



OPEN

## Multilocus sequence analysis reveals genetic diversity in *Staphylococcus aureus* isolate of goat with mastitis persistent after treatment with enrofloxacin

Richard Costa Polveiro<sup>1</sup>, Manuela Maria Cavalcante Granja<sup>1</sup>, Thais Coimbra Borba Roldão<sup>1</sup>, Ilderlane Da Silva Lopes<sup>1</sup>, Pedro Marcus Pereira Vidigal<sup>2</sup>, Magna Coroa Lima<sup>1</sup> & Maria Aparecida Scatamburlo Moreira<sup>1</sup>✉

*Staphylococcus aureus* is one of the main bacterial agents responsible for cases of mastitis in ruminants, playing an important role in the persistence and chronicity of diseases treated with antimicrobials. Using the multilocus sequence typing technique, network approaches and study of the population diversity of microorganisms, we performed analyzes of *S. aureus* (ES-GPM) isolated from goats with persistent mastitis (GPM). The most strains of ES-GPM were categorically different phylogenetically from the others and could be divided into two lineages: one with a majority belonging to ES-GPM and the other to varied strains. These two lineages were separated by 27 nuclear polymorphisms. The 43 strains comprised 22 clonal complexes (CCs), of which the ES-GPM strains were present in CC133, CC5 and a new complex formed by the sequence type 4966. The genetic diversity of some alleles showed be greater diversity and polymorphism than others, such as of the *aroE* and *yqiL* genes less than *glpF* gene. In addition, the sequences ES-GPM to the *arc* gene and *glpF* alleles showed the greatest number of mutations for ES-GPM in relation to non-ES-GPM. Therefore, this study identified genetic polymorphisms characteristic of *S. aureus* isolated from milk of goats diagnosed with persistent mastitis after the failed treatment with the antibiotic enrofloxacin. This study may help in the future to identify and discriminate this agent in cases of mastitis, and with that, the most appropriate antibiotic treatment can be performed in advance of the appearance of persistent mastitis caused by the agent, reducing the chances of premature culling and animal suffering.

*Staphylococcus aureus* is a major opportunistic pathogen in humans and one of the most important pathogenic *Staphylococcus* species in veterinary medicine<sup>1</sup>. It appears on the world stage as a challenge in disease chronicity and persistence in antimicrobial treatments<sup>2,3</sup>. *S. aureus* is part of a small group of bacteria that can be named as 'persisters' in some diseases<sup>4</sup>, since administration of antibiotic treatments based on correct protocols allows the survival of a distinct form of bacterial subpopulation in an infection<sup>5</sup>. This bacterial agent is one of the most commonly found pathogens in clinical mastitis in several livestock species<sup>1</sup>, such as goats<sup>6</sup>. However, it can cause subclinical, clinical, persistent, gangrenous and difficult-to-treat intramammary infections<sup>7</sup>.

This type of persistent infection could also be identified in the types of mastitis, but it is often just referred to as chronic<sup>8</sup>, and not as an antimicrobial persistence, which can results in the cure failure due to the performance of only the exchange of the antimicrobial, and consequently in premature culling or replacement of incurable animals. In chronic goat's mastitis cases, the mammary gland can atrophy due to severe damage to the glandular tissue leading to fibrosis and total loss of the ability to produce milk<sup>9</sup> and can result in the disposal of the animal.

<sup>1</sup>Laboratory of Bacterial Diseases, Sector of Preventive Veterinary Medicine and Public Health, Veterinary Department, Universidade Federal de Viçosa, Viçosa, Minas Gerais 336570-900, Brazil. <sup>2</sup>Núcleo de Análise de Biomoléculas (NuBioMol), Center of Biological Sciences, Universidade Federal de Viçosa, Viçosa, Minas Gerais 336570-900, Brazil. ✉email: [masm@ufv.br](mailto:masm@ufv.br)

Persistent infections are typically multifactorial and may have mechanisms for different classes of antibiotics, such as  $\beta$ -lactams and fluoroquinolones<sup>10</sup>, and mainly occurs due to the selection of different pathogens capable of evading the immune system<sup>5</sup>, which are genetically prepared to activate the series of stress responses for long-term survival<sup>4</sup>. Persistent mastitis in ruminants<sup>11,12</sup> raises concerns because it renders ineffective treatments with antibiotics such as enrofloxacin<sup>12</sup>. Enrofloxacin is a fluoroquinolone exclusively developed for use in veterinary medicine<sup>13</sup> and presents itself as an appropriate choice in some cases of mastitis in goats<sup>14</sup>, being active against major pathogenic bacteria (both Gram-positive and Gram-negative)<sup>15</sup>. Therefore, despite the best possible antimicrobial treatment, failures of bacteriological cure can often occur in mastitis treatments, especially for *S. aureus* mastitis, and antimicrobial resistance is considered one of the reasons for low cure rates<sup>16</sup>.

Consequently, identifying the nuclear polymorphisms of *S. aureus* to characterize it as persistent during mastitis may be a solution in provide the barcode that can be used to track and identify strains with the same or similar polymorphisms, and thus increase the chances of providing a treatment appropriate antimicrobial. Thus, multi-locus sequence typing (MLST), which use phylogenetic procedures on the nucleotide sequences of allelic locus used in MLST are being widely used for identification and determination of phylogenetic relationships between isolates<sup>17</sup>. On the other hand, the haplotype network presents itself as an alternative representation of the clonal genealogy of bacterial isolates<sup>18</sup> and complementary to the phylogeography of *S. aureus*. In addition, reconstructing bacterial haplotypes is required when choosing the right treatments for diseases caused by specific haplotypes in a population<sup>19</sup>.

Population analyses of bacteria based on data from different loci using MLST schemes can provide important epidemiological results, especially if confronted with different methods involving genealogical typing<sup>20</sup>. Therefore, it was the aim of this study is to investigate the genetic diversity of *S. aureus* (ES-GPM) isolated from goats with persistent mastitis (GPM) through MLST scheme and compare with other sequence of *S. aureus* isolated from different other types of mastitis (non-ES-GPM) and hosts that are deposited in the PubMLST database<sup>21</sup>, using phylogeographical, phylogenetic, and diversity approaches.

## Results

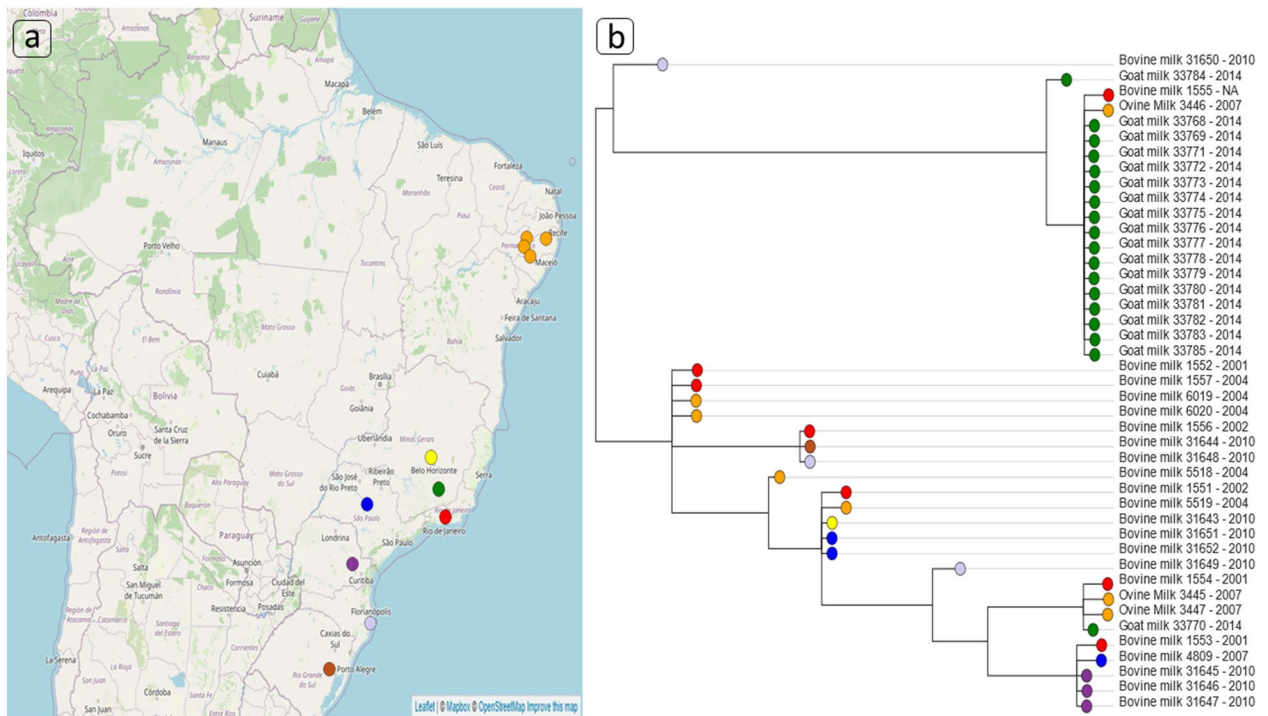
A total of 43 isolates were analyzed in this study. Within these 43, 18 (ES-GPM) isolates from the milk of animals were diagnosed with GPM at Capril UFV, named Minas Gerai—LDBAC (Table S1). In addition, another 25 isolates of animals diagnosed with mastitis were selected from the PubMLST database (<https://pubmlst.org/organisms/staphylococcus-aureus/>); these are shown in supplementary Table S3. As a result of this search of the database, 22 bovine mastitis and three ovine isolates were referenced. The total sequences obtained originated from seven different Brazilian states: Rio de Janeiro (n = 7), Pernambuco (n = 7), São Paulo (n = 7), Minas Gerais (n = 1), Rio Grande do Sul (n = 1), Paraná (n = 3), Santa Catarina (n = 3) and Minas Gerais-LDBAC (n = 18) regarding the ES-GPM sequences (Fig. 1a and b). These analyzed sequences represent all the sequences deposited at (<https://pubmlst.org/organisms/staphylococcus-aureus/>) referring to animal mastitis in Brazil up to 09/28/2019. The MICs ranged from 0.125 to 16  $\mu\text{g}/\text{mL}$  in the isolates prior to treatment, whereas in the isolates after treatment with enrofloxacin antibiotic, the antimicrobial MICs ranged from 0.19 to 16  $\mu\text{g}/\text{mL}$ , with only samples of IDs 33784 and 33785 showing 16  $\mu\text{g}/\text{mL}$ , as previously reported<sup>22</sup>.

