of marked *Mollicutes* genomes.

FIGURES

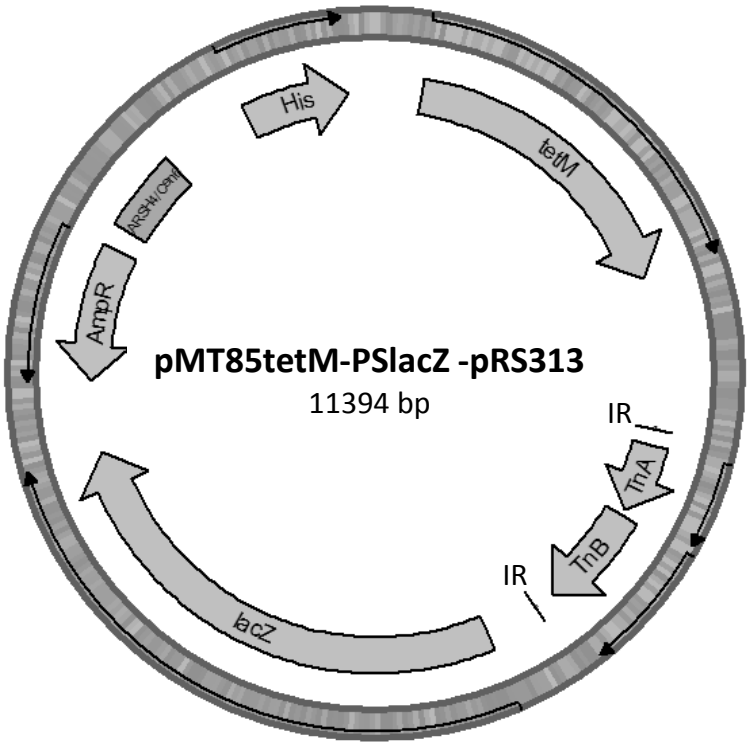

Figure S1. Map of pMT85tetM-PSlacZ-pRS313 plasmid.

