Vet. Res. (2008) 39:14
DOI: 10.1051/vetres:2007052

www.vetres.org

Original article

Genotyping of *Mycoplasma mycoides* subsp. *mycoides* SC by multilocus sequence analysis allows molecular epidemiology of contagious bovine pleuropneumonia

Aboubakar YAYA[1,2], Lucía MANSO-SILVÁN[1], Alain BLANCHARD[3], François THIAUCOURT[1]*

[1] CIRAD UPR15 "Control of animal diseases", OIE & FAO reference laboratory for CBPP, Montpellier, France
[2] LANAVET, Laboratoire national vétérinaire, Garoua, Cameroon
[3] INRA, Université de Bordeaux 2, UMR 1090 Génomique, diversité, pouvoir pathogène, Villenave d'Ornon, France

(Received 4 May 2007; accepted 30 October 2007)

Abstract – *Mycoplasma mycoides* subsp. *mycoides* SC (*Mmm*SC) is the etiological agent of contagious bovine pleuropneumonia (CBPP). Although eradicated in most developed countries, the disease reappeared in Europe in the 1990s. This reappearance may have been caused either by importation from sub-Saharan Africa, where CBPP is still endemic, or by the reemergence of virulent strains in Europe, as suggested by earlier studies. A multilocus sequence analysis scheme has been developed to address this issue and, most importantly, to be able to monitor new epidemics. The alignment of the full genome sequence of the reference strain PG1 and the partial genome sequence of a pathogenic strain allowed the identification of polymorphic sites. Nineteen initial loci were selected within housekeeping genes, genes of unknown function and non coding sequences. The suitability of these loci for genotyping *Mmm*SC strains was first tested on six strains of diverse geographic origin. The analyses showed that the published PG1 sequence contained a number of specific polymorphisms that were therefore of no use for molecular typing. Among the eight informative polymorphic loci finally selected, only one (*ftsY*) was positioned within a housekeeping gene. Three main groups and 31 different allelic profiles were identified among 51 strains and strain variants examined. Cluster analysis confirmed that European strains from the 1990s did not originate from Africa. It also showed a genetic link between a European strain isolated in 1967 and those found in southern Africa and Australia. This was in agreement with historical data showing that CBPP was introduced in these regions during colonisation in the 19th century.

***Mycoplasma mycoides* / contagious bovine pleuropneumonia / multilocus sequence analysis / molecular epidemiology**

1. INTRODUCTION

Mycoplasma mycoides subsp. *mycoides* SC (*Mmm*SC) is the etiological agent of contagious bovine pleuropneumonia (CBPP) [33]. This disease belongs to the list of notifiable animal diseases from the World Organisation for Animal Health (OIE) because of its economic importance (http://www.oie.int). CBPP induces lesions of pneumonia and pleurisy in cattle and domestic buffaloes and, without treatment, mortality rates can reach up to 50%. CBPP gained a worldwide distribution as a result of live animal trade in the middle of the 19th century [11]. The disease was progressively eradicated from many countries in Europe and North America on the eve of the

* Corresponding author: thiaucourt@cirad.fr

20th century thanks to the strict application of stamping out policies. A strategy that combined vaccination campaigns, detection and slaughter of infected animals and control of animal movements allowed Australia to regain a CBPP-free status in 1973 [8]. In sub-Saharan Africa, CBPP was eradicated in the extreme southern part of Africa but remains enzootic elsewhere. The disease was under control in Africa around 1980, when repeated vaccination campaigns were organised to eradicate rinderpest by use of bivalent vaccines that also provided protection against CBPP. However, its prevalence and distribution have extended dramatically since the eradication of rinderpest in 1993, which implies that no more vaccination campaigns have been organised and subsidised [25]. European countries are now free of CBPP, though some of them have experienced sporadic outbreaks at regular intervals. This was the case in Portugal, where CBPP outbreaks were reported in 1935, 1953, and 1993. Outbreaks were also reported at the border between Spain and France in 1967 and 1982, and in Italy in 1993 [34]. On each occasion, field epidemiological inquiries failed to yield clues regarding the origin of the outbreaks and did not permit to elucidate whether they were the result of reintroduction or resurgence of the disease.

*Mmm*SC, the agent of CBPP, belongs to the so-called *Mycoplasma mycoides* cluster, which includes six mycoplasma subspecies and groups of strains that are closely related from both a phenotypic and a genetic point of view [9, 10]. The organisms most closely related to *Mmm*SC are *M. mycoides* subsp. *mycoides* LC (*Mmm*LC) and *M. mycoides* subsp. *capri* (*Mmc*), which are both etiological agents of contagious agalactia in small ruminants. Recent data have shown that these two subspecies could be grouped into a single entity [45]. Differentiation of the strains belonging to these three taxons is difficult when relying only on the standard growth inhibition technique. Concerning intra-species polymorphism, comparative studies by SDS-PAGE have already shown that strain variability within *Mmm*SC is much more limited than within *Mmm*LC and *Mmc* [9]. Although *Mmm*SC is thought to be a very homogeneous taxon, previous studies have shown that *Mmm*SC strains could be differentiated using molecular typing assays. Analysis of genomic polymorphism by endonuclease restriction of whole genomic DNA showed that different profiles were obtained with *Bam*HI and *Pst*I [31] and that the profiles corresponding to European strains differed from those of African origin. The use of *Hind*III allowed distinguishing various vaccine strains [40]. The evidence of multiple copies of insertion elements present in the *Mmm*SC genome allowed the development of new typing tools based on Southern blot hybridisation. The use of IS1296 as a probe allowed the clear differentiation of recent European strains from those of African origin [7]. The observed difference was explained later on by the identification of an 8.8 kbp deletion in the genome of most *Mmm*SC strains of European origin [43], resulting in a missing IS1296 band. The use of another insertion element, IS1634, also led to different Southern blot profiles [24, 46] although the high copy number of this insertion sequence (N = 60), as compared to that of IS1296 (N = 28), gave rise to profiles that were difficult to analyse. The use of a technique based on sequencing multiple loci, designated multilocus sequence analysis (MLSA), allowed the identification of 15 different allelic profiles within a representative number of *Mmm*SC strains (N = 48) of various origins [22]. In 2004 the whole genome sequence of *Mmm*SC reference strain PG1 was published [48], opening new opportunities for the development of typing tools. Analysis of the PG1 genome sequence showed that the loci and primer pairs previously selected for MLSA were not the most adequate. Some of the primers hybridised on multiple sites, whilst other targeted sequences were duplicated in the PG1 genome, hampering result interpretation. Comparison of PG1 sequences with the sequences of a pathogenic *Mmm*SC strain, obtained through an ongoing genome sequencing project, allowed the identification of a number of polymorphic sites. The objective of this study was to select a set of polymorphic loci to develop a more precise and robust MLSA scheme in order to type *Mmm*SC strains from

Table I. List of strains.

Profiles	Strains	Origin	Locality	Year of isolation	Comments	References	Provided by
A00	Gemu Gofa	Ethiopia	Turmi?	1974			NVI Dr Fikre
A00	83162	Chad	Bougor?	1983			CIRAD
A01	91130	CAR	Bouar?	1991			LCV Bangui, Dr Condolas
A02	Afadé	Cameroon	Afade	1965			LaboFarcha Dr Provost
A02	Vom	Nigeria	Vom?	1965	Sheep orign		Vom Laboratory
A03	B17	Chad	N'Djamena	1959			LaboFarcha Dr Balis
A04	8740	Cameroon	Touroua	1987		[50]	LANAVET, Garoua Dr Tulasne
A04	8740-Rita	Cameroon	Garoua	1987	8740 after passage on cattle		LANAVET, Garoua Dr Yaya
A04	8740/11	Cameroon	Touroua	1987	11th passage		CIRAD
A04	8740/53	Cameroon	Touroua	1987	53rd passage		CIRAD
A04	98029	Benin	Kandi?	1998			Dr Hendrikx
A04	9048	Nigeria	Sokoto?	1990			Labocel Niamey Dr Bloch
A05	Muguga	Kenya	Athi-river	1997		[47]	KARI, Muguga, Dr Wesonga
A05	95014	Tanzania	Shinyanga?	1995			FAO Dr Provost
A05	94111	Rwanda	Kigali	1994			Action Nord-Sud, Mme Isakov
A06	2000-033	Ethiopia	Gimbi	2000			NVRIC Sebata, Dr Sentayou
A07	7721	Mauritania	Tiguent	1997			CNERV, Nouakchott, Dr Brun
A08	96010	Mauritania	Noukchott?	1996			NVRIC Sebata, Dr Sentayou
A09	Fatick	Senegal	Fatick	1968			IEMVT (CIRAD)
A10	99042	Mali	Kolokani?	1999			LCR Bamako, Dr Niang
A11	2000-005	Burkina Faso	Ouagadougou?	2000			LDRV, Ouagadougou, Dr Bado
A12	2003-036	Mali	Segou?	2001			LCV Bamako, Dr Niang
A12	87137-10	Burkina Faso	Banfora	1987			LDRV, Ouagadougou, Dr Rouille
A12	98011	Ivory Coast	?	1998			LCPA Bingerville, Dr Cissé
A12	9050-529-1	Ivory Coast	Korhogo?	1990			LCPA Bingerville Dr Domenech
B01	8988-1	Namibia	Oudangwa	1989			CVL Windhoek, Dr Hehnen
B02	97039	Namibia	Rundu?	1994			CVL Windhoek, Dr Huebschle
B02	97041	Botswana	Nata?	1995			CVL Windhoek, Dr Huebschle

all over the world and to elucidate the molecular epidemiology of CBPP.