**Phylogeography, STs and CCs.** In Fig. 1, the phylogeographic distribution of *S. aureus* isolates are characterized, relating the phylogeny of the strains to the isolation sites of the sample, which showed little association between the origin of the isolate and the phylogenetic structure found. Despite the apparent small number of sequences as shown in Fig. 1, these MLST data represent the totality of Brazilian sequences deposited from *S. aureus* that infected dairy animals. However, strains related to ovine and caprine mastitis apparently had isolates from bovine mastitis as common ancestors. In addition, ES-GPM-related strains were highly associated in clades or clusters, which may be because few different strains causing GPM to circulate among animals within the same goat herd.

On the other hand, strains related to bovine mastitis demonstrated phylogenetic ramifications along the branches of the phylogenetic tree (Fig. 1), regardless of the Brazilian region studied, which implies greater nucleotide polymorphism. Strains related to ES-GPM are apparently little related phylogenetically to the other two host species (ovine and bovines) and show no clear phylogenetic distinction, as can be seen by strains from bovine ID 1555 and ovine ID 3446 (Fig. 1). However, even though most of the ES-GPM-related strains remained close phylogenetically, the strain ID 33770 (ST 5; Fig. 1) distanced itself from the other ES-GPM strains, becoming close to those from ovine and bovines. In addition, IDS strains 33784 and 33770 were harvested prior to treatment with enrofloxacin as listed in Table S1, and this possibly determined that they were phylogenetically more distant from other ES-GPM strain and that also presented as a different ST.

The corresponding phylogenetic analyses (Fig. 2a,b) showed high consistency with the Multilocus Sequence Tree (MS tree) (Fig. 2c), providing better resolution when considering the seven (AT) genes characterized for *S. aureus* (Fig. 2a,b) than if a single housekeeping gene were used. In addition, these analyses also presented robust topology with high bootstrap values, dividing the studied population into two lineages (Fig. 2a). The phylogenetic tree based on housekeeping genes or concatenated maintenance genes, as shown in Fig. 2b, indicated moderate differentiation among populations, as also shown in Fig. 1b. Nevertheless, the 43 strains were divided into two clusters, called Lineage 1 (L1) and Lineage 2 (L2), being characterized by four CCs and another 18 CCs, respectively.

Accordingly, each lineage cluster (Fig. 2a) represents one or more phylogenetic associations that, concomitantly, also determined the formation of CCs in Fig. 2c, and thus 20 STs with their concatenated sequences formed L1, and 23 formed L2. The high nucleotide identity caused L1 to form a smaller cluster of CCs (Fig. 2) compared to L2 (Fig. 2).



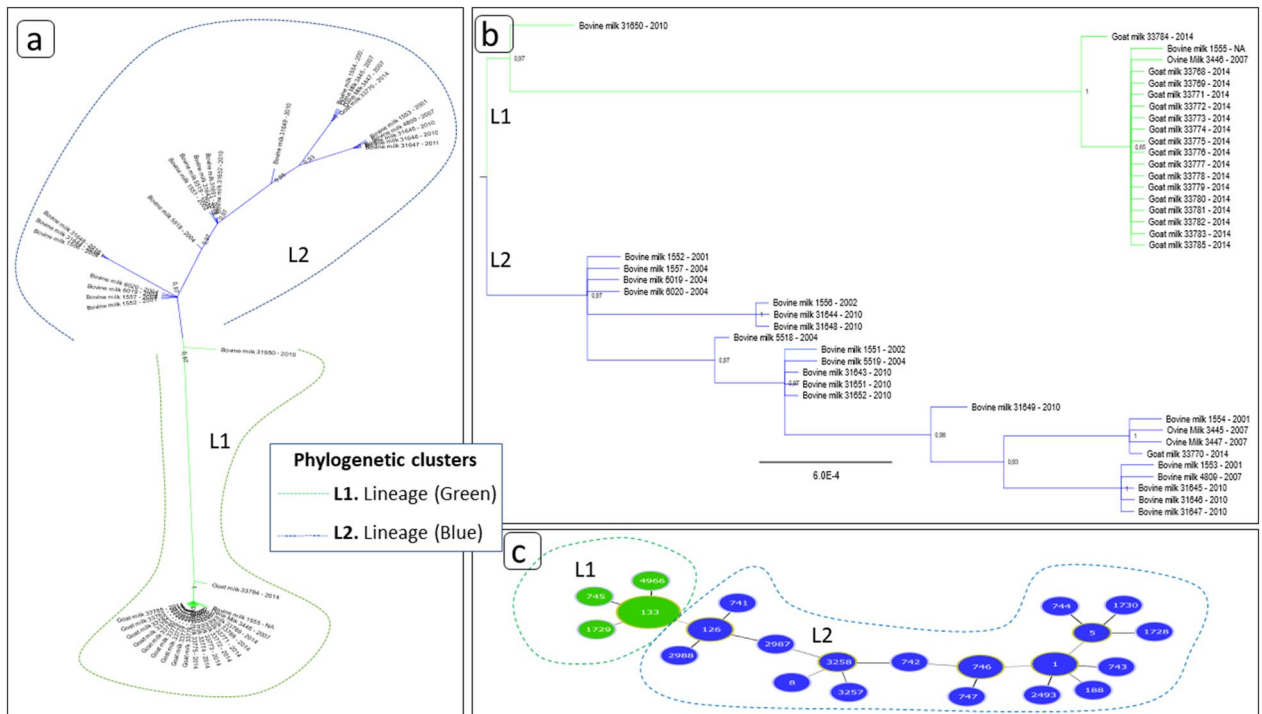
**Figure 1.** Geographic distribution, geospatial information, and phylogenetic analysis of isolates of *Staphylococcus aureus* in milk from Brazilian herds. **(a)** Geographic distribution of the 43 isolates among the seven Brazilian states. Each dot on the map represents the origin location of mastitic milk: Orange—Pernambuco (Northeast / BR), yellow—Minas Gerais (Southeast / BR), and green (ES-GPM)—Minas Gerais (Southeast / BR)—Minas Gerais (Southeast / BR), red—Rio de Janeiro (Southeast / BR), blue—São Paulo (Southeast / BR), purple—Paraná (South / BR), gray—Santa Catarina (South / BR), brown—Rio Grande do Sul (South / BR). Created with Microreact version 5.93.0 (at <http://microreact.org>). **(b)** Phylogenetic tree structured with metadata in Microreact, with the colors referenced to each Brazilian state analyzed. The tip labels were detailed with the source of the samples, identification of the isolates (ID) and lastly this is the year in which it was collected and analyzed. The tree was visualized using Figtree software and the fully interactive version of our phylogenetic tree and including geospatial information can be found at <https://microreact.org/project/puA-MDz8o>. Created with Microreact version 5.93.0 (at ) and by Figtree software version 1.4.4.

Among the 22 isolates from bovines, 16 different STs were identified, while for the three ovine isolates, three STs were identified. Of the 18 goat isolates from LDBAC-UFV, three STs were also identified, which totaled 22 STs. These isolates were characterized in 22 CCs as shown in Fig. 2c. Other information, such as the AT variation of each of the seven housekeeping that characterize ST, can be found in Table S3.

A larger number of smaller-sized blue CC circles are shown in Fig. 3a, which demonstrates a greater variability between the ATs of *S. aureus* STs that cause bovine mastitis. On the other hand, the STs of the same bacteriological agent that affects goats and ovine are at opposite ends of the MS tree, and in some cases with CCs of larger size, showing less variability and clustering of STs by several isolates. Among these, we highlight CC133, CC5 and CC4966 in green in Fig. 3b, which correspond to GPM and ES-GPM. In addition, the fact that the STs of goat and ovine ES-GPM are positioned at two opposite points of the MS tree and connected by the CCs in blue (bovine samples), can demonstrate that the colonization of *S. aureus* in ovine and goats, and ES-GPM, are genetic conditions that arose from bovine strains at different times, determining the same pathogenic condition.

In the MS tree in Fig. 3b, the description of the CCs can be observed in relation to the states of origin of the samples. Some states showed the greatest variability in STs and, because of that, larger CCs as well, such as the state of Rio de Janeiro (red) and Pernambuco (orange), which each held seven different STs. The CC746, of IDs 1556, 31644 and 31648 (from bovine samples—Table S3), was the most present in the largest number of states, such as in Santa Catarina, Rio de Janeiro and Rio Grande do Sul, and therefore CC126, of IDs 31643, 31651 and 31652, were present in the states of Minas Gerais and São Paulo.