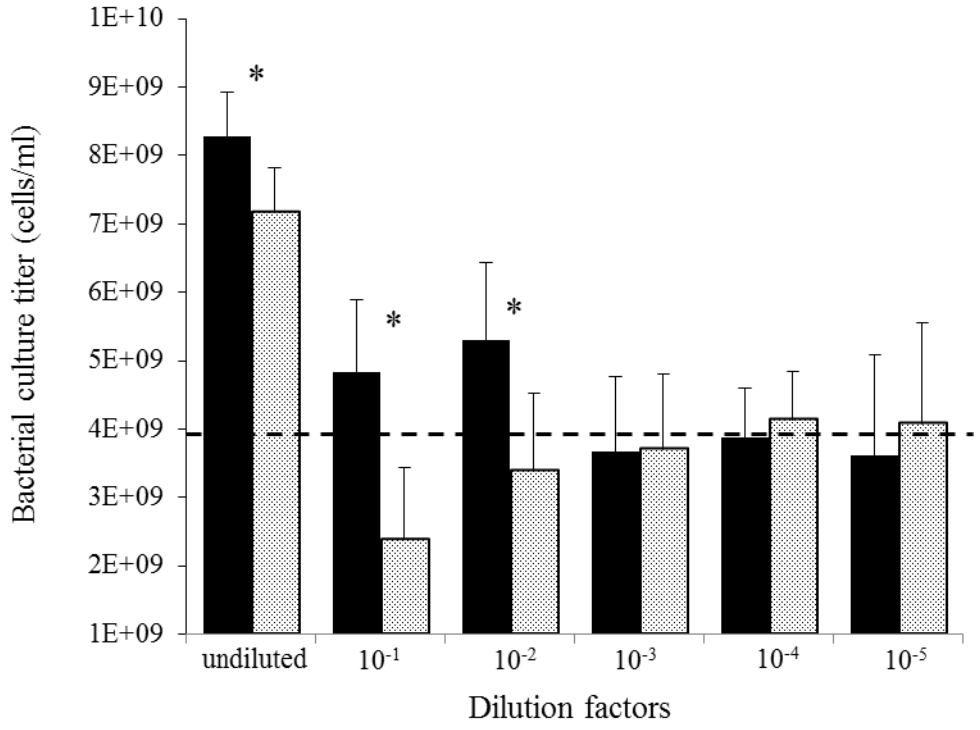

Figure S2. Titer estimations of *Mollicutes* cultures using quantitative real-time PCR prior embedding of donor genomes in plugs. As an example, gDNA preparations of a *Mmc* GM12 culture were carried out before (black bars) or after (dotted bars) a 2h chloramphenicol treatment. gDNA preparations were used undiluted or after 5 successive serial dilutions (10^{-1} to 10^{-5}) as DNA templates. The dashed line ($3.99E10^{9}$ cells/ml) represents the result of a comparative cell counting experiment done in parallel to show qPCR accuracy. Stars represent the statistical difference obtained ($p<0.001$, paired t-test) between the beginning and the end of the chloramphenicol treatment with low-dilution samples.

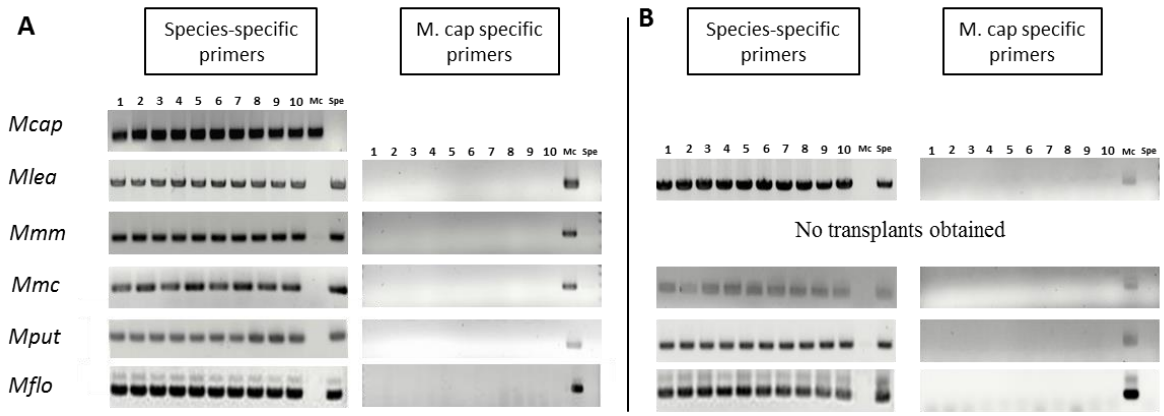

Figure S3. Genotype analysis of putative bacterial transplants obtained during (A) from-bacteria or (B) from-yeast GT assays using PCR. PCRs were conducted with two different sets of primers: (i) species-specific primers characteristic of the bacterial species used as donor genome and (ii) *Mcap*-specific primers used as negative controls. Lanes 1 to 10 correspond to 10 different transplants tested. Mc: *Mcap* used as positive (*Mcap*-specific primers) or negative (species-specific primers) controls. Spe: species tested among the 6 species for which putative transplants were obtained (*Mcap, Mlea, Mmm, Mmc, Mput and Mflo*). Amplification sizes are summarized in Table S6.