2. MATERIALS AND METHODS

2.1. *Mmm*SC strains and isolates

Table I lists the names and origin of the 51 strains and strain variants of *Mmm*SC used in this study. The strains, selected in order to cover the entire geographic distribution of CBPP, were isolated from Africa (N = 41), Europe (N = 5), Australia (N = 2), and India (N = 1), whereas the origin of strains Lederle, Asmara, and PG1 is unknown. The oldest strain of known origin was isolated in 1936 (strain V5 from Australia) and the most recent in 2004 (strain 2004-003 from Zambia). The date of isolation of six strains was not precisely known. Although the majority of the isolates were of bovine origin, a few isolates from small ruminants (Vom, 94158 and 99048) were also included.

Strains were grown in a modified Hayflick medium containing sodium pyruvate and glucose [39]. At the end of the exponential phase of growth, 2 mL of culture were centrifuged at 12 000 × *g* for 15 min at 4 °C. The pellet was washed once in PBS, and re-suspended in 100 µL of distilled water, to which 150 µL of lysis buffer (100 mM Tris HCl, pH 8.5; 0.05% Tween 20; 0.24 mg/mL proteinase K) were added. Following incubation at 60 °C for 30 min, proteinase K was heat-inactivated at 95 °C for 5 min. Samples in lysis buffer were stored at –20 °C until analysis and they were added to the PCR mix without further DNA extraction.

2.2. Identification of suitable loci for MLSA

Partial genomic sequences of the *Mmm*SC pathogenic strain 8740-Rita, obtained from an ongoing whole genome sequencing project, were used to evaluate the polymorphism existing between this strain and the reference strain PG1 (NC_005364). Although the final assembly of the 8740-Rita genome has not yet been obtained, the available contigs encompassed most of the DNA

Table I. Continued.

Profiles	Strains	Origin	Locality	Year of isolation	Comments	References	Provided by
B02	97038	Botswana	?	1995		[1]	CVL Windoek, Dr Huebschle
B03	T2/34	Tanzania	Morogoro?	1956			CIRAD
B03	99048	India	Izatnagar?	1999	Goat origin	[38]	IVRI, Izatnagar, Dr Srivastava
B03	Gladysdale	Australia	Gladysdale	< 1965			CIRAD
B03	Pillai	Sudan	Malakal?	1960			CIRAD
B04	2004-003	Zambia	Sesheke	2004			FAO Dr Amanfu
B05	V5	Australia	?	1938	Vaccine	[6]	CIRAD
B06	Asmara	Eritrea	Asmara	unknown	Vaccine		CIRAD
B07	99021	Tanzania	T1/B	1999	T1 variant	[47]	CIRAD
B07	T1sr	Tanzania	Korogwe?	1951	Vaccine		CIRAD
C01	94158	Portugal	?	1994		[5]	LNIV, Lisboa Dr Brandao
C01	PO2	France	Pyrennees	1981			CIRAD
C01	99065	Italy	Brescia	1992			IZS Teramo Dr Pini
C02	9335-170	Italy	?	1992			IZS Teramo Dr Santini
C03	PO1967	France	Pyrennees	1967			CIRAD
D01	97009	Ethiopia	Makale	1997			NVI Dr Berhe
D02	2003-011	Eritrea	Assab?	2003			MOA Asmara, Dr Tesfaalem
E01	DK32	Senegal	Dakar	unknown	Vaccine	[32]	CIRAD
E02	9373-804	Guinea	Kankan?	1993			LDV Conakry, Dr Condé
F01	PG1	Unknown	?	< 1931	Reference strain		CIRAD
G01	Lederle	Unknown	Puigcerda?	< 1967	Vaccine		CIRAD
H01	Filfili	Senegal	Filfili	1965			CIRAD
I01	KH3J	Sudan	Juba	< 1948	Vaccine	[16]	CIRAD

Localities followed by a ? were chosen to obtain a geographic positionning. CAR = Central African Republic; CIRAD = centre de coopération internationale en recherche agronomique pour le développement; CVL = central veterinary laboratory; FAO = food and agriculture organisation; IVRI = Indian veterinary research institute; IZS = instituto zooprofilactico sperimentale; KARI = Kenya agricultural research institute; LABOCEL = laboratoire central d'élevage; LANAVET = laboratoire national vétérinaire; LCPA = laboratoire central de pathologie animale; LCV = laboratoire central vétérinaire; LDRV = laboratoire départemental de recherche vétérinaire; LDV = laboratoire de diagnostic vétérinaire; LNIV = laboratorio nacional de investigacao veterinaria; MOA = ministry of agriculture; NVI = national veterinary institute; NVRI = national veterinary research institute.

sequences homologous to PG1. Only insertion sequences and large duplicated zones in the PG1 genome, unsuitable for MLSA, were excluded. As a whole, 800 kbp were available for comparison. Alignments were performed using the AlignX program of Vector NTI software (Vector NTI v10.3.0, Invitrogen corporation, Carlsbad, CA, USA). Polymorphic sites were recorded and positioned in the PG1 genomic sequence map. Nineteen polymorphic positions were chosen for initial validation. Seven of them resided in housekeeping genes, six within lipoprotein or transmembrane protein genes and a further six within non-coding sequences and they were all distributed along the entire PG1 genome, as shown in Figure 1. The suitability of these loci for MLSA was first validated by sequencing the homologous regions of five additional *Mmm*SC strains: the original PG1 strain used for genome sequencing in Sweden and four other strains from Europe (PO1967), East Africa (94111), West Africa (Filfili), and Australia (Gladysdale). Eight potentially interesting loci were then chosen for final validation using the whole set of strains (Tab. II).

2.3. PCR and sequencing

PCR was carried out in a 50 μL reaction containing: 0.5 unit of Taq DNA polymerase (Qiagen SA, Courtaboeuf, France) in its corresponding buffer (including 1.5 mM $MgCl_2$), 300 μM of dTTP and dATP, 150 μM of dGTP and dCTP, 0.4 μM of each primer and 1 μL of sample diluted 1/100. Amplifications were performed using GeneAmp PCR Systems 2400 and 9700 (Perkin Elmer, Courtaboeuf, France). Thermal cycling consisted in an initial denaturation step at 94 °C for 5 min, followed by

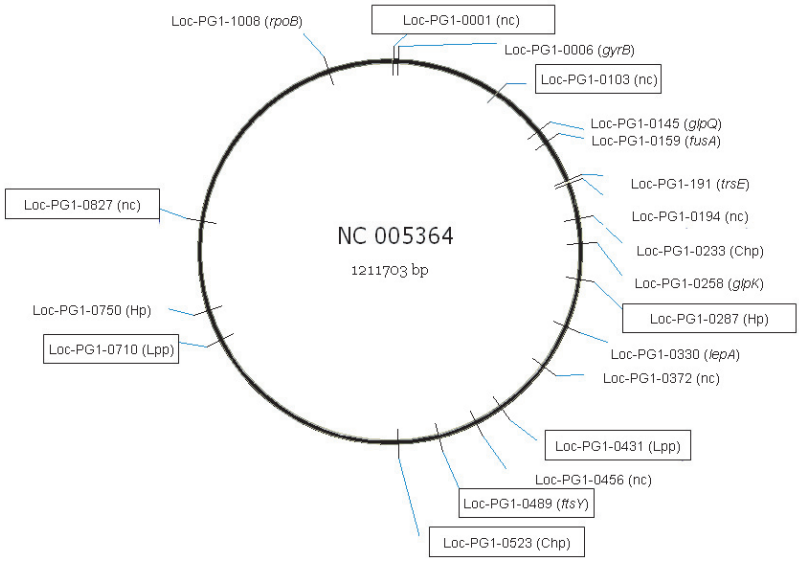

Figure 1. Distribution of polymorphic positions in the PG1 genome (NC_005364) and loci used for MLSA. Polymorphic positions are identified giving the corresponding CDS number in the PG1 sequence. The loci retained for the final MLSA scheme are boxed.

35 cycles of denaturation at 94 °C for 15 s, annealing at 52 °C for 30 s and extension at 72 °C for 90 s. The final extension step was maintained at 72 °C for 7 min. The size and purity of the amplicons were controlled by electrophoretic separation in 1% agarose gels and samples with relevant features were sent to Genome Express (Genome Express Cogenics, Meylan, France) for sequencing with the corresponding primers. Both forward and reverse strands were sequenced systematically.

2.4. Allelic profile analysis

The sequences obtained from each corresponding forward and reverse primers were assembled using Vector NTI Suite™ (InfoMax, 2001) and the extremities corresponding to single strand sequences or showing aberrant features were trimmed. The sequences obtained from different strains for each locus were aligned using ClustalW (Vector NTI). Polymorphic sites were recorded, carefully checking the corresponding sequence chromatograms. A new allele number was assigned to each change in the nucleotide sequence. In the end, each strain was characterised by an allelic profile, corresponding to the combination of allele numbers for each of the eight selected loci. This resulted in a matrix comprising 31 different allelic profiles (Tab. III), which was analysed using eBURST V3, a software developed for the analysis of multilocus sequence typing (MLST) datasets[1]. Standard parameters were used, except for the number of loci, which was fixed at eight. Each strain was then characterised by a letter, corresponding to its group, followed by a number, corresponding to the allelic profile.