**Haplotypic network of *S. aureus* for source animal and geographical distribution.** In this analysis, with the concatenation of the seven MLST genes from *S. aureus*, the 43 sequences were grouped into 22 haplotypes (Table 1). The H20 haplotype (Figs. 4 and 5), referring to the ES-GPM sequences, was the point in the network that presented the greatest number of aggregated samples, being exclusively of goats, and thus of the ES-GPM from Minas Gerais. The bovines showed a greater number of different haplotypes (72.73%), followed by goats (13.64%) and ovine (13.64%; Fig. 4). The haplotypic diversity (Hd) of the broad population of *S. aureus* (seven MLST housekeeping genes) was higher (0.8571; Table 1), represented by the high number of haplotypes



**Figure 2.** Phylogenetic tree and MS tree. (a). Tree and phylogenetic structure in line with Multilocus Sequence Tree (MLST tree) subdivided phylogenetically into clades (L1 and L2). Created with Figtree software version 1.4.4. | (b). Phylogenetic tree represented in a cladogram with more information on the isolates, such as: sample source, sample identification (ID) and year of sample collection. Created with Figtree software version 1.4.4. | (c). MS tree. The MS tree of the 22 STs was inferred by PhyloViz on the basis of allelic profiles. Each circle indicated a ST (node), and a larger size of the circle corresponded to a larger number of strains included. The dash lines separated two lineages L1 to L2. Minimum spanning tree reflects the clonal relationships of the sequence types (STs) for 43 *Staphylococcus aureus* bacteria subdivided between two strains (L1/green and L2/blue) constructed using goeBURST characterized as clonal complex (CCs). The background colors indicated different CCs. Created with PHYLOViZ software.

formed. The haplotypes referring to the goat and ovine samples show that their ancestors have always belonged to *S. aureus* from bovine isolates (Fig. 4).

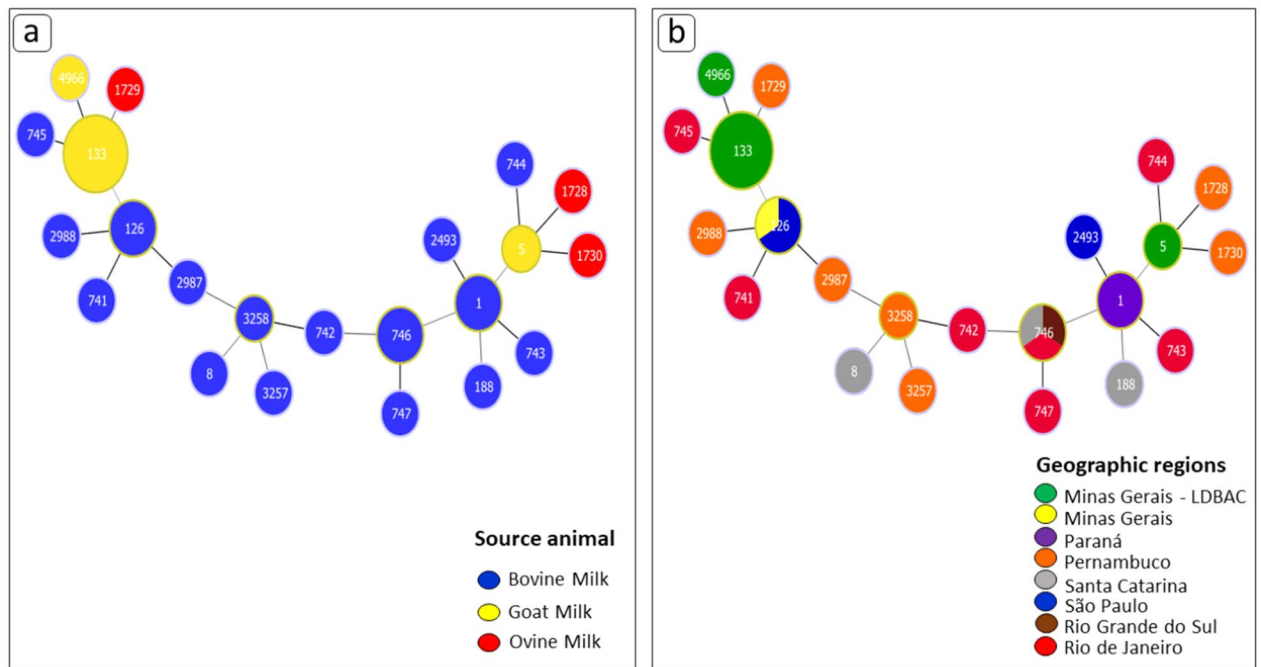
Interestingly, the distribution patterns of CCs found in the MS tree (Fig. 3) were like those of the haplotype networks (Figs. 4 and 5). Thus, H20, H22, H5 and H9 (Fig. 4) corresponded to CC133, CC4966, CC745, CC1729, respectively, while H21, H8, H10 and H4 corresponded to C5, CC1730, CC1728 and CC744. In addition, C133 (Fig. 3) or H20 (Fig. 4) had between 24 and 27 mutations separating them from the other haplotypes, and the groups corresponding to C5 (Fig. 3) or H21 (Fig. 4) were separated from the rest of the network groups by 6 to 7 mutations. However, most haplotypes (72.73%) were separated by one mutation, while the rest (23.81%) were separated by between 2 and 27 mutations.

The high number of mutations presented from H22 established the separation of the groups, and Line 1 (L1; 22.73%) was formed by haplotypes H5, H9, H20 and H22. With this, H22 was positioned exactly where the two different clusters/strains are separated (Figs. 4 and 5), between H22 and H19, a result corresponding to the analysis of Fig. 2b. On the other hand, the H21, H8, H10 and H4 haplotypes were part of the Line 2 (L2) lineage, with H21 or C5 being the only GPM and ES-GPM haplotypes that did not group with L1.

Of the 22 haplotypes found, only 13% were goats and H20 was the most common ( $n = 16$ , 37.2%), formed by ES-GPM collected at UFV, Minas Gerais. The H20 haplotype (Minas Gerais—LDBAC) from goat strains proved to be a common ancestor of the H22 (Minas Gerais—LDBAC), H9 (Pernambuco), and H5 (Rio de Janeiro) haplotypes (Fig. 5). However, the H19 haplotype (Santa Catarina) was the haplotype that differed most in the number of mutations from ancestors and other haplotypes; however, a greater number of samples from this state could help to reduce this isolation and complement the network.

Other haplotypes belong to the lineage L1, H20 and H5 with bovine and ovine sequences, respectively, showed only one mutation point between them. The H19 haplotype, also belonging to this strain, was the most distinct from the others. The L2 line (77.27%) was formed by the other haplotypes, with cattle representing 82.35% of this line (Fig. 4). The number of mutations represented in the haplotypic network of Figs. 4 and 5, suggest that the L1 strain can be highly divergent from the isolates of the L2 strain.

The H6 haplotype, referring to the bovine *S. aureus* sequences, was the haplotype with the greatest dispersion among the states, presenting itself in the states of the southern region (Santa Catarina and Rio Grande do Sul) and in the southeast region (Rio de Janeiro). In addition, H16 presented itself as a common ancestor and divisor of the haplotypic network, branching out several haplotypes, such as H13 and H12 (Pernambuco) and H1



**Figure 3.** The MS tree of the 22 STs was inferred by PhyloViZ based on allelic profiles. Each circle indicated a ST (node) or complex clonal (CC), and a larger size of the circle corresponded to a larger number of strains included, in other words, size of the circle is logarithmically proportional to the number of strains represented by the ST. The background colors indicated different CCs. (a) Each color represents a different host for *Staphylococcus aureus* (bovine, caprine and ovine) obtained from milk of animals with mastitis. The circles with blue colors denote the source of the sample from bovine, yellow from goat and red from ovine. Created with PHYLOViZ software. (b) The isolates are colored according to isolated areas, distinguishing between samples from 7 Brazilian states. The CC referenced in green are from milk of goats with persistent goat mastitis (GPM). Created with PHYLOViZ software.

