Comparative Analysis of *agr* Locus Diversification and Overall Genetic Variability among Bovine and Human *Staphylococcus aureus* Isolates

Philippe Gilot¹* and Willem van Leeuwen²

Unite de Pathologie Infectieuse et Immunologie, Institut National de la Recherche Agronomique, F-37380 Nouzilly, France, and Erasmus MC, Rotterdam, The Netherlands²

Received 9 July 2003/Returned for modification 23 September 2003/Accepted 26 November 2003

The accessory gene regulator (*agr*) is a central system that controls the expression of *Staphylococcus aureus* virulence factors. In this study, the distribution of *agr* alleles, defined by *agr* restriction fragment length polymorphism within *agr* interference groups, among *S. aureus* isolates from bovine and human origin was measured with PCR-based techniques. Statistically highly significant associations of some *agr* alleles with the infection of a specific host were found. The genetic difference between the two *S. aureus* populations was further stressed after analysis of their genetic background by binary typing.

*Staphylococcus aureus* is an important cause of a variety of diseases in humans and animals and is a large economic problem worldwide (9, 14). Whereas the mechanism of pathogenicity of *S. aureus* is still not completely understood, the bacterium is hypothesized to invade the tissue by starting to adhere directly to epithelial cells (13, 18). The expression of such adhesive proteins, which recognize macromolecules of the host tissue, is globally controlled by the *agr* locus (17). The *agr* operon encodes a two-component signaling system, which is driven by a quorum-sensing autoinducing peptide (AIP), encoded by the *agrD* gene. *S. aureus* isolates can be classified in four major *agr* interference groups on the basis of AIP specificity for its signal receptor (AgrC) (4, 6, 7). In a laboratory setting, Ji et al. showed that AIP from strains belonging to a specific *agr* group activate each other’s *agr* response, whereas AIP from strains of different *agr* groups compete with each other at the level of *agr* expression, since each AIP inhibits the *agr* response in strains from other groups (7). This type of bacterial interference is unusual, because it affects the expression of a subset of genes rather than inhibiting growth. Diverse works have noted the existence of distinct *agr* alleles, as defined with *agr* restriction fragment length polymorphism, within *agr* groups both in clinical *S. aureus* isolates that originated with humans and in those from cows (2, 15, 16, 19, 24, 29). In a recent study, one of us identified 12 distinct *agr* alleles in an epidemiologically unrelated collection of bovine mastitis isolates. The majority of strains from this collection are nevertheless represented by one particular *agr* allele (R III A1) from *agr* group 1 (2). These observations suggest the occurrence of host-adapted (or tissue-adapted) *S. aureus* strains in which the *agr* restriction type (allele) may play a significant role.

In this study, we compared the distribution of *agr* alleles in relation to the genotype among *S. aureus* strains of human and bovine origin by using restriction endonuclease-PCR of the *agr* locus (*agr* RE-PCR) and binary typing (BT).

---

*agr* polymorphism. Two hundred fifty-one human *S. aureus* isolates, originating in diverse collections, were analyzed by *agr* RE-PCR. The results were compared with those obtained from a collection of bovine mastitis isolates, previously analyzed with the same method (Table 1). In brief, a 1,070-bp variable region of the *agr* operon (a sequence encoding *agrD*, the N-terminal two-thirds of *agrC*, and the C-terminal part of *agrB*) was amplified by PCR, as described previously (2). PCR amplifications of all tested human isolates resulted in amplicons of the expected molecular weight. After restriction with *AluI* or *RsaI*, these amplicons were electrophoresed through a 3% agarose gel, giving nine (A1, A2, A4, A5, A7 to A10, and A13) and five (R I’ and R III to R VI) different profiles, respectively. Most of these patterns were previously identified after restriction of the same region of the *agr* locus of *S. aureus* bovine mastitis isolates (2). Three of them (A9, A10, and A13) were uniquely found in human isolates. Patterns A9 (502, 244, 177, and 149 bp) and A10 (772, 144, 74, 73, and 11 bp) correspond to restriction profiles of the *agr* locus from *agr* group reference strains and from strains whose genomes were sequenced, all isolated from humans (A9: *agr* group 2 reference strain with GenBank accession no AF001782, strains N315 and Mu50; A10: *agr* group 3 reference strain with GenBank accession no AF001783, strain MRSA-252 [2]). Pattern A13 (338, 274, 168, 155, 113, and 25 bp) was newly identified during this study (Fig. 1).

The combination of the *AluI* and *RsaI* restriction patterns allows the construction of restriction maps and the definition of *agr* restriction types (alleles) (2). Eleven *agr* restriction types were defined among the strains isolated from humans (Fig. 2). Two of these types, type R I’-A4 (a R I’-A1 type which has lost the *AluI* site at position 437) and type R I’-A13 (a R I’-A1 type which has lost the *AluI* site at position 890) were newly identified during this study. *agr* restriction types were previously classified into an evolutionary scheme, which allows them to be placed in one of the four *agr* interference groups (2). By this kind of analysis, we found that the majority of strains isolated from humans fall into the classification *agr* group 1 (61.7%) and group 2 (33.5%). Group 3 contains only 4.8% of the analyzed strains, and no group 4 strains were identified in our collection of strains isolated from humans. The above *agr*
group classification of strains was confirmed by an agr group-specific multiplex PCR (2). The repartition of human isolates in the agr interference group is quite similar to what was found for bovine mastitis isolates (69% agr group 1 strains, 23.9% agr group 2 strains, 2.8% agr group 3 strains, and 1.4% agr group 4 strains) (2). Whereas the repartition of strains in agr groups is quite similar in both populations, the 17 different agr types identified are very differently distributed in each of them. The most striking differences (statistically highly significant) were found in type R I'-A1 (35.4% of strains of human origin but only 4.2% of strains of bovine origin), type R IV-A9 (26.7% of strains of human origin but 0% of strains of bovine origin), type R III-A1 (8.8% of strains of human origin but 56.3% of strains of bovine origin), and type R IV-A7 (0.4% of strains of human origin but 8.4% of strains of bovine origin). Some types, containing only a few strains, were also found to be unique either to the population of human origin or to the population of bovine origin (Fig. 2).