2.5. Molecular epidemiology analysis

When the exact origin of the strain was known, the geographic coordinates were inferred from a dedicated website[2]. When this information was not available, a putative geographic location was assigned to the strain, taking into account the known distribution of CBPP in the country. For example, strains originating from Namibia were positioned north of the veterinary fence that separates CBPP

[1] http://eburst.mlst.net/v3/

[2] http://earthsearch.net/intSearch/

Table II. List of primers used for PCR and sequencing. The positions refer to the nucleotide sequence NC_005364 of reference strain PG1.

Loci names	Gene	Primers	Position	Sequences (5'- 3')	Amplicon size (bp)
Loc-PG1-0001	nc	Loc-PG1-0001-F	1272	AACAAAAGAGATCTTAAATCACACTTTA	538
		Loc-PG1-0001-R	1809	CCTCTTGTTTAACTTCTAGATCAGAAT	
Loc-PG1-0103	Lpp	Loc-PG1-0103-D	121796	GATGGATATAATCTATACTAGCATTTA	1321
		Loc-PG1-0103-F	123116	CCTTATATAGATAAAACTCCTCCTTA	
Loc-PG1-0287	nc	Loc-PG1-0287-F	328750	GATTGCTTTAATCAATTTCTTACTGA	545
		Loc-PG1-0287-R	329294	GGATAACCTTGATTTTTTATTTGCTTTA	
Loc-PG1-0431	Lpp	Loc-PG1-0431-F	485201	CAATTCTTTAAATTTGGGTTTGTT	608
		Loc-PG1-0431-R	485808	CTTGCAAGAGTATTTAGATTTGATTAAA	
Loc-PG1-0489	ftsY	Loc-PG1-0489-F	555289	GTTAGTTGTTGAAATGTTAGATAT	756
		Loc-PG1-0489-R	556044	CCCATATCAGTTTGGATTAA	
Loc-PG1-0523	Chp	Loc-PG1-0523-F	597315	ACAGCATTTGATCAAGATTTAAGTAGTT	824
		Loc-PG1-0523-R	598138	TTACCTAGGTGTTTAAAACTTCATTTG	
Loc-PG1-0710	Hp	Loc-PG1-0711 –F	816093	CCAGTTGAACCATTTATTTTATATATACCT	643
		Loc-PG1-0711-R	816735	AAATATAAGTGGTGCTGGAATAACA	
Loc-PG1-0827	nc	Loc-PG1-0827-F	940048	AGTTGTACAACTGTATGAATCTATGATTAT	619
		Loc-PG1-0827-R	940666	CAGGATATACTTCAAAAATTAAAGGTTT	

Lpp: Lipoprotein, Chp: conserved hypothetical protein, Hp: hypothetical protein, nc: non coding sequence.

infected regions in the North from CBPP-free zones in the South. Strains were then positioned on a map, displaying group and allele numbers.

3. RESULTS

3.1. Choice of MLSA loci and initial validation

The first objective of this study was to find suitable loci for multilocus sequence analysis of *Mmm*SC strains. Nineteen polymorphic sites observed by comparing the genome sequence of *Mmm*SC reference strain PG1 with the partial genome sequence of pathogenic strain 8740-Rita were chosen for initial validation. Five strains, including the original PG1 used for genome sequencing in Sweden and four other strains from diverse geographic origins, were used for this initial validation.

Two loci were considered unsuitable because the polymorphisms that were initially observed between PG1 and 8740-Rita were apparently due to errors in the published PG1 sequence. In Loc-PG1-0330 (*lepA*) both the G in position 379030 and the T in position 379050 are actually A. These two modifications result in a single amino acid modification in the deduced protein sequence. In Loc-PG1-0191 *(trsE*), the T in position 227175 should be deleted and a T should be inserted after position 227214. These changes modify significantly the sequence of the *trsE* product (12 amino acids). Interestingly, when this locus was sequenced in the vaccine strain T1/44, an insertion sequence (IS1634) was found interrupting the *trsE* gene at position 227363. The putative function of the *trsE* gene product is not well known, though it may be involved in the transport of large molecules, intracellular trafficking and secretion. The disruption of *trsE* in strain T1/44 indicates that this gene is not essential, though whether this has contributed to the attenuation of the vaccine strain is yet to be elucidated. Four additional loci were excluded as the polymorphism was found to be limited to the sequence of strain PG1, whereas the other five *Mmm*SC strains exhibited identical sequences. Three of these loci where located within the housekeeping genes *fusA* (Loc-PG1-0159), *glpK* (Loc-PG1-0258), and *rpoB* (Loc-PG1-1008), whereas Loc-PG1-0372 corresponded to a non-coding sequence. Another four loci (Loc-PG1-0145, Loc-PG1-0194, Loc-PG1-0456, and Loc-PG1-0750) were eliminated because they did not add to the discriminatory power of Loc-PG1-0489. Loc-PG1-0233 was also excluded, since only

Table III. Matrix showing the various allelic profiles obtained through MLSA analysis of the eight loci.

Profiles	Type strains	Loc-PG1-0001	Loc-PG1-0103	Loc-PG1-0287	Loc-PG1-0431	Loc-PG1-0489	Loc-PG1-0523	Loc-PG1-0710	Loc-PG1-0827
A00	Gemu Gofa	1	2	1	1	2	1	1	1
A01	91130	5	2	1	1	2	1	1	1
A02	Afadé	1	3	1	1	2	1	1	1
A03	B17	1	3	1	1	2	1	1	2
A04	8740	4	3	1	1	2	1	1	1
A05	Muguga.	1	1	1	1	2	1	1	1
A06	2000/033	1	1	1	1	3	1	1	1
A07	7721	1	2	4	1	2	1	1	1
A08	96010	2	2	4	1	2	1	1	1
A09	Fatick	1	2	4	1	4	1	1	1
A10	99042	1	2	4	1	2	2	1	1
A11	2000/005	2	2	4	1	2	2	1	1
A12	87137	3	2	4	1	2	2	1	1
B01	C8988	1	3	1	1	4	3	2	1
B02	970039	1	2	1	1	4	3	2	1
B03	Gladysdale	1	2	1	5	4	3	2	1
B04	2004/003	1	4	1	1	4	3	2	1
B05	V5	1	3	1	5	4	3	2	1
B06	Asmara	1	3	1	1	4	3	1	1
B07	T1SR	1	5	1	1	4	3	1	1
C01	94158	1	2	1	1	5	5	2	2
C02	9335	1	2	1	2	5	5	2	2
C03	PO1967	1	2	1	1	5	3	2	2
D01	97009	4	2	1	1	4	3	1	1
D02	2003/011	4	2	1	1	4	3	1	2
E01	DK32	3	2	1	1	1	1	1	1
E02	9373/804	3	3	1	1	1	1	1	1
F01	PG1	1	2	5	6	6	3	2	2
G01	Lederlé	6	1	3	4	4	4	2	3
H01	Filfili	4	1	2	1	1	1	1	1
I01	KH3J	1	2	1	3	4	1	1	1

two alleles could be defined in this sequence. All strains of European origin, and none of the other, bore allele 2 in this locus, making it a good tag for this origin. It must be noted that PG1, the origin of which is unknown, also displayed allele 2 in this locus. On the contrary, strain Lederle, which had been used as a vaccine in Spain, displayed allele 1, suggesting that it may not be of European origin. Finally, in Loc-PG1-0194, three alleles were defined, though the results were redundant and did not allow a better differentiation of the strains. In the end, eight loci were selected (Tab. II) to perform MLSA on the whole set of *Mmm*SC strains.

Interestingly, analysis of sequence variations amongst *Mmm*SC strains showed that many coding sequences (CDS) were interrupted in PG1, whereas larger CDS were identified in the other *Mmm*SC strains. This was the case of four out of the eight selected loci. The CDS MSC_0103 and MSC_0104 constituted a single CDS of around 370 amino acids, as well as MSC_0287 and MSC_0288, forming a single CDS of 713 amino acids, and MSC_0710 and MSC_0711, which constituted a single CDS of 934 amino acids. The CDS MSC_0827 was extended from 99 to 344 amino acids.

3.2. Robustness of the MLSA

To evaluate the robustness of the method, cultures of pathogenic strain 8740, the same strain at the 11th and 53rd passages and after a passage in bovines (8740-Rita), were analysed. These were considered variants of the same strain, differing only by a few, well characterised, in vitro or in vivo passages. In addition, three strains (94111, Filfili, Gladysdale) were tested in duplicate at three months of interval. No differences were evidenced either between the duplicates or between the four strain variants differing in the number of passages. This result was important to show that the selected loci were stable over a few passages, implying that the sequence differences were correlated to the natural evolution of the strains in vivo. In the case of high mutation rates the signal would be blurred or biased through in vitro passage, whereas in the case of low mutation rates the loci would have been unsuitable for use as molecular epidemiology tools.