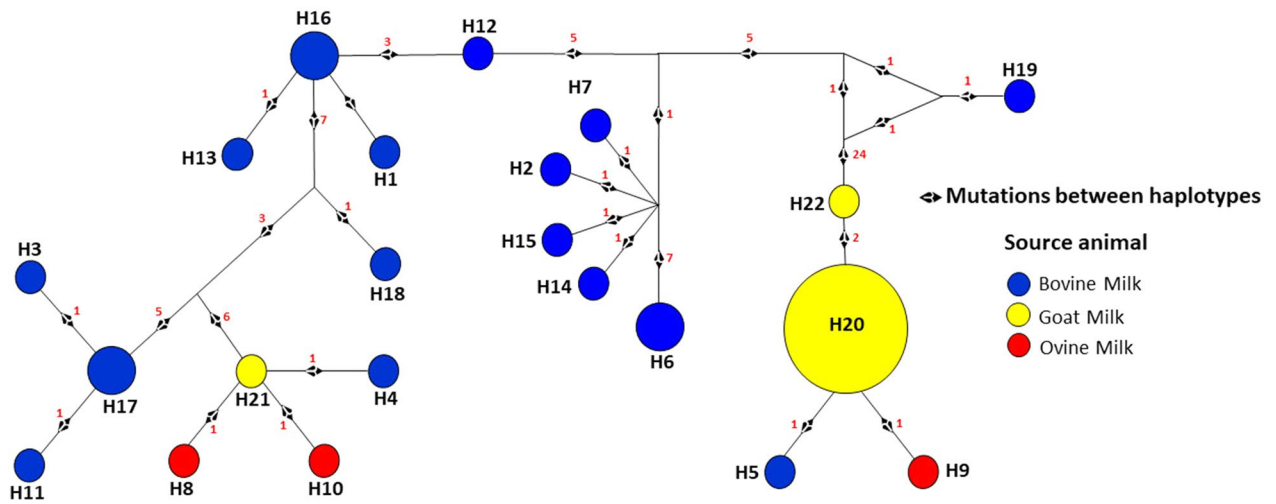
	<i>arcC</i>	<i>aroE</i>	<i>glpF</i>	<i>gmk</i>	<i>pta</i>	<i>tpi</i>	<i>yqiL</i>	Concatenated
$T^a$	456	456	465	417	474	402	516	3186
$N^b$	43	43	43	43	43	43	43	43
$h^c$	5	8	4	5	6	4	10	22
$Hd^d$	0.6844	0.7375	0.5548	0.6844	0.6534	0.6744	0.7907	0.8571
$S^e$	6	13	5	9	6	7	15	61
$S^*$	4	10	3	7	4	7	12	47

**Table 1.** Molecular diversity indices of wide populations by multilocus sequence typing genes of *Staphylococcus aureus*.  $T^a$ : length of DNA sequences (pb);  $N^b$ : number of sequences;  $h^c$ : number of haplotypes;  $Hd^d$ : haplotypic diversity;  $S^e$ : polymorphic sites;  $S^*$  number of parsimony informational sites.

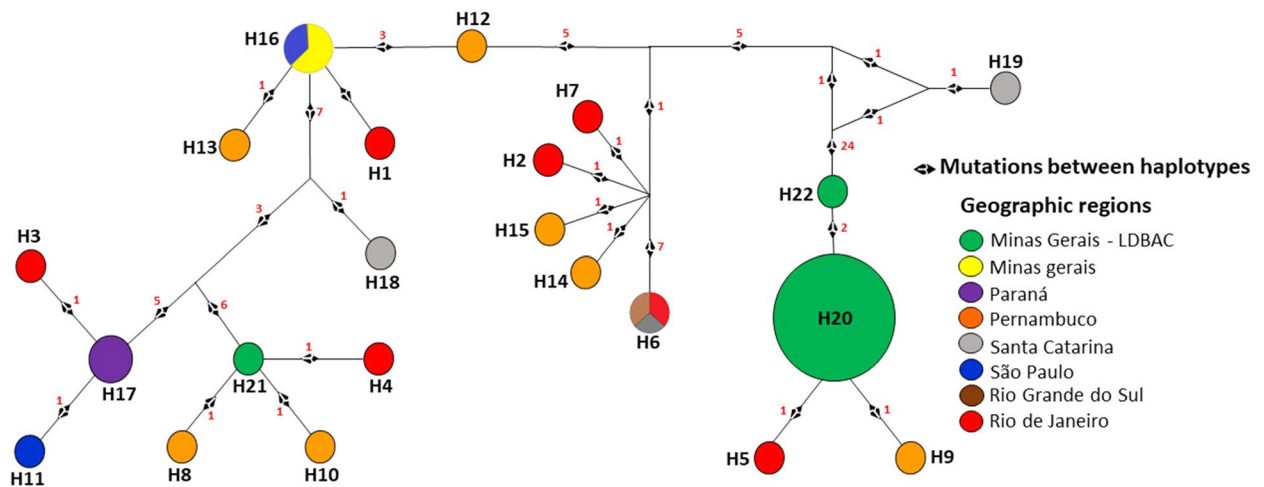
(Rio de Janeiro). In the entire haplotype network referring to the different Brazilian regions (Fig. 5), the states of Rio de Janeiro and Pernambuco were the ones that presented the greatest number of haplotype diversities.

**Haplotypic networks of individual genes by hosts.** The individual analysis of the seven genes of *S. aureus* isolates in the different hosts revealed the formation of 4–10 different haplotypes (Table 1). While the haplotype network analysis of the concatenated housekeeping genes (large population) revealed a greater specificity of the haplotypes for their host, in the analysis of the seven housekeeping genes separately there was predominantly an aggregation of the isolates in a non-host-specific manner (Fig. S1 and S2). The haplotypes referring to milk samples of cow or cattle origin appear to be more specific; at the time, the haplotypes referring to ovine and goats showed greater plurality.

The frequency of different hosts that shared AT of *S. aureus* in the same haplotype was low: 9% on average. Only the H4 haplotype for the *yqiL* gene showed 100%, in which it grouped three different hosts in the same proportion. The *yqiL* and *aroE* genes showed the highest values of haplotypic diversity: 0.7907 and 0.7375, respectively, while *glpF* had the lowest value: 0.5548 (Table 1). The greatest genetic variability observed for the *yqiL* and *aroE* genes, to the detriment of the haplotypic diversity values, demonstrated that these genes are more mutable in relation to *glpF*.



**Figure 4.** Haplotype network of multilocus sequence typing alleles *arcC*, *aroE*, *gmk*, *pta*, *tpi*, *glpF* e *yqiL* of *Staphylococcus aureus* obtained from milk of animals with mastitis. Hosts: Yellow—Caprine; Blue—Bovine; and Red—Sheep. Red numbers between haplotypes—polymorphic sites. The colors separating the haplotypes refer to the groups formed in the MS tree. Constructed with Network 4.6.1.0 software.



**Figure 5.** Haplotype network of multilocus sequence typing alleles *arcC*, *aroE*, *gmk*, *pta*, *tpi*, *glpF* and *yqiL* of *Staphylococcus aureus*, obtained from milk of animals with mastitis and based on the geographic distribution of Brazil. The distribution refers to the states of Minas Gerais—yellow and green for Minas Gerais—LDBAC, Pernambuco—orange, Paraná—purple, Rio de Janeiro—red, Rio Grande do Sul—brown, Santa Catarina—gray and São Paulo—blue. Red numbers between haplotypes—mutations sites. The colors separating the haplotypes refer to the groups formed in the phylogenetic tree and MS tree. Constructed with Network 4.6.1.0 software.

The genetic diversity shown in the *arc*, *tpi*, *pta* and *gmk* genes were very similar (Table 1). The *arc* and *tpi* genes showed a remarkably similar distribution of haplotypes in each host. However, the *pta* and *gmk* genes showed more specific haplotypes for ovines and goats than the other genes shown in supplementary Figure S1 and S2. However, the *pta* gene (Fig. S2) appears to have a greater haplotype specificity for goats and ovines than the other genes.

**Sequence diversity.** In order to assess the general sequence diversity of the seven MLST loci of the isolates under study, the average GC (genomic DNA base composition) content, the number of polymorphic sites, the ratio of non-synonymous (dN) to synonymous (dS) substitutions and nucleotide diversity ( $\pi$ ) were calculated and are shown in Table 2. Sequence alignment of each MLST locus showed no insertion/deletion. Concatenated sequences for the seven loci were 3186 bp in length with a diversity index  $\pi$  of 0.00670. The mean GC content of the MLST gene fragments ranged from 30.10% for *aroE* to 40.9% for *glpF*. The nucleotide diversity index  $\pi$  ranged from 0.00346 (*glpF*) to 0.01119 (*aroE*; Table 2). The number of polymorphic sites varied from 1.07% for *glpF* to 2.93% for *aroE*, the most polymorphic locus (Table 2). Interestingly, the *glpF* sequence showed the lowest haplotypic diversity of all (Table 1) and the highest average G + C content (Table 2).

Locus	Length (bp)	Average G + C content (%)	S/M	No. of polymorphic sites (%)	SNPs	Average dN/dS ratio (%)	No. of nonsense mutation	Nucleotide diversity ( $\pi$ )
<i>arc</i>	456	38.30	6	1.32	2	0.239	1	0.00354
<i>aroE</i>	443	30.10	14	2.93	3	0.432	2	0.01119
<i>glpF</i>	465	40.90	5	1.07	2	0.228	1	0.00346
<i>Gmk</i>	417	33.50	9	2.15	2	0.151	–	0.00762
<i>Pta</i>	474	36.20	6	1.26	2	0.383	–	0.00447
<i>Tpi</i>	402	37.50	7	1.74	0	0.095	–	0.00766
<i>yqiL</i>	516	37.90	15	2.90	3	0.173	2	0.00900
Concatenated	3186	36.40	62	1.94	14	0.261	6	0.00670

**Table 2.** Nucleotide and allelic sequence diversity for whole population of seven housekeeping genes of *Staphylococcus aureus* in multilocus sequence typing. Variable sites (S)/mutations(M); Number of singletons mutations (SNPs); Nucleotide diversity, Pi ( $\pi$ ); All locus contains 43 alleles.

Population(n)	<i>Phi</i> test	Recombination analysis					Linkage disequilibrium	
		theta/site	rho/site	rho/theta	LB 95%	UB 95%	IAs	P value
All (22 STs)	4.053E–5	4.854E–03	1.05E–03	2.16E–01	4.534E–04	1.698E–03	0.0052	<1.00E10–04
L1 (4 STs)	NA	6.849E–04	5.65E–01	826.053E+00	7.301E–04	2.777E+00	0.0001	1.98E10–02
L2 (18 STs)	5.413E–03	3.618E–03	2.26E–03	6.24E–01	5.134E+00	11.102 E+00	0.0014	<1.00E10–04

**Table 3.** Recombination test and estimation of *Staphylococcus aureus*. Contains 22 Sequence types (All); Lineage 1 (L1); Lineage 2 (L2); Standard association of index (IAs); Upper bound 95% (LB); Lower bound 95% (UB).