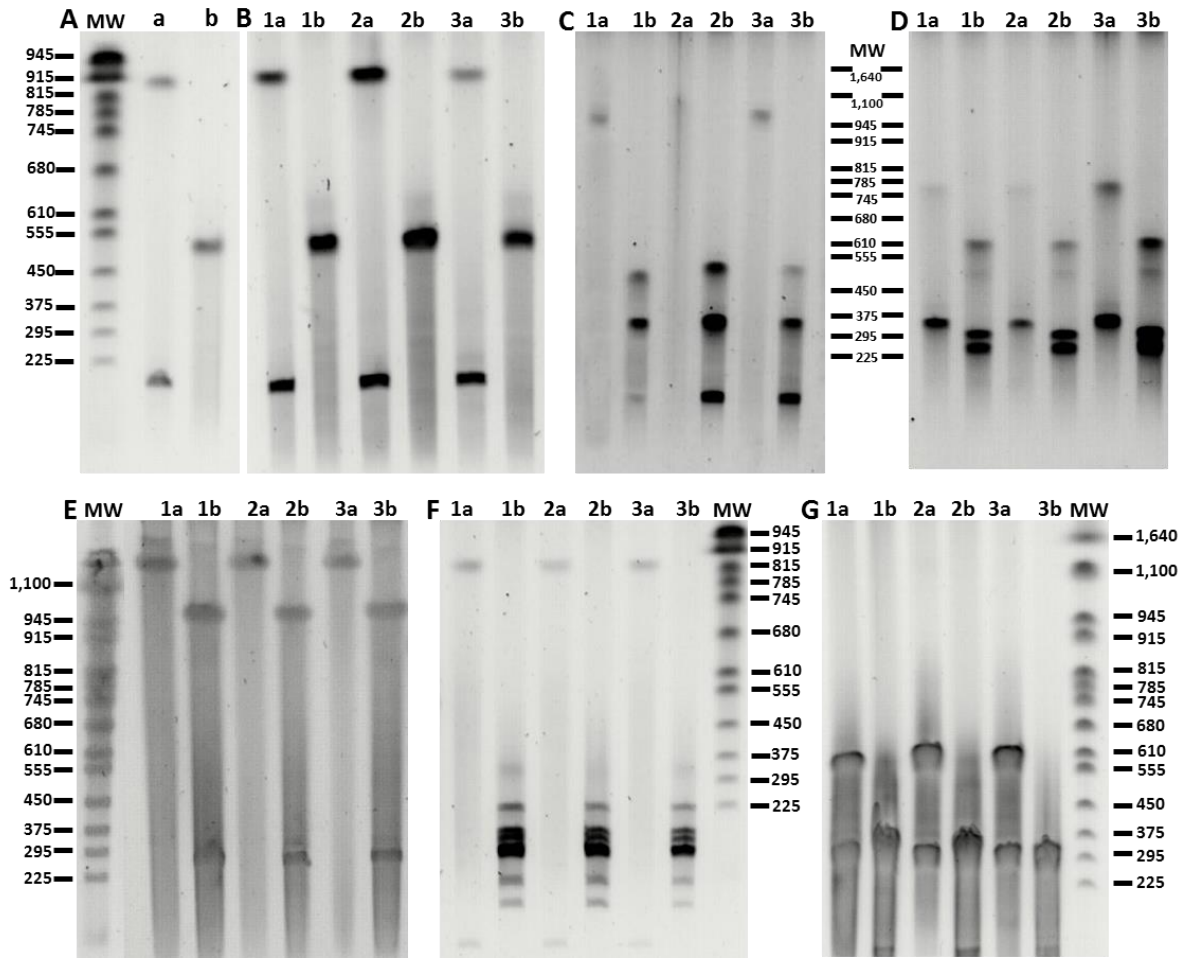

Figure S4. Confirmation of transplant genotypes using PFGE analyses. For each donor species tested in this study, PFGE was conducted on the donor genome (lane 1) and digestion profiles were compared to those of two putative transplants (lanes 2 and 3), previously tested by species-specific PCR. gDNA plugs for *Mcap*ΔRE (A, recipient cell), *Mcap* (B), *Mlea* (C), *Mmc* (D), *Mmm* (E) and *Mflorum* (F) were digested with ApaI (a) and XhoI (b) prior to electrophoresis. ApaI and XhoI-digested fragment sizes for each species are summarized in Table S7. MW: yeast chromosome PFG marker (New England Biolabs) in kpb.

Mycoplasma species	Strain	Natural Host	Genome size (Mbp)	% G+C	# CDS	GenBank #
M. leachii	PG50	Ruminant	1.009	23.7	883	CP002108.1
M. mycoides subsp. *mycoides*	PG1	Ruminant	1.212	24.0	917	BX293980.2
M. mycoides subsp. *capri*	95010	Ruminant	1.154	23.8	921	FQ377874.1
M. capricolum subsp. *capricolum*	ATCC27343	Ruminant	1.010	23.8	817	CP000123.1
M. putrefaciens	CIRAD	Ruminant	0.860	27.0	684	CP004357.1
Mesoplasma florum	L1	Plants	0.793	27.0	683	AE017263.1
Spiroplasma citri	GII-3	Insects&plants	1.840	26.0	2050	PRJNA13930
M. hominis	PG21	Human	0.665	27.1	531	FP236530.1
M. mobile	163K	Fish	0.777	25.0	637	AE017308.1
M. pulmonis	UAB CTIP	Rodent	0.964	26.6	783	AL445566.1
M. bovis	PG45	Ruminant	1.003	29.3	766	CP002188.1
M. agalactiae	PG2	Ruminant	0.877	29.7	714	CU179680.1
M. pneumoniae	M129	Human	0.816	40.0	689	U00089.2
M . genitalium	G37	Human	0.580	32.0	482	L43967.2
M. gallisepticum	R (Low)	Poultry	0.996	31.4	727	AE015450.2
Candidatus Phytoplasma mali	AT	Insect&plant	0.602	21.4	480	CU469464.1
Candidatus Phytoplasma australiense	rp-A	Insect&plant	0.880	27.4	844	AM422018.1
Acholeplasma laidlawii	PG-8A	Ubiquitous	1.497	31.9	1380	CP000896.1