3.3. Alleles defined in the different loci

3.3.1. Alleles defined in non-coding sequences

Five alleles were defined in locus Loc-PG1-0287. Two alleles gathered the majority of the strains (46 out of 51), with allele 4 found only in Western Africa, while allele 1 had no geographic specificity.

Three alleles were defined in locus Loc-PG1-0827. Forty-one strains bore the same allele. The other alleles corresponded to strains found in Europe and Africa, as well as PG1.

Six alleles were defined in Loc-PG1-0001, with the greatest variability found among strains from West Africa.

3.3.2. Alleles defined in genes of unknown function: Lipoproteins (Lpp) and Conserved hypothetical proteins (Chp)

The smallest number of alleles was 2 for locus Loc-PG1-0710. The observed mutation induced an amino acid modification in the predicted putative lipoprotein (L to S). There was a clear geographic segregation of the alleles (Fig. 2). With the exception of strain Pillai, all strains found in inter-tropical Africa bore allele 1.

Four different alleles were defined in Loc-PG1-0523. However, some *Mmm*SC strains could not be typed, as this locus was located within the 8.8 kbp deletion that was found to be specific to recent European strains. A fifth allele was therefore created in order to type this last group of strains using eBURST. Interestingly, the only strain of European origin for which this locus could be sequenced, PO1967, bore the same allele as the strains found in southern Africa and Australia. This allele was characterised by six polymorphic positions that differed in other *Mmm*SC strains (Tab. IV). This gave more strength to the genetic link demonstrated by this allele.

Five alleles were defined in locus Loc-PG1-0103. They corresponded to a variable number of five to nine trinucleotide repeats (AAT) coding for asparagine (N) within a gene coding for

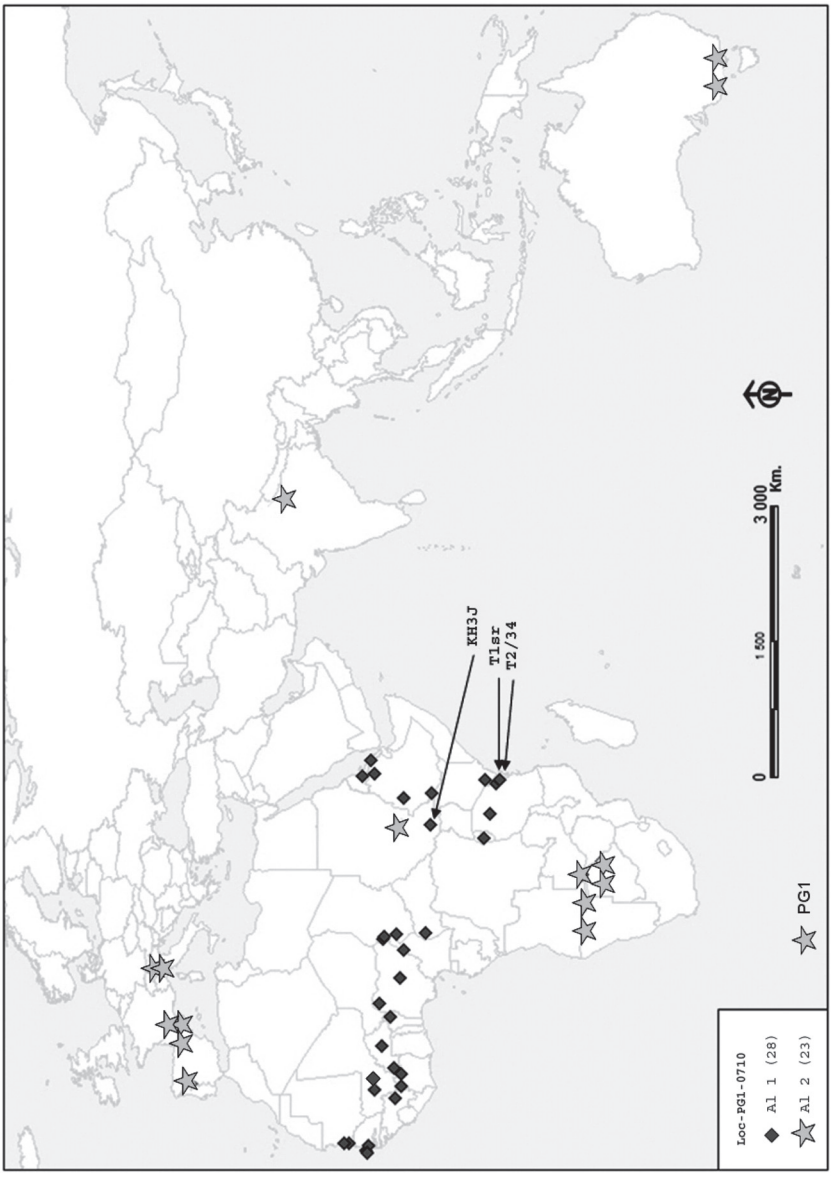

Figure 2. Geographic distribution of alleles for Loc-PG1-0710. Except for strain Pillai, from Sudan, all strains of African origin display allele 1 in this locus, whilst the others display allele 2. Note that strain PG1, located by the legend, was not positioned on the map because its geographic origin is unknown.

Table IV. Polymorphism in the different allelic profiles. The bottom line represents the consensus sequences. Nucleotides identical to the consensus sequence are left blank. Deletions are indicated by][. The second line shows the position of polymorphic sites in the PG1 genome (sequence NC_005364). K: number of oligonucleotide repeats (TCAACAAGA); LL: GAACCTGTTGTTAATCCAAGATAAACAAGTTGAAATA.

Groups	Profiles	Type strains	Loc-PG1-0001										Loc-PG1-0103	Loc-PG1-0287			Loc-PGA-0431							Loc-PG1-0489					Loc-PG1-0523								Loc-PG1-0710	Loc-PG1-0827		Number of strains
			1500	1518	1523	1524	1527	1528	1529	1530	1531	1532	122439	328887	329008	329113	485305	385470	485517	485524	485567	485572	485615	555508	555733	555796	555799	555914	597455	597609	597614	597615	597616	597618	597626	597647	816297	940184	940351	
A	A00	Gemu Gofa																							4																2
	A01	91130			A																				4																1
	A02	Afadé											6ATT												4																2
	A03	B17											6ATT												4															C	1
	A04	8740				T							6ATT												4													G			6
	A05	94111											4ATT												4																3
	A06	2000/033											4ATT												4			A													1
	A07	7721												T											4																1
	A08	96010					][	][	][	][	][	][		T											4																1
	A09	Fatick												T											3																1
	A10	99042												T											4				G												1
	A11	2000/005					][	][	][	][	][	][		T											4				G												1
	A12	2003/036				T				A				T											4				G												4
B	B01	C8988											6ATT												3					A	T	T	C	A	T		G			1
	B02	970041																							3					A	T	T	C	A	T		G			3
	B03	Pillai																			C	C			3					A	T	T	C	A	T		G			4
	B04	2004/003											7ATT												3					A	T	T	C	A	T		G			1
	B05	V5											6ATT								C	C			3					A	T	T	C	A	T		G			1
	B06	Asmara											6ATT												3					A	T	T	C	A	T					1
	B07	T1SR											8ATT												3					A	T	T	C	A	T					2
C	C01	94158																							3	][	T						No amplification				G		C	3
	C02	9335															T						T		3	][	T						No amplification				G		C	1
	C03	PO1967																			C	C			3	][	T			A	T	T	C	A	T		G		C	1
D	D01	97009				T																			3					A	T	T	C	A	T					1
	D02	2003/011				T																			2					A	T	T	C	A	T				C	1
E	E01	DK32				T																			5															1
	E02	9373/804				T							6ATT												5															1
F	F01	PG1													T					A				A	3	][	T			A	T	T	C	A	T		G		C	1
G	G01	Lederle	T	T			][	][	][	][	][	][	4ATT		][			G							3					A	T	T	C	A	T	T	G	A		1
H	H01	Filfili				T							4ATT			A									5															1
I	I01	KH3J																	A						3															1
		Consensus	C	A	G	A	T	A	G	T	A	A	5ATT	A	A	G	C	A	G	G	T	G	C	T	K	LL	A	G	T	G	A	A	A	G	G	C	A	G	G	

a putative lipoprotein. There was no apparent geographic segregation of the various alleles.

3.3.3. Allele defined in the housekeeping gene **ftsY**

Six alleles were defined in Loc-PG1-0489, corresponding to *ftsY*. The majority of the strains were segregated in two alleles, comprising similar numbers (23 and 18 strains respectively). Allele 2 was found only on *Mmm*SC strains from sub-Saharan origin. Allele 4 was displayed by strains from Southern Africa, Australia, and India, as well as by some strains from East Africa. It was also found on one old strain from West Africa and a strain of unknown origin (Lederle), which was formerly used as a vaccine in Europe.

All strains isolated in Europe bore allele 5 on this locus (Fig. 3), strain Lederle, formerly used as a vaccine in Spain, has an unknown origin.