The dN/dS values ranged from 0.095 (*tpi*) to 0.432 (*aroE*; Table 2) indicating that all seven MLST loci exhibit purifying selection, since the dN/dS ratio indicates purifying selection if dN/dS < 1<sup>17</sup>. In general, the greatest number of mutations was observed for the whole population of sequences (Table 2), with emphasis for *aroE* and *yqiL* alleles. In addition, the sequences of ES-GPM (supplementary Table S4) also demonstrated that the *aroE* and *yqiL* alleles were present more polymorphisms. Thereby, the sequences ES-GPM to the *arcC* (66%) and *glpF* (60%) alleles showed the greatest number of mutations for ES-GPM in relation to non-ES-GPM (Table S4).

**Nonsense mutations in *arcC*, *aroE*, *glpF*, *yqiL* and test of recombination.** Six nonsense mutations were unexpectedly found in four distinct alleles and, interestingly, only in the lineage L2, being denoted in *arcC* (CAA → TAA) at position 111 of ST 2988 (bovine); *aroE* (CAG → TAG, AGA → TGA) at positions 274 and 57, STs 742 and 3258, respectively (bovine); *glpF* (CAA → TAA) at positions 363 in ST 1728 (ovine); *yqiL* (CAG → TAG, AAA → TAA) at positions 170 in ST 1730 and at positions 495 for the STs 1730 (ovine), 744 (bovine), 1728 (ovine), 1730 (ovine) and 5 (goat). Interestingly, nonsense mutations were not identified in the new ST 4966.

The *phi* test is a rapid statistically efficient test for recombination. The P-value generated from the *phi* test for all 22 STs was  $p = 4.053E-5$  (Table 3) and for L2 was  $p = 5.413E-03$ , which indicates significant incidence of recombination across the whole population and L2. However, for L1 it was not possible to establish the value (NA) by the program used, which suggested that there was no significant evidence of recombination for L1 or, if not, partly since there are very few informative sites. Furthermore, the *phi* test for each AT did not demonstrate a significant result.

The detecting per-site  $\rho/\theta$  (rho/theta) value for the 22 STs was 2.16E01 (Table 3), suggesting that point mutation was 4.63-fold more likely to occur than recombination at the level of whole population. The values of  $\rho/\theta$  ratio were 6.24E–01 for lineage L2, which likewise suggests point mutation in this lineage to be 1.62-fold more likely to occur than recombination. However, the  $\rho/\theta$  ratio for L1 could indicate a high recombination rate, but the *phi* test did not indicate that there would be evidence, as previously reported in the text. The IAs values were 0.0052 ( $P = < 1.00E10-04$ ) and 0.0014 ( $P = < 1.00E10-04$ ; Table 3) for all 22 STs and L2, respectively, indicating a tendency toward linkage disequilibrium between the alleles of L2 at the level of whole population (all 22 STs). This result indicates that clonal relationship and recombination were not sufficient to break down the linkage disequilibrium for all 22 STs and L2. However, for L1 a tendency toward free recombination between the alleles in lineage was suggested.

## Discussion

In recent years, persistent infections have shown an important role in the relapse and recalcitrance of infections, moreover they are likely to help spread antibiotic resistance<sup>23</sup>, being persistence a potential critical trigger for therapeutic failures<sup>4</sup>. Overall, phylogenetic analyses resulted in the establishment of two lineages of cases of bovine, ovine and caprine mastitis in different states of Brazil, and highly clonal ES-GPM unresponsive to the antibiotic enrofloxacin, in a single herd of goats. The 43 isolates that formed the 22 CCs provided further

evidence that geographic isolation was not the primary factor leading to moderate genetic differentiation of *S. aureus* ES-GPM and Non-ES-GPM.

*S. aureus* isolates from animals are commonly assigned to host-specific CCs<sup>24</sup>, and CC133 was the main group of characterized isolates belonging to ES-GPM in this study, in addition to the presence of the CC5 complex. In previous studies, CC133 has been associated with small ruminants<sup>25</sup> and has been specifically assigned to goat, ovine and bovine isolates in several different countries<sup>24</sup> and in Brazil<sup>26</sup>. According Aires-de-Sousa et al.<sup>27</sup>, the CC133 of *S. aureus* may have adapted to small ruminants, with human ancestry, due to the adaptive diversification of the genome resulting from allelic variation, the loss of genes or horizontal acquisition of mobile genetic elements.

The CC5 was also identified as belonging to the ES-GPM association but in a separate lineage from CC133. In addition, CC5 is recognized as a common clonal complex and generalized to several hosts, and is among the most prevalent clones that cause hospital infections in humans and causes methicillin-resistant *S. aureus* (MRSA)<sup>28</sup>. CC5 has also been characterized in bovine mastitis<sup>29</sup>, in buffalo milk<sup>30</sup> and also isolated in foods such as samples of milk and dairy products<sup>31</sup>. Consequently, the persistent *S. aureus* lineage, may be silently acting on persistent infections along with CC133 and CC5 in other cases of mastitis or even in other clonal complexes not studied here.

RABELLO et al.<sup>32</sup> suggests that the prevalence of a limited number of clones is strictly related to mastitis in different herds and HOEKSTRA et al.<sup>33</sup> demonstrated that the same genotype of *S. aureus* can cause clinical and subclinical mastitis in goats and ovine. However, diversity analyses implemented using different techniques indicate that there is a difference between sequences by type of mastitis, such as ES-GPM. This indicates that questions of agent selection or challenge<sup>3</sup> can trigger the persistence<sup>34</sup> of mastitis, and with this, questions related to the history of choosing certain classes of antibiotics in treatments and adaptations of the agent must be taken into account in the analysis of MLST in mastitis, as in cases of MRSA or MSSA (methicillin-susceptible *S. aureus*). As the mutations that lead bacteria to persist to enrofloxacin and other fluoroquinolones are still being determined<sup>12,35–37</sup>, more complete studies, with total genome sequencing<sup>38</sup> and biocomputational analyses should be better implemented in the future to elucidate these issues.

As was reported above, we did not find high genetic variability in this study, but nevertheless, it has been reported that cases of animal mastitis by *S. aureus* in small regions of Brazil are related to low genetic variability and a small clonal population<sup>27</sup>. Furthermore, a greater number of sequences could better demonstrate the genetic diversity, patterns of distribution and evolutionary history of this agent in dairy animals. On the other hand, in cases with a large number of isolates collected from different regions of Norway, the strains were closely related genetically, and their clonal population was responsible for most cases of mastitis by *S. aureus* in domestic ruminants<sup>39</sup>. The low diversity of *S. aureus* in milk samples in studies with ruminants<sup>24,27,40</sup> may be related to its later diversification from *S. aureus* associated with humans through a combination of foreign DNA acquisition and gene decay<sup>41</sup> and also that their strains, in general, are young in relation to the species<sup>38</sup>. However, in this study the strains related to bovine mastitis demonstrated greater nucleotide polymorphism possibly associated with evolutionary pressures under the pathogen due to factors related to adaptation of a species to optimize the process of infection, escape host immune response, and also as possible adaptation to a different environmental niches and use of antibiotics that impacts on the evolution of certain core genes<sup>42</sup>.

In general, for isolates of *S. aureus* from invasive disease the r/m per allele parameter is approximately 1:1<sup>43</sup>, which means that the isolates would have the same probability of diversifying in their large population by recombination as by point mutation. However, in humans, the genetic variability of *S. aureus* may be mainly associated with point mutation, since alleles are up to 15 times more likely to change by point mutation than by recombination<sup>44</sup>, similar to the results of our study. Conversely, in *S. aureus* adapted to cattle, it was found by MLST alleles that a nucleotide substitution was more likely to be due to recombination than to point mutation, and equally likely in humans<sup>45</sup>. Nevertheless, there is an extensive difference between the *S. aureus* genomes associated with cattle and those isolated from humans<sup>41</sup>. These differences between genomes and their r/m ratios demonstrate that there are gaps in the understanding of the diversification behavior of *S. aureus* between different hosts. In addition, subtle changes in strain due to single non-synonymous point mutation in *S. aureus* may be involved with persistence to antibiotics<sup>46</sup> as diagnosed in GPMs unresponsive to treatment with enrofloxacin.

Nevertheless, there is evidence of widespread homologous recombination in the core genome of *S. aureus* in studies of animal-associated strains of ovines, bovines, and poultry<sup>38</sup>, due to the performance of mobile genetic elements (MGEs), which generate a landscape of hotspots in the core genome<sup>38</sup>. Furthermore, there are concerns about the exchange of *S. aureus* CCs between animals and humans<sup>47,48</sup>, in which strains of this agent have been previously described in the literature as being capable of causing disease in animals and humans<sup>49</sup>. Therefore, the probability of widespread homologous recombination between *S. aureus* from different hosts may be high, and the potential of this agent to infect both humans and animals may indicate the chances of a greater or lesser degree of recombination or point mutation not having a noticeably clear pattern in *S. aureus*, unless characteristics other than the infection chain are intrinsically defined in the analysis metadata, such as those of ES-GPM.