Table S1. List of *Mollicutes* genomes used to reconstruct the phylogenetic tree and to generate the similarity percentages of the core proteome. Species tested for whole genome transplantation studies are indicated in grey.

Species	*Mcap*	*Mlea*	*Mmc*	*Mmm*	*Mputr*	*Mflorum*	*Scitri*	*Magalactiae*	*Mbovis*	*Mpulmonis*	*Mhominis*	*Mmobile*	*Mpneumoniae*	*Mgenitalium*	*Mgallisepticum*	*Phytomali*	*Phytoaustraliense*	*Alaidlawii*
Mcap	**100.00**	**99.68**	**99.25**	**99.06**	**95.43**	**92.47**	**85.29**	**76.60**	**76.46**	**77.68**	**77.69**	**77.84**	**74.41**	**74.37**	**76.66**	**75.29**	**75.68**	**77.88**
Mlea		100.00	99.23	99.13	95.40	92.49	85.30	76.68	76.53	77.71	77.64	77.77	74.32	74.35	76.71	75.36	75.67	77.91
Mmc			100.00	99.56	95.47	92.51	85.28	76.67	76.55	77.73	77.68	77.83	74.35	74.31	76.73	75.17	75.68	77.89
Mmm				100.00	95.15	92.33	85.15	76.62	76.46	77.63	77.54	77.71	74.24	74.20	76.71	75.07	75.47	77.75
Mputr					100.00	93.07	85.27	76.57	76.49	77.56	77.62	77.67	74.24	74.16	76.38	75.33	75.82	77.78
Mflorum						100.00	85.07	76.20	76.20	77.49	77.44	77.40	73.87	73.89	76.32	74.69	75.48	77.45
Scitri							100.00	76.67	76.65	78.05	78.03	77.92	74.26	74.01	76.84	76.17	76.53	78.71
Magalactiae								100.00	99.01	84.18	82.99	82.61	71.67	71.51	73.85	72.32	72.74	74.67
Mbovis									100.00	84.17	82.89	82.60	71.66	71.50	73.80	72.27	72.79	74.63
Mpulmonis										100.00	84.28	84.12	73.21	73.31	75.17	73.45	73.78	75.39
Mhominis											100.00	84.04	72.65	72.51	74.98	73.26	73.89	75.28
Mmobile												100.00	72.43	72.18	74.56	73.48	74.19	75.49
Mpneumoniae													100.00	93.58	81.49	70.53	70.62	71.56
Mgenitalium														100.00	81.22	70.84	70.91	71.40
Mgallisepticum															100.00	72.29	72.47	74.19
Phytomali																100.00	85.24	82.01
Phytoaustraliense																	100.00	83.02
Alaidlawii																		100.00

Table S2. Pairwise similarity matrix of core proteins involved in translation. Similarities were calculated using BLASTP 2.4.0 from multiple alignments of 79 core proteins involved in translation. Species belonging to the phylogenetic group Spiroplasma and used in this study are indicated in grey.