3.4. Strain groups by MLSA analysis

All 51 strains were characterised by an allelic profile and 31 different allelic profiles were identified (Tab. III). The genetic events underlying the description of the various alleles are described in Table IV. Analysis using eBURST allowed the identification of three main groups of strains (comprising 43 strains). Eight strains were included into groups gathering only two strains or considered as singletons. The geographic positioning of the various groups and allelic profiles is shown in Figure 4.

Group A was the largest, including 25 strains (Fig. 5). Thirteen different allelic profiles were clustered within this group. The largest (A04 and A12) gathered up to four strains. Group A strains were found mostly in West and central Africa, though some of them were found in East Africa (Fig. 4). It must be noted that all *Mmm*SC strains isolated after 1994 in East Africa presented the same allelic profile (A05).

Group B was the second largest, including 13 strains that were subdivided into 7 allelic profiles (Fig. 6). Group B strains were found in southern Africa, East Africa, Australia and India (Fig. 4). The most common allelic profile (B03) was shared by four strains of various origins (Australia, Tanzania, Sudan, and India). Most of group B strains found in East Africa had been isolated before 1970 (Pillai, T1sr, T2-34, and a vaccine strain received from a laboratory at Asmara, the origin of which remains undetermined).

Group C corresponded to strains found solely in Europe (Fig. 4 and 6).

Finally, eight strains could not be included into any of the three main groups, as they displayed peculiar allelic profiles. In West Africa peculiar profiles corresponded to old strains isolated in Senegal and Guinea and in East Africa to a vaccine strain (KH3J) also isolated long ago. Two strains isolated recently from northern Ethiopia and Eritrea also bore peculiar allelic profiles. Finally, PG1 was the strain that exhibited the most peculiar allelic profile.

4. DISCUSSION

4.1. Genomic analysis

The objective of this study was to establish a more robust MLSA scheme for *Mmm*SC strains based on whole genome sequence comparison. Analysis of PG1 sequences showed that many primers designed on a previous MLSA scheme [22] hybridised at multiple sites or targeted sequences that were duplicated in the PG1 genome. The initial validation of 19 selected loci in this study revealed that sequencing errors may remain in the published PG1 genome. Out of the 19 initial polymorphic loci, two corresponded to sequencing errors. It must be noted that, if this result was extrapolated to the 267 polymorphic sites found between the sequences of strains PG1 and 8740-Rita, there may be around 20 errors in the published PG1 sequence. Another four out of the 19 polymorphic positions seemed to be specific to the PG1 sequence. The peculiarity of the PG1 genome has already been shown. For example, this strain possesses a large DNA repeat of 24 kbp, which is absent in other *Mmm*SC strains [3]. The peculiarities of PG1 may be due to a natural divergent evolution of the strain. However, it is difficult to substantiate this hypothesis, since

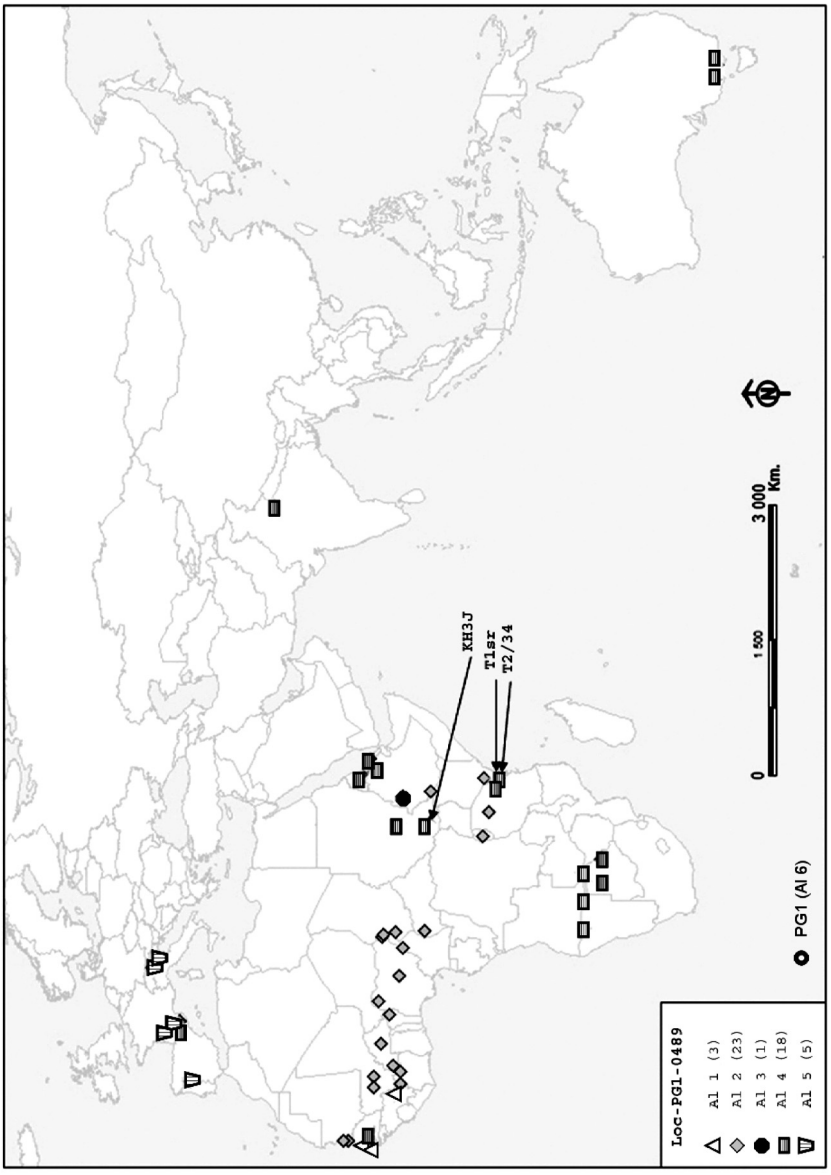

Figure 3. Geographic distribution of alleles for Loc-PG1-0489. Six alleles were defined in this locus. Allele 2 was found only on *Mmm*SC strains from sub-Saharan origin. Allele 4 was displayed by strains from Southern Africa, Australia, and India, as well as by some strains from East Africa. It was also found on one old strain from West Africa and a strain of unknown origin (Lederle), which was formerly used as a vaccine in Europe. Note that strain PG1, located by the legend, was not positioned on the map because its geographic origin is unknown.

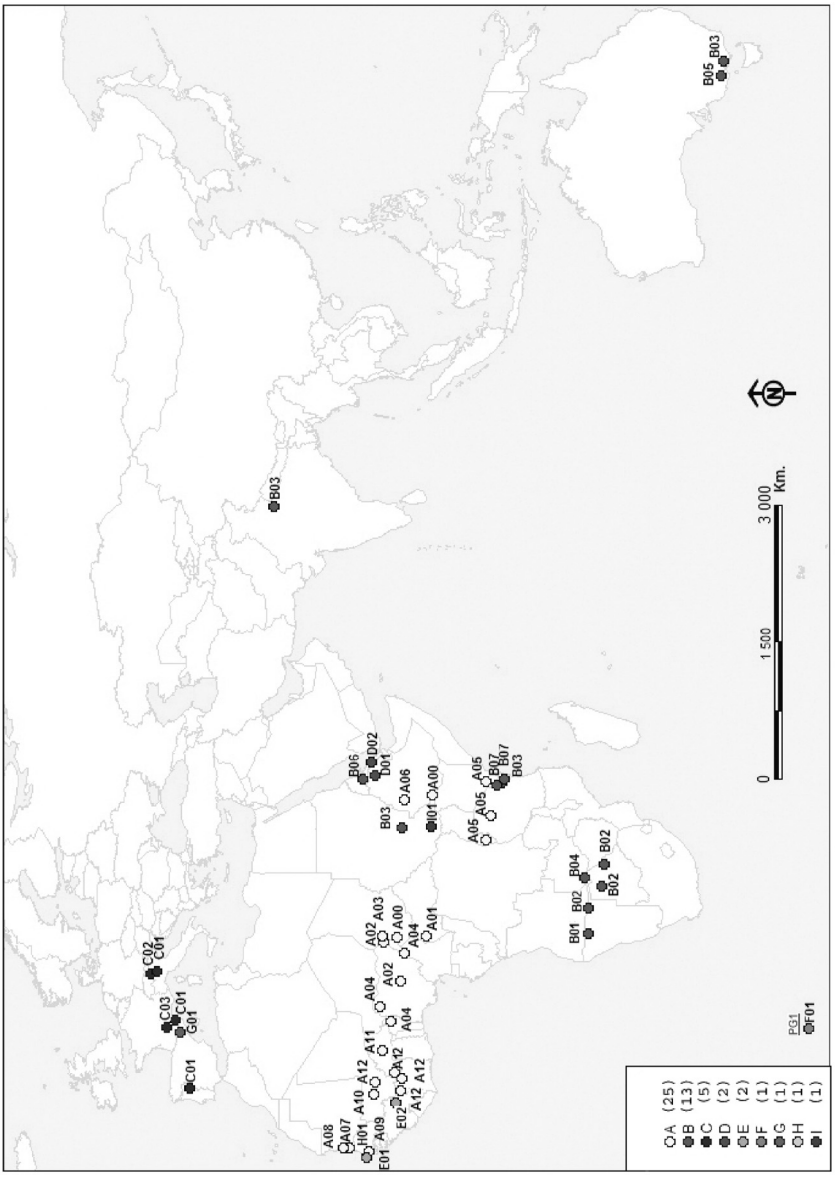

Figure 4. Geographic distribution of allelic profiles. There is a clear correlation between the allelic profile of *Mmm*SC strains and their geographic origin. Strains of group A are found in sub-Saharan Africa and nowhere else. This could be an indication that CBPP was present in Africa prior to colonisation time. On the other hand the genomic link between strains from European origin (group C) and those found in Southern Africa, Australia, India and some found in East Africa (group B) can be correlated with historical data describing CBPP expansion during colonisation in the 19th century. Note that strain PG1, located by the legend, was not positioned on the map because its geographic origin is unknown.