The greater diversity and polymorphism of some alleles of *S. aureus* may be associated with adaptive mutations due to response to environmental changes or a switch in host species, since this agent presents tropism to several hosts and in particular, antibiotic exposure<sup>50</sup>. In addition, housekeeping genes are genes associated with metabolic maintenance, which demonstrates that adaptations in their genomes may be affecting metabolism in response to distinct nutrient availability<sup>50</sup>. Of all genes, *yyjL* and *aroE* showed the highest values of non-synonymous substitutions and mutations. The *yyjL* locus in *S. aureus* demonstrated a potential signature of recombination<sup>45</sup> compared to the other six gene fragments<sup>38</sup>, but the role of this locus in infections is poorly understood. However, a high number of non-synonymous substitutions may suggest that the removal of deleterious mutations by purification selection should be relatively slow. Gene *aroE* has also been shown to contribute to the chronicity of *S. aureus* infection, such as the invasiveness and cytotoxicity of the agent, with an increase in the



load of intracellular bacteria<sup>51</sup>, and because of that, this same locus can play an important role in the persistence of infection such as ES-GPM in goats. Moreover, the *aroE* gene locus is shown to be in a region of the genome with an excess of homologous recombination, most likely in MGEs<sup>38</sup> and with that it can confer an increased capacity to colonize and infect ruminants<sup>52</sup>.

Interestingly, the *glpF* sequence, among all others, was shown to be more conserved, with the lowest haplotypic diversity and the highest average G + C content, and also presenting low numbers of mutations, even though it is present in a region of the genome with excess homologous recombination<sup>38</sup>. The highest average G + C content may be linked to the pathogenic potential for the host genome<sup>53</sup>. However, we can emphasize that the *glpF* gene has a role in the formation of the L form in bacteria<sup>54</sup>, which is directly involved in antibiotic tolerance or persistence<sup>55</sup> in *S. aureus*, such as to ampicillin or norfloxacin<sup>3</sup>. This demonstrates that *glpF* sequence, even though it is a member of the metabolic housekeeping genes and is present in the core genome, it is directly linked to persistence in infections; therefore, a high *glpF* sequence polymorphism may not be advantageous for colonization of the host as in ES-GPM for enrofloxacin and thus is shown to be more conserved within the other six alleles.

## Conclusion

In this study we identified 27 specific genetic mutations for strains of ES-GPM (*S. aureus* isolated from goats with persistent mastitis) that may help in the future to discriminate *S. aureus* in cases of persistent goat mastitis (GPM). In addition, in 22 CCs that we found, CC133, CC5 and a new ST 4966 were specifically related to ES-GPM. We describe polymorphisms in specific alleles *arcC* and *glpF* genes, that showed a greatest number of mutations for ES-GPM in relation to non-ES-GPM. Furthermore, the identification of *S. aureus* and these polymorphisms genetics in persistent bacterial infections together with the MLST, can assist in decision making for the appropriate choice of protocol and of an antibiotic for the treatment of mastitis persists in goats. We hope that future studies can better clarify the persistence of this agent in certain antibiotic treatments.

## Methods

**Bacterial strains.** In this study, we used two data sources, the first being cataloged by our research group with 18 isolates of *S. aureus* from goats diagnosed with GPM, recovered from milk samples, and treated with enrofloxacin antibiotic. The animals are kept under intensive farming in a free stall regime, with a high-level mechanical milking system and automatic cleaning of milk pipes. The harvest of samples was established in the Capril UFV, located in the mesoregion of Zona da Mata of Minas Gerais, Brazil. The isolates were identified by phenotypic and genotypic tests, as well as by methods of bacteriological examination and antibiotic sensitivity tests as previously reported<sup>12,56</sup>.

Briefly, the animals selected were examined and diagnosed with clinical mastitis caused by *S. aureus*. Firstly, were evaluated for signs of clinical mastitis and the presence of at least visually abnormal milk (i.e., the presence of flakes, clots, blood, or serous milk), as well as changes in the mammary gland, such as an increased volume and body temperature, and the presence of pain, redness during forestripping performed at the milking parlour, in the presence of a veterinarian. After antibiogram results, these animals were treated with enrofloxacin (KINETOMAX – Bayer). Minimal inhibitory concentrations (MICs) for enrofloxacin were performed using the Etest method (BioMérieux, Marcy l'Étoile, France). However, due to acute mastitis and the need for rapid treatment, only after completion of seven days of enrofloxacin treatment were these MIC results available. Twenty-one days after the completion of treatment, these animals continued to have clinical mastitis. New milk samples were collected, and *S. aureus* was isolated again. Therefore, the diagnosis of persistent mastitis (GPM) was based on the detection of the same species of agent in more than one consecutive sampling<sup>57–59</sup>. Thus, 18 isolates of *S. aureus* were obtained (nine before treatment and nine after treatment) and were grouped in this study as ES-GPM.

The second source of data for the analyzes were another 25 sequences of *S. aureus* isolates recovered from milk samples, these being selected from mastitis cases obtained from ovine and bovine isolates from the MLST online database (<https://pubmlst.org/organisms/staphylococcus-aureus/>). The search terms used were: country (Brazil), source (Milk) and disease (Mastitis), totaling 43 isolates from seven Brazilian states (Table S3).

**Genomic DNA extraction, MLST locus amplification and sequencing.** The seven housekeeping genes of *S. aureus* isolates obtained from the GPMs were sequenced: carbamate kinase (*arcC*), shikimate dehydrogenase (*aroE*), glycerol kinase (*glpF*), guanylate kinase (*gmk*), phosphate acetyltransferase (*pta*), triosephosphate isomerase (*tpi*), acetyl coenzyme A acetyltransferase (*yqiL*; Table S2), as previously described in methods and evaluated at <https://pubmlst.org/organisms/staphylococcus-aureus/><sup>60</sup>. DNA from *S. aureus* isolates were extracted using the PROMEGA kit following the manufacturer's protocol and fragments amplified according to the protocol described by ENRIGHT et al.<sup>61</sup>. The gene and primer specifications are shown in Table S2. MLST locus amplification was performed in 50 µL reaction volumes containing 0.5 µL DNA, 0.5 µg of each primer, 1U Taq DNA polymerase (Qiagen, Crawley, UK), 5 µL buffer 103 (supplied with Taq Polymerase) and 0.2 mM deoxynucleoside triphosphates (Perkin-Elmer Applied Biosystems; Foster City, California). Initial denaturation was for 5 min at 95 °C, followed by 30 cycles at 55 °C for 1 min, extension at 72 °C for 1 min and denaturation at 95 °C for 1 min, followed by the final extension at 72 °C for 5 min. The amplified products were sent for sequencing at MACROGEN, INC. (Seoul, South Korea) using capillary electrophoresis.

**Alignment, editing and curation of *S. aureus* MLST sequences.** The 18 forward and reverse sequences of the seven *S. aureus* MLST genes obtained from ES-GPM were trimmed and aligned with their respective allele type (AT) and sequence type (ST), obtained from the PubMLST database (<https://pubmlst.org/organisms/staphylococcus-aureus/>)<sup>21</sup> in the UGENE software<sup>62</sup> along with the 25 deposited sequences.

**Expanded multi-locus sequence typing (E-MLST).** The ST, allelic profile and CC were confirmed by the E-MLST database (<https://pubmlst.org/organisms/staphylococcus-aureus/>)<sup>21</sup>. Any ST or allele not previously described were submitted to the database and assigned new AT and ST numbers.

**Setting the best-fit model of DNA evolution and phylogenetic tree.** Phylogenetic reconstructions by Bayesian Inference (BI) and Maximum Likelihood (ML) require the setting of a DNA evolution model to calculate the probabilities of nucleotide changes<sup>63</sup>. “HKY + I” was selected as the best evolution model of the MLST genes by the jModeltest 2 program<sup>64</sup>. The phylogenetic tree was inferred by the Markov Chain Monte Carlo (MCMC) method using MrBayes v3.1.2<sup>65</sup> and phylogenetic trees were calculated in two runs with 1,000,000 (one million) generations and a sampling frequency of 100 (one hundred). The parameter convergence was analyzed in Tracer version 1.6 (<http://tree.bio.ed.ac.uk/software/tracer>) and 25% of the trees generated were burned to produce the consensus tree. The phylogenetic tree and geospatial information was visualized together with associated metadata using Microreact Web server version 5.93.0<sup>66</sup> and by Figtree software version 1.4.4<sup>67</sup>.

**Sequence diversity analyses.** The G + C content, variable sites (Ss) or mutations (Ms), number of polymorphic sites, average pairwise nucleotide, number of singleton (SNPs) and difference per site ( $\pi$ ) were calculated with DnaSP Version 6.12.03<sup>68</sup>. The average non-synonymous/synonymous rate ratio (dN/dS) was calculated with KaKs Calculator Version 2.0<sup>69</sup> to infer the direction and magnitude of natural selection. The ratio of non-synonymous and synonymous substitutions (dN/dS) measures the level of selection in a protein-coding gene. Further, the dN/dS ratio indicates purifying selection if dN/dS < 1, positive selection if dN/dS > 1, and neutral evolution if values are close to 1<sup>17</sup>.

**Population structure and recombination analyses.** Strain relationships were analyzed using the goeBURST algorithm<sup>70</sup>, as implemented in the software PHYLOViZ<sup>71</sup> to cluster the STs into CCs based on the most stringent definition. Global optimal eBURST implemented by PHYLOViZ was used to cluster STs, generating a multi-locus sequence tree (MS tree) to visualize possible evolutionary relationships between STs. The pairwise homoplasy index (*phi*) test<sup>72</sup> implemented in SplitTree4<sup>73</sup> for recombination was performed, and a *P*-value of < 0.05 indicated that recombination existed.