GT type	Species	% mapped reads	Number of mapped reads	% genome covered	uncovered bases	% of unmapped reads aligning on *Mcap*	Number of unmapped reads aligning on Mcap recipient	variants detected
Bacteria to bacteria	*Mcap*	97.21	2,589,607	98.91	10,960	0.00	0	0
		97.47	3,222,511	99.57	4,312	0.00	0	4
	Mmc	98.72	2,602,551	99.21	8,553	0.51	477	8
		98.77	2,206,648	99.24	8,280	0.04	33	6
	Mmm	97.49	2,638,223	99.83	2,013	0.11	37	10
		97.78	2,169,053	99.90	1,261	0.21	58	10
	Mlea	96.51	2,333,468	99.00	10,049	0.26	133	9
		96.41	2,870,261	98.73	12,776	0.18	126	9
	Mputr	97.87	2,074,821	99.96	338	0.04	22	1
		97.61	3,153,099	99.95	375	0.06	41	0
	Mflorum	99.80	1,832,600	99.41	4,706	1.27	46	4
		99.35	2,605,319	99.74	2,037	0.53	89	2
Yeast to bacteria	*Mmc*	98.71	2,518,831	99.62	4,150	0.08	28	10
		98.80	2,651,570	99.60	4,356	0.11	36	21
		99.10	2,544,315	99.77	2,509	0.08	18	18
		98.84	1,824,578	99.53	5,156	0.11	24	12
	Mlea	96.37	2,784,149	99.51	4,925	0.11	72	6
		96.11	2,472,548	98.94	10,682	0.16	91	4
		95.58	2,450,091	98.42	15,967	0.14	88	3
		96.69	2,457,706	99.86	1,378	0.12	84	7
	Mputr	97.70	2,601,959	99.47	4,404	0.02	23	7
		97.73	1,819,786	99.91	780	0.02	18	1
		97.43	2,345,550	99.96	361	0.02	20	1
		97.20	2,521,034	99.90	858	0.03	26	0
	Mflorum	98.06	3,202,240	99.98	157	0.18	114	3
		97.52	2,305,321	99.81	1,533	0.15	90	4
		97.14	2,983,976	99.93	555	0.11	93	4
		97.63	1,150,607	99.64	2,896	0.19	52	3

Table S3. Whole-genome sequencing of transplanted genomes.

Species	Plasmid abbreviation	Primer name	Sequence (5'-3')	Size product (bp)
Mycoplasma leachii	pMleaOriC	Mlea-oriC-F	CATGGGATCCGTAGCCATTCTTGCTCTA	1947
		Mlea-oriC-R	CATGGGATCCAATCAATTACTTTAGCAGC	
Mycoplasma putrefaciens	pMputrOriC	Mputr-oriC-F1	CATGGGATCCCCATGAACACGTGCGTGT	2017
		Mputr-oriC-R1	CATGGGATCCCCTGTCAAGCTAGGATTAC	
Mesoplasma florum	pMflOriC	Mfl-oriC-F	CATGGGATCCAACCGTGAACACGAGCGTGT	1925
		Mfl-oriC-R	CATGGGATCCATACTTAATCTCATATAACACCTC	

Table S4. Primers for *oriC*-based replicative plasmids construction.

Species	Transposon localization	Genome annotation
Mycoplasma mycoides* subsp. *mycoides	**MSC_0635**	**Conserved hypothetical protein (predicted lipoprotein)**
Mycoplasma leachii	**MSB_A0138**	**Putative beta-lactamase**
Mycoplasma putrefaciens	**MPUT_0553/MPUT_0554**	**Intergenic region**
Spiroplasma citri	**SPICI-01B-028**	**Hypothetical transmembrane protein**

Primers used for direct sequencing)		
Primer name	Sequence (5'-3')	Annealing temp. (°C)
MT85.1	ACAGTAATTGCGGGTGGATC	60
MT85.2	TTACCGCCTTTGAGTGAGCT	60

Table S5. Transposon localization in genomes selected as donor for genome transplantation experiments.

Species	Primer name	Sequence (5'-3')	Annealing temp. (°C)	Size product (bp)
Mycoplasma capricolum subsp. capricolum	McapCK-F1 McapCK-R1	GTTGATGGTGTAGAAGAACCT GACATCTTCTTCAAAACCGT	55	272
Mycoplasma mycoides* subsp. *capri	MmcGM12-F1 MmcGM12-R1	CTACTATTGAAATTGTAACTGG GATCTGTTGATCAAAATACTCA	55	311
Mycoplasma mycoides* subsp. *mycoides	MmmPG1-F1 MmmPG1-R1	GTGTTGGGTATGAATGCTTATG CATAACTTCAAGCAGAATCACC	50	574
Mycoplasma leachii	P67BG7-L P67BG7-R	GGTAATTCGAATAATGATCCT TAAGTTTATTGAATTAAAGCG	50	1406
Mycoplasma putrefaciens	Mput-F1 Mput-R1	AATGATTTGCCTTTTGTCGAA TCAATAGCTTCATCAGCTCCAA	50	371
Mesoplasma florum	MflorumL1-F1 MflorumL1-R1	CTATTGGAGCAACTCCGGCT GCACCCACTAAACCTGTTCC	50	466
Spiroplasma citri	ScitriGII3-F1 ScitriGII3-R1	CAGAATCTGCCTTATTAAAAGC TATATCCAGCACTAGTATTACC	50	395