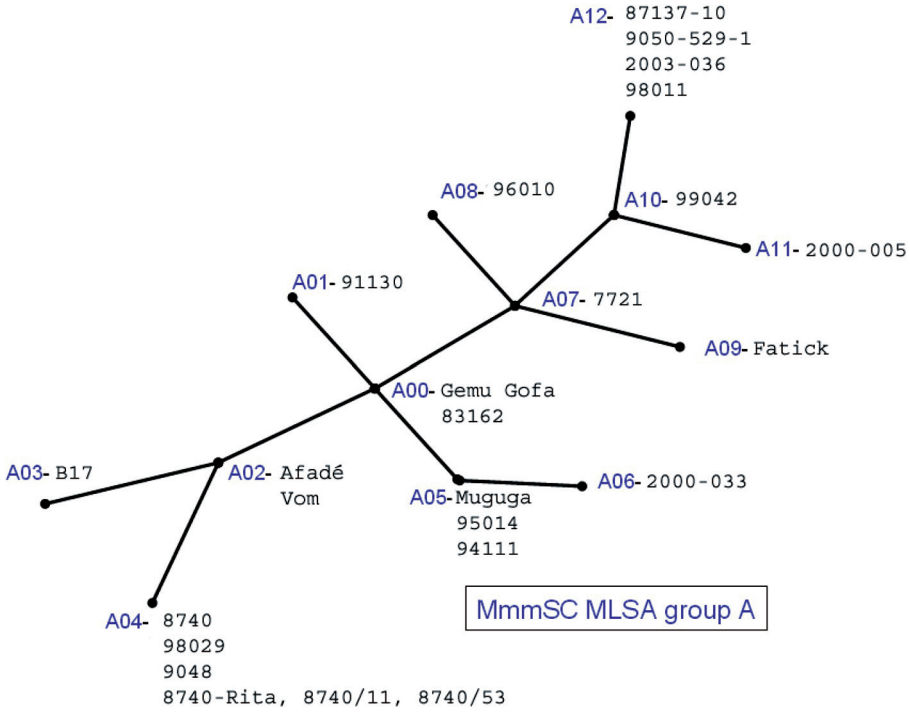

Figure 5. Tree obtained by eBURST for group A strains. Group A is subdivided into 13 allelic profiles and gathers strains found only in sub-Saharan Africa. The largest allelic profile (A12) comprises four strains. However, these strains were collected in neighbouring countries on well known cattle trade routes. This is a likely explanation for finding strains of similar allelic profiles at different locations.

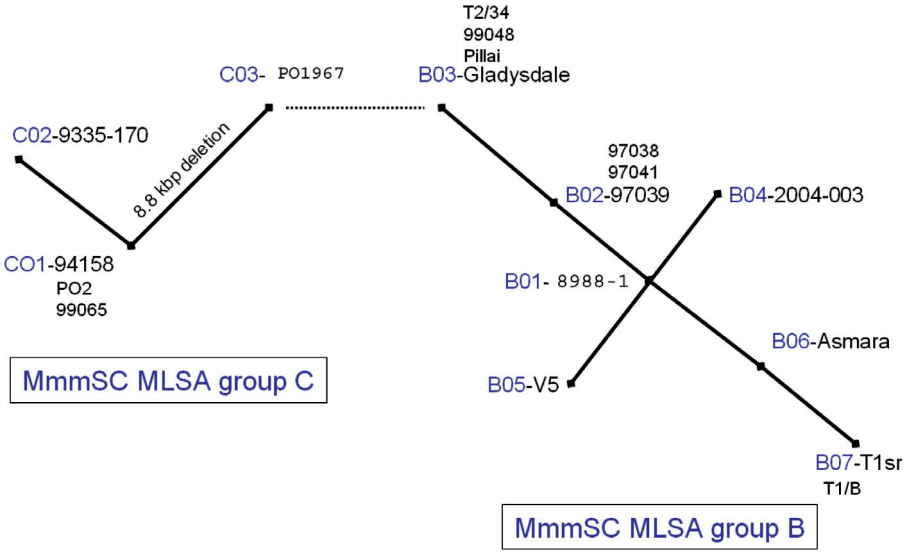

Figure 6. Tree obtained by eBURST for group B and C strains. Group B is subdivided into seven allelic profiles and gathers strains found in Southern and East Africa, as well as strains isolated in Australia and India. Group C gathers European strains. Strain PO1967 is an intermediate between group B and group C.

the exact origin and year of isolation (before 1931) of the strain are unknown. More likely, the divergence of PG1 may result from the many in vitro passages. A single housekeeping gene (*fts*Y) was validated for use in the new MLSA scheme. It was noteworthy that polymorphisms in this gene included the presence of duplicated nonanucleotides coding for a stretch of three amino acids (DQQ) present in three to five copies, as well as the deletion of a rather long DNA fragment (35 bp). Since this gene, coding for a bacterial signal-recognition particle receptor, plays a fundamental role in mediating membrane targeting and insertion of nascent proteins [2], it is likely that the polymorphic sites are not positioned on the active sites of this molecule.

4.2. Epidemiological analysis

The presence of group A strains only in West and central Africa may indicate that these strains have evolved separately from those present in Europe. This may also indicate that CBPP existed in the continent prior to colonisation and presumed introduction of CBPP with live cattle. Such a hypothesis was already raised when a traditional "vaccination" procedure, consisting in inoculating pleural fluid subcutaneously at the bridge of the nose, was described in West Africa [4]. The local inflammation generally led to a reaction of the periosteum, with development of a pseudo-horn, which resulted in the erroneous description of a new cattle breed called *Bos triceros* [35]. These traditional procedures were unknown in Europe, where the inoculation point, as described by Willems, was the tip of the tail [49]. Furthermore, a possible ancestral allelic profile (A00) was identified in strains isolated in Chad and the Gemu-Gofa region in Western Ethiopia. In these regions the influx of exotic cattle was minimal in the 19th century. In contrast, the presence of group B strains in East Africa could reflect an influx of affected cattle from Europe. This may have taken place directly from European countries (i.e. from Italy to its Eritrean colony) or indirectly, as when the 1868 British expedition to Abyssinia used oxcarts from India to pull artillery in the mountainous regions [19] and might have thus introduced CBPP in East Africa.

The geographic location of strains sharing the same allelic profile agreed with known epidemiological data. Therefore, the presence of multiple strains with the same allelic profile certainly does not result from an MLSA lack of discriminatory power. For example, the presence of allelic profile A05 in many countries in Eastern Africa can be explained by a 'clonal expansion' of the initial strain of Kenyan origin that led to the re-contamination of countries such as Tanzania and Rwanda in the 1990s [25]. Similarly, the presence of the same allelic profile (A12) in Mali, Burkina Faso, and Ivory Coast is in agreement with the transhumance and trade routes followed by cattle herds that are raised in Sahelian countries and exported to meat markets near large cities in the South.

Although strain PO1967, of European origin, belongs to group C, its sequence in Loc-PG1-0233 showed that it is closely related to Group B strains. The localisation of Group B strains in this MLSA scheme corroborated what was known regarding the transmission of CBPP from Europe to overseas colonies during the 19th century. It is well known that CBPP was exported from Europe to Southern Africa [41], Australia [15, 28], and India [36] at that time. The European strain linked to this group, PO1967, is the oldest strain from our collection, which was isolated in 1967 on the French-Spanish border. This strain may represent the closest relative to the ancestral strains that circulated in Europe in the 19th century. More recent European strains differ from PO1967 notably by a deletion of 8.8 kbp, which includes genes encoding a glycerol transporter that may be associated with *Mmm*SC virulence [30, 44]. Unfortunately *Mmm*SC strains circulating prior to 1967 in Europe are not available. They could have been used to monitor the genetic drift and to check if they were genetically closer to *Mmm*SC strains of Southern Africa.