The LDhat program<sup>74</sup> implemented in Recombination Detection Program (RDP) v.4.97<sup>75</sup> was used to calculate the per-site  $\rho/\theta$  ratio based on the concatenated sequences of seven loci with 1,000,000 MCMC updates. The parameters  $\rho$  and  $\theta$  represent the rates of recombination and mutation, respectively. Linkage disequilibrium from allelic data was evaluated by calculating the standardized index of association (IAs) using LIAN v3.735<sup>76</sup> in web interface (<http://guanine.evolbio.mpg.de/cgi-bin/lian/lian.cgi.pl/query>). The null hypothesis of complete linkage equilibrium (IAs > 0; presence of linkage disequilibrium or clonality) was tested by using Monte Carlo methods with 10,000 iterations on allelic profile<sup>17</sup>. If there is linkage equilibrium because of frequent recombination events, the expected value of IAs is zero, which suggests no association between alleles at different loci; if IAs are statistically significant different from zero, alleles are suggested with genetic linkage<sup>77</sup>.

**Genetic network.** Genetic networks present an alternative view of genealogies represented by bifurcated structures of phylogenetic trees, and the possible dispersal routes of *S. aureus* isolates in the Brazilian dairy milk were predicted following the methodology described by Vidigal et al.<sup>78</sup>. To reconstruct the network, sequences of AT and ST genes were grouped into haplotypes using DnaSP v6<sup>68</sup>. Following this, the network was constructed using the network 4.6.1.0 (<http://www.fluxus-technology.com>) and the Median Joining algorithm (MJ)<sup>79</sup>.

**Ethics statement.** The experimental protocol was approved by the Ethics Committee (Comissão de ética no uso de animais - CEUA) of the Federal University of Viçosa, according to the protocol number 43/2016. The methods were carried out in accordance with the approved guidelines. In addition, this experiment was conducted by Bacterial Diseases Laboratory (LDBAC) and the Molecular Biology Laboratory (BIOMOL) located at the Veterinary Department (DVT) of the Universidade Federal de Viçosa (UFV), Viçosa, Minas Gerais.

### Availability of data and materials

All data generated and/or analyzed during this study are included in this published article, and at Microreact interactive viewer<sup>66</sup> [<https://microreact.org/project/puA-MDz8o>]. Additionally, the accession numbers of the sequences or reference codes used in this study are called IDs and can be seen in the supplementary tables S1 and S3, with their isolated characteristics and their corresponding AT and ST. These data are publicly available and accessible online at the *S. aureus* PubMLST database [<https://pubmlst.org/organisms/staphylococcus-aureus/>] (IDs: 33768 - 33785)<sup>21</sup>.

Received: 25 March 2021; Accepted: 28 July 2021

Published online: 26 August 2021

### References

- Peton, V. & Le Loir, Y. Staphylococcus aureus in veterinary medicine. *Infect. Genet. Evol.* **21**, 602–615 (2014).
- Rollin, G. et al. Intracellular survival of staphylococcus aureus in endothelial cells: A matter of growth or persistence. *Front. Microbiol.* **8**, 1354 (2017).
- Han, J. et al. Glycerol uptake is important for L-form formation and persistence in *Staphylococcus aureus*. *PLoS ONE* **9**, e108325 (2014).
- Peyrsson, F. et al. Intracellular Staphylococcus aureus persists upon antibiotic exposure. *Nat. Commun.* **11** (2020).

5. Balaban, N. Q. *et al.* Definitions and guidelines for research on antibiotic persistence. *Nat. Rev. Microbiol.* **17**, 441–448 (2019).
6. Moroni, P. *et al.* Characterization of *Staphylococcus aureus* isolated from chronically infected dairy goats. *J. Dairy Sci.* **88**, 3500–3509 (2005).
7. Côté-Gravel, J. & Malouin, F. Symposium review: Features of *Staphylococcus aureus* mastitis pathogenesis that guide vaccine development strategies. *J. Dairy Sci.* **102**, 4727–4740 (2019).
8. Bergonier, D., de Crémoux, R., Rupp, R., Lagriffoul, G. & Berthelot, X. Mastitis of dairy small ruminants. *Vet. Res.* **34**, 689–716 (2003).
9. Alawa, J. P., Ngele, M. B. & Ogwu, D. Chronic caprine mastitis in Nigerian goat breeds: Microbiological flora and histopathological findings. *Small Rumin. Res.* **35**, 203–207 (2000).
10. Wolfson, J. S., Hooper, D. C., McHugh, G. L., Bozza, M. A. & Swartz, M. N. Mutants of *Escherichia coli* K-12 exhibiting reduced killing by both quinolone and  $\beta$ -lactam antimicrobial agents. *Antimicrob. Agents Chemother.* **34**, 1938–1943 (1990).
11. Grunert, T. *et al.* Distinct phenotypic traits of *Staphylococcus aureus* are associated with persistent, contagious bovine intramammary infections. *Sci. Rep.* **8**, 15968 (2018).
12. Polveiro, R. C. *et al.* Effects of enrofloxacin treatment on the bacterial microbiota of milk from goats with persistent mastitis. *Sci. Rep.* **10**, 1–13 (2020).
13. Altreuther, P. Data on chemistry and toxicology of Baytril. *Vet. Med. Rev.* **2**, 87–89 (1987).
14. Elsheikh, H. A., Taha, A. A. W., Khalafallah, A. I. & Osman, I. A. M. Disposition kinetics of enrofloxacin (Baytril 5%) in sheep and goats following intravenous and intramuscular injection using a microbiological assay. *Res. Vet. Sci.* **73**, 125–129 (2002).
15. Babaahmady, E. & Khosravi, A. Toxicology of baytril (enrofloxacin). *Afr. J. Pharm. Pharmacol.* **5**, 2042–2045 (2011).
16. Barkema, H. W., Schukken, Y. H. & Zadoks, R. N. Invited review: The role of cow, pathogen, and treatment regimen in the therapeutic success of bovine *Staphylococcus aureus* mastitis. *J. Dairy Sci.* **89**, 1877–1895 (2006).
17. Gautam, V. *et al.* Multilocus sequence analysis reveals high genetic diversity in clinical isolates of *Burkholderia cepacia* complex from India. *Sci. Rep.* **6**, 35769 (2016).
18. Morrison, D. A. Networks in phylogenetic analysis: New tools for population biology. *Int. J. Parasitol.* **35**, 567–582 (2005).
19. Li, X., Saadat, S., Hu, H., Li, X. & Birol, I. BHap: A novel approach for bacterial haplotype reconstruction. *Bioinformatics* **35**, 4624–4631 (2019).
20. Didelot, X., Urwin, R., Maiden, M. C. J. & Falush, D. Genealogical typing of *Neisseria meningitidis*. *Microbiology* **155**, 3176–3186 (2009).
21. Jolley, K. A., Bray, J. E. & Maiden, M. C. J. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications [version 1; referees: 2 approved]. *Wellcome Open Res.* **3** (2018).
22. Lima, M. C. *et al.* Profiles of *Staphylococcus aureus* isolated from goat persistent mastitis before and after treatment with enrofloxacin. *BMC Microbiol.* **20**, 127 (2020).
23. Fisher, R. A., Gollan, B. & Helaine, S. Persistent bacterial infections and persister cells. *Nat. Rev. Microbiol.* **15**, 453–464 (2017).
24. Smyth, D. S. *et al.* Molecular genetic typing reveals further insights into the diversity of animal-associated *Staphylococcus aureus*. *J. Med. Microbiol.* **58**, 1343–1353 (2009).
25. Eriksson, J., Espinosa-Gongora, C., Stamphøj, I., Larsen, A. R. & Guardabassi, L. Carriage frequency, diversity and methicillin resistance of *Staphylococcus aureus* in Danish small ruminants. *Vet. Microbiol.* **163**, 110–115 (2013).
26. Silva, N. C. C. *et al.* Molecular characterization and clonal diversity of methicillin-susceptible *Staphylococcus aureus* in milk of cows with mastitis in Brazil. *J. Dairy Sci.* **96**, 6856–6862 (2013).
27. Aires-de-Sousa, M. *et al.* Characterization of *Staphylococcus aureus* isolates from buffalo, bovine, ovine, and caprine milk samples collected in Rio de Janeiro State, Brazil. *V. Appl. Environ. Microbiol.* **73**, 3845–3849 (2007).
28. Challagundla, L. *et al.* Phylogenomic classification and the evolution of clonal complex 5 methicillin-resistant *Staphylococcus aureus* in the Western Hemisphere. *Front. Microbiol.* **9** (2018).
29. Li, T. *et al.* Molecular characteristics of *Staphylococcus aureus* causing bovine mastitis between 2014 and 2015. *Front. Cell. Infect. Microbiol.* **7**, 127 (2017).
30. El-Ashker, M. *et al.* *Staphylococci* in cattle and buffaloes with mastitis in Dakahlia Governorate, Egypt. *J. Dairy Sci.* **98**, 7450–7459 (2015).
31. Basanisi, M. G., La Bella, G., Nobili, G., Franconieri, I. & La Salandra, G. Genotyping of methicillin-resistant *Staphylococcus aureus* (MRSA) isolated from milk and dairy products in South Italy. *Food Microbiol.* **62**, 141–146 (2017).
32. Rabello, R. F. *et al.* Multilocus sequence typing of *Staphylococcus aureus* isolates recovered from cows with mastitis in Brazilian dairy herds. *J. Med. Microbiol.* <https://doi.org/10.1099/jmm.0.47357-0> (2007).
33. Hoekstra, J. *et al.* Differences between *Staphylococcus aureus* lineages isolated from ovine and caprine mastitis but not between isolates from clinical or subclinical mastitis. *J. Dairy Sci.* **102**, 5430–5437 (2019).
34. Domingue, G. J. & Woody, H. B. Bacterial persistence and expression of disease. *Clin. Microbiol. Rev.* **10**, 320–344 (1997).
35. Frimodt-Møller, J. *et al.* Mutations causing low level antibiotic resistance ensure bacterial survival in antibiotic-treated hosts. *Sci. Rep.* **8**, 1–13 (2018).
36. Barrett, T. C., Mok, W. W. K., Murawski, A. M. & Brynildsen, M. P. Enhanced antibiotic resistance development from fluoroquinolone persists after a single exposure to antibiotic. *Nat. Commun.* **10**, 1–11 (2019).
37. Redgrave, L. S., Sutton, S. B., Webber, M. A. & Piddock, L. J. V. Fluoroquinolone resistance: Mechanisms, impact on bacteria, and role in evolutionary success. *Trends Microbiol.* **22**, 438–445 (2014).
38. Everitt, R. G. *et al.* Mobile elements drive recombination hotspots in the core genome of *Staphylococcus aureus*. *Nat. Commun.* **5**, 1–9 (2014).
39. Mørk, T., Tollersrud, T., Kvitle, B., Jørgensen, H. J. & Waage, S. Comparison of *Staphylococcus aureus* genotypes recovered from cases of bovine, ovine, and caprine mastitis. *J. Clin. Microbiol.* **43**, 3979–3984 (2005).
40. Ben Zakour, N. L. *et al.* Genome-wide analysis of ruminant *Staphylococcus aureus* reveals diversification of the core genome. *J. Bacteriol.* **190**, 6302–6317 (2008).
41. Herron-Olson, L., Fitzgerald, J. R., Musser, J. M. & Kapur, V. Molecular correlates of host specialization in *Staphylococcus aureus*. *PLoS ONE* **2**, e1120 (2007).
42. Anisimova, M., Bielawski, J., Dunn, K. & Yang, Z. Phylogenomic analysis of natural selection pressure in *Streptococcus* genomes. *BMC Evol. Biol.* **7**, 154 (2007).
43. Spratt, B. G., Hanage, W. P. & Feil, E. J. The relative contributions of recombination and point mutation to the diversification of bacterial clones. *Curr. Opin. Microbiol.* **4**, 602–606 (2001).
44. Feil, E. J. *et al.* How clonal is *Staphylococcus aureus*? *J. Bacteriol.* **185**, 3307–3316 (2003).
45. Budd, K. E. *et al.* Extensive genomic diversity among bovine-adapted *Staphylococcus aureus*: Evidence for a genomic rearrangement within CC97. *PLoS ONE* **10**, e0134592 (2015).
46. Mechler, L. *et al.* A novel point mutation promotes growth phase-dependent daptomycin tolerance in *Staphylococcus aureus*. *Antimicrob. Agents Chemother.* **59**, 5366–5376 (2015).
47. Ye, X. *et al.* Livestock-associated methicillin and multidrug resistant *S. aureus* in humans is associated with occupational pig contact, not pet contact. *Sci. Rep.* **6**, 1–9 (2016).
48. Price, L. B. *et al.* *Staphylococcus aureus* CC398: Host adaptation and emergence of methicillin resistance in livestock. *MBio* **3**, 1–6 (2012).