Table S6. Primers for species-specific PCR reactions for transplant genotype confirmation.

Species	Genome size (bp)	Restriction enzymes	Number of restriction fragments	Expected fragments sizes (bp)
Mycoplasma capricolum subsp. capricolum	1,010,023	ApaI	2	848,466 + 161,587
		XhoI	2	496,985 + 513,068
Mycoplasma mycoides* subsp. *capri	1,084,610	ApaI	2	770,085 + 314,525
		XhoI	3	589,902 + 269,000 + 225,708
Mycoplasma mycoides* subsp. *mycoides	1,211,730	ApaI	1	1,211,730
		XhoI	3	958,722 + 176,119 + 76,889
Mycoplasma leachii	1,008,965	ApaI	1	1,008,965
		XhoI	3	548,154 + 339,533 + 121,287
Mycoplasma putrefaciens	860,013	ApaI	2	849,494 + 10,519
		XhoI	8	184,057 + 150,946 + 133,365 + 121,130 + 85,882 + 85,753 + 83,960 +14,920
Mesoplasma florum	793,224	ApaI	2	541,387 + 251,849
		XhoI	4	278,693 + 241,953 + 238,126 +34,464

Table S7. ApaI and XhoI-digested transplant genotypes profiles during PFGE assays.

SUPPLEMENTARY REFERENCES

1. Gibson,D.G., Young,L., Chuang,R.-Y., Venter,J.C., Hutchison,C.A. and Smith,H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods*, **6**, 343–5.

2. Janis,C., Lartigue,C., Frey,J., Wróblewski,H., Thiaucourt,F., Blanchard,A. and Sirand-Pugnet,P. (2005) Versatile use of oriC plasmids for functional genomics of *Mycoplasma capricolum subsp. capricolum*. *Appl. Environ. Microbiol.*, **71**, 2888–93.

3. Dordet-Frisoni,E., Sagné,E., Baranowski,E., Breton,M., Nouvel,L.X., Blanchard,A., Marenda,M.S., Tardy,F., Sirand-Pugnet,P. and Citti,C. (2014) Chromosomal transfers in mycoplasmas: when minimal genomes go mobile. *MBio*, **5**, e01958.

4. Zimmerman,C.-U. and Herrmann,R. (2005) Synthesis of a small, cysteine-rich, 29 amino acids long peptide in *Mycoplasma pneumoniae*. *FEMS Microbiol. Lett.*, **253**, 315–21.

5. Aboklaish,A.F., Dordet-Frisoni,E., Citti,C., Toleman,M.A., Glass,J.I. and Spiller,O.B. (2014) Random insertion and gene disruption via transposon mutagenesis of *Ureaplasma parvum* using a mini-transposon plasmid. *Int. J. Med. Microbiol.*, **304**, 1218–25.

6. Lartigue,C., Vashee,S., Algire,M.A., Chuang,R.-Y., Benders,G.A., Ma,L., Noskov,V.N., Denisova,E.A., Gibson,D.G., Assad-Garcia,N., *et al.* (2009) Creating bacterial strains from genomes that have been cloned and engineered in yeast. *Science*, **325**, 1693–6.

7. Murphy,L.D. and Zimmerman,S.B. (2001) A limited loss of DNA compaction accompanying the release of cytoplasm from cells of *Escherichia coli*. *J. Struct. Biol.*, **133**, 75–86.

8. Zusman,D.R., Carbonell,A. and Haga,J.Y. (1973) Nucleoid condensation and cell division in *Escherichia coli* MX74T2 ts52 after inhibition of protein synthesis. *J. Bacteriol.*, **115**, 1167–78.