Recent strains of European origin displayed peculiar alleles that were not observed in non-European strains, notably on Loc-PG1-0233 which was not retained in the MLSA

scheme. This finding indicates further that recent CBPP outbreaks in Europe were not due to a reintroduction, but were more likely to be a resurgence of the disease. This conclusion should be of major concern to European countries, particularly with regards to the widening of EU borders. The unnoticed persistence of *Mmm*SC in Europe may be explained either by the existence of another, yet unidentified, reservoir or by the circulation of strains of low-pathogenicity, which may sometimes regain virulence. The presence of *Mmm*SC strains in hosts other than cattle and domestic buffaloes has already been described. This is notably the case of goats, from which *Mmm*SC strains have been regularly isolated both in Africa and in Europe and in which lesions of CBPP can sometimes be reproduced [13]. In an attempt to examine whether goats may be used in place of cattle as models for CBPP, various inoculation routes were tested by Yaya et al. [51]. Some inoculated goats seroconverted and one animal, inoculated intra-peritoneally, developed lesions. However, the role of goats in the natural epidemiology of CBPP is certainly marginal, if not completely irrelevant. This was evidenced in Botswana in 1994, when this country decided to slaughter the entire cattle population from the infected zone [27]. Although the goat population was not targeted in the stamping-out campaign, the subsequent reintroduction of naive cattle in this country was not followed by CBPP outbreaks, hence showing that goats did not play a significant role in CBPP persistence and long term transmission, at least in this context.

On the contrary, the unnoticed circulation of low-pathogenic *Mmm*SC strains in Europe remains a possible explanation for the resurgence of CBPP. It is well known that *Mmm*SC strains show various degrees of virulence. However, strain virulence is difficult to measure since (i) there is no small animal model that can mimic the lesions observed in cattle, (ii) endobronchial inoculations in cattle are not always successful, even when using strains that have shown to be highly virulent in the field, and (iii) in-contact transmission trials are usually carried out with a reduced number of animals, which may not guarantee the evaluation of the strain virulence in a single experiment. This last point has been verified *in silico* with stochastic transmission models showing that successive trials performed with the same initial parameters may lead to very different outcomes [21].

MLST is an unambiguous procedure for characterising isolates of bacterial species based on the sequences of the internal fragments of house-keeping genes [23, 37]. The sequences of approximately 500 bp are normally used, as these can be accurately sequenced on both strands using an automated DNA sequencer. For each house-keeping gene, the different sequences present within a bacterial species are assigned as distinct alleles and each isolate is characterised by an allelic profile, corresponding to the alleles defined in each of the loci. In MLST analyses the number of nucleotide differences found between alleles is disregarded and sequences are given different allele numbers whether they differ at a single or at many nucleotide positions. The rationale for this is that a single genetic event resulting in a new allele can occur by a point mutation, altering only a single nucleotide site, or by a recombinational replacement that will often result in the modification of multiple sites. This typing scheme has been applied successfully to a wide variety of bacterial pathogens such as *Staphylococcus aureus* [14], *Campylobacter jejuni* [12], and *Haemophilus influenzae* [26], to cite only a few. This approach has multiple applications such as phylogeny [18], molecular structure analysis [42], and molecular epidemiology. In molecular epidemiology, MLST can be used to study the evolution of antibiotic resistant strains [17], the temporal trends in strain expansion [29], or the distribution of strains of various lineages within a population [20].

In the case of *Mmm*SC, this approach was not successful, since the variability within housekeeping genes was too limited, if present at all. This was the case of the genes *fusA*, *lepA*, and *rpoB*. In terms of an evolutionary perspective this may indicate that *Mmm*SC genomes are extremely homogeneous, suggesting that this mycoplasma has adapted very recently to its bovine host. From a practical

point of view, this means that genes of unknown function or non coding sequences had to be selected in this study to differentiate *Mmm*SC strains. This is why the strategy has alternatively been named "multilocus sequence analysis". This typing scheme is less universal than MLST analyses, since the gene targets used were selected for typing strains within a biotype and would not be suitable for typing strains of higher taxonomic ranks. However, like MLST, MLSA has the advantage of being a portable and very robust approach.

This new *Mmm*SC typing scheme proved more robust than the previous one developed in our laboratory [22]. No double bands were obtained after PCR amplification, which ensured that PCR products could be sent without any purification step. This was sometimes the case with the previous MLSA. On the contrary, the new scheme allowed the description of 31 profiles instead of only 15 out of a comparable number of strains (51 and 48 respectively).

MLSA analysis may open new perspectives for a better understanding of *Mmm*SC genome plasticity. In this respect, it would be very interesting to obtain ancestral DNA from samples of CBPP-affected lungs preserved in formalin for histology. This should allow the characterisation of the ancestral allelic profiles circulating in Europe and Africa and may permit measuring evolution rates for the various loci.

This MLSA strategy may also be applied to better understand CBPP epidemiology, improving thus the surveillance and control of the disease. New loci may be added to increase the discriminating power of the MLSA tool on subsets of strains that are very closely related. Including a large number of strains isolated from goats in various regions of the world would be very interesting in order to examine whether the MLSA profiles of *Mmm*SC strains from goat origin are similar or, on the contrary, more closely related to other *Mmm*SC strains circulating in the same region. The fact that caprine species may serve as a reservoir should be an incentive to multiply mycoplasma isolation and identification from this host, which is known to harbour mycoplasmas in the ear canal in the absence of pathological signs. Such a procedure is already in place in France, conducted by an epidemiological network called VIGIMYC, managed by the *Agence Française de Sécurité Sanitaire des Aliments* in Lyon.

The results obtained in this study will be included in a web tool dedicated to molecular epidemiology within an EU-funded project called EPIZONE. This tool should facilitate the dissemination of typing methods for pathogens threatening Europe, allow any laboratory to compare in-house data with large databank sets and, in turn, generate phylogenetic trees and actualised maps, provided that the geographic coordinates of all isolates are known. Such a tool should naturally be linked to the websites of the OIE (http://www.oie.int/wahid-prod/) and FAO (http://www.fao.org/ag/againfo/programmes/en/empres/disease_cbpp.asp), which provide updated information on new outbreaks for notifiable diseases.

Finally, this new MLSA tool will be particularly useful to all countries at risk of CBPP reintroduction or resurgence and to those engaged in eradication campaigns, allowing them to trace back the origin of any remaining CBPP focus.

Acknowledgements. We are indebted to a number of projects that made this study feasible: the AU/IBAR CBPP PACE research project, funded by the European Union (REG/5007/005), the LABOVET PSF project, funded by the French Ministry of Foreign Affairs and the EPIZONE Rex project, funded by the European Union. We are particularly grateful to the French Embassy in Cameroon for supporting the PhD thesis work of Yaya Aboubakar.

REFERENCES

[1] Amanfu W., Masupu K.V., Adom E.K., Raborokgwe M.V., Bashiruddin J.B., An outbreak of contagious bovine pleuropneumonia in Ngamiland district of north-western Botswana, Vet. Rec. (1998) 143:46–48.

[2] Angelini S., Deitermann S., Koch H.G., *FtsY*, the bacterial signal-recognition particle receptor, interacts functionally and physically with the SecYEG translocon, EMBO Rep. (2005) 6:476–481.

[3] Bischof D.F., Vilei E.M., Frey J., Genomic differences between type strain PG1 and field strains of *Mycoplasma mycoides* subsp. *mycoides* small-colony type, Genomics (2006) 88:633–641.

[4] Blancou J., Early methods of surveillance and control for contagious bovine pleuropneumonia, Rev. Sci. Tech. (1996) 15:1241–1282.

[5] Brandao E., Isolation and identification of *Mycoplasma mycoides* subspecies *mycoides* SC strains in sheep and goats, Vet. Rec. (1995) 136:98–99.

[6] Campbell A.D., Contagious bovine pleuropneumonia. A report on the use of new antigens for the complement fixation and agglutiantion tests, J. Conc. Sci. Ind. Res. (1938) 11:112–118.

[7] Cheng X., Nicolet J., Poumarat F., Regalla J., Thiaucourt F., Frey J., Insertion element IS1296 in *Mycoplasma mycoides* subsp. *mycoides* small colony identifies a European clonal line distinct from African and Australian strains, Microbiology (1995) 141:3221–3228.

[8] Clay A.L., Lloyd L.C., The eradication of CBPP from Australia, Bull. Off. Int. Epizoot. (1974) 91:533–546.

[9] Costas M., Leach R.H., Mitchelmore D.L., Numerical analysis of PAGE protein patterns and the taxonomic relationships within the '*Mycoplasma mycoides* cluster', J. Gen. Microbiol. (1987) 133:3319–3329.

[10] Cottew G.S., Breard A., DaMassa A.J., Erno H., Leach R.H., Lefevre P.C., Rodwell A.W., Smith G.R., Taxonomy of the *Mycoplasma mycoides* cluster, Isr. J. Med. Sci. (1987) 23:632–635.

[11] Curasson G., Péripneumonie bovine, in: Traité de pathologie exotique vétérinaire et comparée, Vigot Frères (Ed.), Paris, France, 1942, pp. 276–353.

[12] Dingle K.E., Colles F.M., Wareing D.R., Ure R., Fox A.J., Bolton F.E., Bootsma H.J., Willems R.J., Urwin R., Maiden M.C., Multilocus sequence typing system for *Campylobacter jejun*i, J. Clin. Microbiol. (2001) 39:14–23.

[13] Dujardin-Beaumetz E., Transmission de la péripneumonie des bovidés aux espèces ovine et caprine, Ann. Inst. Pasteur (Paris) (1906) 20:449–466.

[14] Enright M.C., Day N.P.J., Davies C.E., Peacock S.J., Spratt B.G., Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*, J. Clin. Microbiol. (2000) 38:1008–1015.

[15] Fisher J., The origins, spread and disappearance of contagious bovine pleuro-pneumonia in New Zealand, Aust. Vet. J. (2006) 84:439–444.