49. Manara, S. *et al.* Whole-genome epidemiology, characterisation, and phylogenetic reconstruction of *Staphylococcus aureus* strains in a paediatric hospital. *Genome Med.* **10**, 82 (2018).
50. Richardson, E. J. *et al.* Gene exchange drives the ecological success of a multi-host bacterial pathogen. *Nat. Ecol. Evol.* **2**, 1468–1478 (2018).
51. McLean, K. *et al.* Artificial selection for pathogenicity mutations in *Staphylococcus aureus* identifies novel factors relevant to chronic infection. *Infect. Immun.* **87** (2019).
52. Herron, L. L. *et al.* Genome sequence survey identifies unique sequences and key virulence genes with unusual rates of amino acid substitution in bovine *Staphylococcus aureus*. *Infect. Immun.* **70**, 3978–3981 (2002).
53. Mann, S. & Chen, Y. P. P. Bacterial genomic G + C composition-eliciting environmental adaptation. *Genomics* **95**, 7–15 (2010).
54. Glover, W. A., Yang, Y. & Zhang, Y. Insights into the molecular basis of L-form formation and survival in *Escherichia coli*. *PLoS ONE* **4**, e7316 (2009).
55. Spoering, A. L., Vulić, M. & Lewis, K. GlpD and PlsB participate in persister cell formation in *Escherichia coli*. *J. Bacteriol.* **188**, 5136–5144 (2006).
56. Lima, M. C. *et al.* Profiles of *Staphylococcus aureus* isolated from goat persistent mastitis before and after treatment with enrofloxacin. *BMC Microbiol.* **20** (2020).
57. Contreras, A., Corrales, J. C., Sanchez, A. & Sierra, D. Persistence of subclinical intramammary pathogens in goats throughout lactation. *J. Dairy Sci.* **80**, 2815–2819 (1997).
58. Koop, G. *et al.* Differences between coagulase-negative *Staphylococcus* species in persistence and in effect on somatic cell count and milk yield in dairy goats. *J. Dairy Sci.* **95**, 5075–5084 (2012).
59. Moroni, P. *et al.* Subclinical mastitis and antimicrobial susceptibility of *Staphylococcus caprae* and *Staphylococcus epidermidis* isolated from two Italian goat herds. *J. Dairy Sci.* **88**, 1694–1704 (2005).
60. Saunders, N. A. & Holmes, A. Multilocus sequence typing (MLST) of *Staphylococcus aureus*. In *Methods in molecular biology* (Clifton, N.J.) vol. 391, 71–85 (2007).
61. 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.* **38**, 1008–1015 (2000).
62. Okonechnikov, K. *et al.* Unipro UGENE: A unified bioinformatics toolkit. *Bioinformatics* **28**, 1166–1167 (2012).
63. Guindon, S., Delsuc, F., Dufayard, J. F. & Gascuel, O. Estimating maximum likelihood phylogenies with PhyML. *Methods Mol. Biol.* **537**, 113–137 (2009).
64. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: More models, new heuristics and parallel computing. *Suppl. mat. Nat. Methods* **9**, 772–772 (2012).
65. Huelsenbeck, J. P. & Ronquist, F. MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* **17**, 754–755 (2001).
66. Argimón, S. *et al.* Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb. genomics* **2** (2016).
67. Rambaut, A. FigTree v1.4.3. *Molecular Evolution, Phylogenetics and Epidemiology* (2016).
68. Rozas, J. *et al.* DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* **34**, 3299–3302 (2017).
69. Wang, D., Zhang, Y., Zhang, Z., Zhu, J. & Yu, J. KaKs\_Calculator 2.0: A toolkit incorporating gamma-series methods and sliding window strategies. *Genom. Proteom. Bioinform.* **8**, 77–80 (2010).
70. Francisco, A. P., Bugalho, M., Ramirez, M. & Carriço, J. A. Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. *BMC Bioinform.* **10**, 152 (2009).
71. Francisco, A. P. *et al.* PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods. *BMC Bioinform.* **13**, 87 (2012).
72. Bruen, T. C., Philippe, H. & Bryant, D. A simple and robust statistical test for detecting the presence of recombination. *Genetics* **172**, 2665–2681 (2006).
73. Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254–267 (2006).
74. Auton, A. & McVean, G. Recombination rate estimation in the presence of hotspots. *Genome Res.* **17**, 1219–1227 (2007).
75. Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. RDP4: Detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **1** (2015).
76. Haubold, B. & Hudson, R. R. LIAN 3.0: detecting linkage disequilibrium in multilocus data. *Bioinformatics* **16**, 847–849 (2000).
77. Guo, C. *et al.* MLST-based inference of genetic diversity and population structure of clinical *Klebsiella pneumoniae*, China. *Sci. Rep.* **5**, 7612 (2015).
78. Vidigal, P. M. P. *et al.* Tripping over emerging pathogens around the world: A phylogeographical approach for determining the epidemiology of Porcine circovirus-2 (PCV-2), considering global trading. *Virus Res.* **163**, 320–327 (2012).
79. Bandelt, H. J., Forster, P. & Röhl, A. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* **16**, 37–48 (1999).

## Acknowledgements

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001. The authors also acknowledge the financial support from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brasília, Brazil), FAPEMIG (Fundação de Amparo à Pesquisa de Minas Gerais, Belo Horizonte, Brazil). Maria Aparecida Scatamburlo Moreira is supported by CNPq.

## Author contributions

R.C.P. conceived the study, R.C.P., M.M.C.G., T.C.B., I.S.L., and M.C.L. conducted the experiment, R.C.P., P.M.P.V., and M.M.C.G. analyzed the results. R.C.P., and M.A.S.M. wrote the manuscript. M.A.S.M., and P.M.P.V. reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-96764-z>.

**Correspondence** and requests for materials should be addressed to M.A.S.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021