[16] Gambles R.M., Studies on the contagious bovine pleuropneumonia with special reference to the complement fixation test, Br. Vet. J. (1956) 112:34-40, 78-86, 120-127, 162–169.

[17] Gherardi G., Fallico L., Del Grosso M., Bonanni F., D'Ambrosio F., Manganelli R., Palu G., Dicuonzo G., Pantosti A., Antibiotic-resistant invasive pneumococcal clones in Italy, J. Clin. Microbiol. (2006) 45:306–312.

[18] Hanage W.P., Kaijalainen T., Herva E., Saukkoriipi A., Syrjanen R., Spratt B.G., Using multilocus sequence data to define the *Pneumococcus*, J. Bacteriol. (2005) 187:6223–6230.

[19] Hozier H.M., British expedition to Abyssinia, MacMillan (Ed.), London, UK, 1869, p. 271.

[20] Lacher D.W., Steinsland H., Blank T.E., Donnenberg M.S., Whittam T.S., Molecular evolution of typical enteropathogenic *Escherichia coli*: clonal analysis by multilocus sequence typing and virulence gene allelic profiling, J. Bacteriol. (2007) 189:342–350.

[21] Lesnoff M., Laval G., Bonnet P., Abdicho S., Workalemahu A., Kifle D., Peyraud A., Lancelot R., Thiaucourt F., Within-herd spread of contagious bovine pleuropneumonia in Ethiopian highlands, Prev. Vet. Med. (2004) 64:27–40.

[22] Lorenzon S., Arzul I., Peyraud A., Hendrikx P., Thiaucourt F., Molecular epidemiology of contagious bovine pleuropneumonia by multilocus sequence analysis of *Mycoplasma mycoides* subspecies *mycoides* biotype SC strains, Vet Microbiol. (2003) 93:319–333.

[23] Maiden M.C.J., Bygraves J.A., Feil E., Morelli G., Russell J.E., Urwin R., Zhang Q., Zhou J., Zurth K., Caugant D.A., Feavers I.M., Achtman M., Spratt B.G., Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms, Proc. Natl. Acad. Sci. USA (1998) 95:3140–3145.

[24] March J.B., Clark J., Brodlie M., Characterization of strains of *Mycoplasma mycoides* subsp. *mycoides* small colony type isolated from recent outbreaks of contagious bovine pleuropneumonia in Botswana and Tanzania: evidence for a new biotype, J. Clin. Microbiol. (2000) 38:1419–1425.

[25] Masiga W.N., Domenech J., Windsor R.S., Manifestation and epidemiology of contagious bovine pleuropneumonia in Africa, Rev.-Off. Int. Epizoot (1996) 15:1283–1308.

[26] Meats E., Feil E.J., Stringer S., Cody A.J., Goldstein R., Kroll J.S., Popovic T., Spratt B.G., Characterization of encapsulated and noncapsulated *Haemophilus influenzae* and determination of phylogenetic relationships by multilocus sequence typing, J. Clin. Microbiol. (2003) 41:1623–1636.

[27] Mullins G.R., Fidzani B., Kolanyane M., At the end of the day. The socioeconomic impacts of eradicating contagious bovine pleuropneumonia from Botswana, Ann. N.Y. Acad. Sci. (2000) 916:333–344.

[28] Newton L.G., Contagious bovine pleuropneumonia in Australia: some historic highlights from entry to eradication, Aust. Vet. J. (1992) 69:306–317.

[29] Perez-Losada M., Crandall K.A., Zenilman J., Viscidi R.P., Temporal trends in gonococcal population genetics in a high prevalence urban community, Infect. Genet. Evol. (2007) 7:271–278.

[30] Pilo P., Vilei E.M., Peterhans E., Bonvin-Klotz L., Stoffel M.H., Dobbelaere D., Frey J., A metabolic enzyme as a primary virulence factor of *Mycoplasma mycoides* subsp. *mycoides* small colony, J. Bacteriol. (2005) 187:6824–6831.

[31] Poumarat F., Solsona M., Molecular epidemiology of *Mycoplasma mycoides* subsp. *mycoides* biotype small colony, the agent of contagious bovine pleuropneumonia, Vet. Microbiol. (1995) 47:305–315.

[32] Provost A., Prophylaxis and vaccination in bovine pleuropneumonia. Evolution of technics and their current practical applications, Rev. Elev. Med. Vet. Pays Trop. (1974) 27:145–161.

[33] Provost A., Perreau P., Breard A., Le Goff C., Martel J.L., Cottew G.S., Contagious bovine pleuropneumonia, Rev.-Off. Int. Epizoot.(1987) 6:625–679.

[34] Regalla J., Caporale V., Giovannini A., Santini F., Martel J.L., Goncalves A.P., Manifestation and epidemiology of contagious bovine pleuropneumonia in Europe, Rev.-Off. In. Epizoot. (1996) 15:1309–1329.

[35] Rochebrune A.T. (de), Formation de races nouvelles. Recherche d'ostéologie comparée, sur une race de boeufs domestiques observée en Sénégambie, C. R. Acad. Sci. (Paris) (1880) 91:304–306.

[36] Shirlaw J.F., Observations on the existence of contagious bovine pleuropneumonia in British India, with an account of preliminary pathological investigation of cases of this disease reported from Assam, Indian J. Vet. Sci. (1939) 9:139–150.

[37] Spratt B.G., Multilocus sequence typing: molecular typing of bacterial pathogens in an era of rapid DNA sequencing and the Internet, Curr. Opin. Microbiol. (1999) 2:312–316.

[38] Srivastava N.C., Thiaucourt F., Singh V.P., Sunder J., Isolation of *Mycoplasma mycoides* small colony type from contagious caprine pleuropneumonia in India, Vet. Rec. (2000) 147:520–521.

[39] Thiaucourt F., Di Maria A., A new microtitration method for the enumeration of contagious bovine pleuropneumonia (CBPP) vaccines, Biologicals (1992) 20:11–13.

[40] Thiaucourt F., Lorenzon S., David A., Tulasne J.J., Domenech J., Vaccination against contagious bovine pleuropneumonia and the use of molecular tools in epidemiology, Ann. NY Acad. Sci. (1998) 849:146-151.

[41] Thiaucourt F., Van der Lugt J.J., Provost A., Contagious bovine pleuropneumonia, in: Infectious diseases of livestock., J.A. Coetzer, R.C. Tustin (Eds.), Oxford University Press Southern Africa, Cape Town, pp. 2045–2059.

[42] Tourasse N.J., Helgason E., Okstad O.A., Hegna I.K., Kolstø A.B., The *Bacillus cereus* group: novel aspects of population structure and genome dynamics, J. Appl. Microbiol. (2006) 101:579–593.

[43] Vilei E.M., Abdo E.M., Nicolet J., Botelho A., Goncalves R., Frey J., Genomic and antigenic differences between the european and African/Australian clusters of *Mycoplasma mycoides* subsp. *mycoides* SC, Microbiology (2000) 146:477–486.

[44] Vilei E.M., Frey J., Genetic and biochemical characterization of glycerol uptake in *Mycoplasma mycoides* subsp. *mycoides* SC: its impact on H_2O_2 production and virulence, Clin. Diagn. Lab. Immunol. (2001) 8:85–92.

[45] Vilei E.M., Korczak B.M., Frey J., *Mycoplasma mycoides* subsp. *capri* and *Mycoplasma mycoides* subsp. *mycoides* LC can be grouped into a single subspecies, Vet. Res. (2006) 37:779–790.

[46] Vilei E.M., Nicolet J., Frey J., IS1634, a novel insertion element creating long, variable-length direct repeats which is specific for *Mycoplasma mycoides* subsp. *mycoides* small-colony type, J. Bacteriol. (1999) 181:1319–1323.

[47] Wesonga H., Thiaucourt F., Experimental studies on the efficacy of T1sr and T1/44 vaccine strains of *Mycoplasma mycoides* subsp. *mycoides* SC against a field isolate causing contagious bovine pleuropneumonia in Kenya- effect of a revaccination, Rev. Elev. Med. Vet. Pays Trop. (2000) 53:313–318.

[48] Westberg J., Persson A., Holmberg A., Goesmann A., Lundeberg J., Johansson K.E., Pettersson B., Uhlen M., The genome sequence of *Mycoplasma mycoides* subsp. *mycoides* SC type strain PG1T, the causative agent of contagious bovine pleuropneumonia (CBPP), Genome Res. (2004) 14:221–227.

[49] Willems L., Mémoires sur la pleuropneumonie épizootique du gros bétail, Recl. Méd. Vét. Pratique, Maisons Alfort, France, (1852) 3:401–434.

[50] Yaya A., Golsia R., Hamadou B., Amaro A., Thiaucourt F., Essai comparatif d'efficacité de deux souches vaccinales T1/44 et T1sr contre la péripneumonie contagieuse bovine, Rev. Elev. Med. Vet. Pays Trop. (1999) 52:171-179.

[51] Yaya A., Hamadou B., Yaya D., Abdoulkadiri S., Thiaucourt F., Inoculation expérimentale de l'agent de la péripneumonie contagieuse bovine à des chèvres, Rev. Elev. Med. Vet. Pays Trop. (2000) 53:319–324.