The distribution of agr restriction types suggests that the two populations are different and indicates that some types probably expand in one or the other population due to their possession of unique genetic characteristics.

### TABLE 1. Strains analyzed in this study

<table>
<thead>
<tr>
<th>Collection</th>
<th>Geographic origin&lt;sup&gt;a&lt;/sup&gt;</th>
<th>No. of strains</th>
<th>Origin</th>
<th>Description&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Worldwide</td>
<td>55</td>
<td>Human</td>
<td>MRSA strains</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>New York City Hospital</td>
<td>64</td>
<td>Human</td>
<td>MRSA outbreak</td>
<td>27</td>
</tr>
<tr>
<td>3</td>
<td>USA</td>
<td>49</td>
<td>Human</td>
<td>Multicenter collection of MRSA and MSSA strains</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>Denmark</td>
<td>34</td>
<td>Human</td>
<td>Nasal carrier</td>
<td>26</td>
</tr>
<tr>
<td>5</td>
<td>The Netherlands</td>
<td>30</td>
<td>Human</td>
<td>MRSA</td>
<td>28</td>
</tr>
<tr>
<td>6</td>
<td>France (mainly), USA, UK, Japan</td>
<td>71</td>
<td>Bovine</td>
<td>Mastitis isolates</td>
<td>2</td>
</tr>
</tbody>
</table>

<sup>a</sup> USA, United States; UK, United Kingdom.

<sup>b</sup> MRSA, methicillin-resistant *S. aureus*; MSSA, methicillin-susceptible *S. aureus*.

BT. The genetic difference between the *S. aureus* populations of bovine origin and those of human origin was analyzed in more detail by BT, a procedure involving strain-differentiating DNA probes which were initially identified on the basis of RAPD analysis of a large collection of *S. aureus* strains isolated from humans (28). The method was proved to be able to elucidate overall genetic polymorphisms and clonal relatedness among *S. aureus* strains and was extensively described elsewhere (30). In brief, 12 cloned strain-specific DNA fragments (AW1 to AW9, AW11, AW14, and AW15) were used as probes and were hybridized to total genomic DNA of the 71 strains isolated from cows with mastitis previously analyzed by agr RE-PCR. Hybridization of the probes was scored with a 1 or a 0 according to the presence or absence of a hybridization signal, resulting in a 12-digit binary code for each *S. aureus* isolate.

All bovine mastitis isolates but two were typeable by the BT method. All typeable isolates were found to hybridize with at least 3 and up to 11 of the tested probes. Probe AW9 hybridized to all typeable bovine mastitis isolates and, as suggested by a previous work, did not contribute to the discrimination power of the method for strains of bovine origin (31). BT was found to be very discriminatory and allowed the division of the collection of bovine mastitis isolates into 33 different binary types. All agr types comprising more than one strain could be subdivided with BT. The highly prevalent R III-A1 type was divided into 15 (16 if nontypeable strains were included) different binary types, whereas the following other two prevalent types, R IV-A5 and R IV-A7, were subdivided into eight and four BT types, respectively. The frequency of most BT types is only one or two strains, except for six BT types containing nine, eight, seven, five, three, and three strains (Fig. 3).

The binary types of the strains isolated from cows with mastitis were compared to the binary types of 50 strains isolated from humans, picked randomly among all strain collections analyzed by RE-PCR of the *agr* locus (28). These randomly selected strains represent all major agr types as mentioned before (20 strains of the R IV-A9 type, 12 strains of the R I'-A1 type, 6 strains of the R III-A2 type, 5 strains of the R III-A1 type, 3 strains of the R IV-A5 type, 2 strains of the R VI-A8 type, and 2 strains of the R V-A4 type). All of the agr types found among these human isolates were also subdivided by BT. The prevalent R I'-A1 and R IV-A9 types were divided into 11 and 16 BT types, respectively. Types R III-A2, R III-A1, R IV-A5, R VI-A8, and R V-A4 contain six, five, three,
two, and two different BT types, respectively (Fig. 3). No binary type common to strains of human and bovine origin was found.

Hierarchical clustering. Genetic relatedness between strains was analyzed by hierarchical clustering of binary types using the squared Euclidian distance of absolute value (weights of 0 and 1 were identical), and a dendrogram was constructed (SPSS 9.0 program for Windows software) (Fig. 3). The overall clonality of all strains (except for one) based on BT results was shown by their clustering into one group for a relative genetic similarity of 79%. At a genetic similarity of 83%, three clusters, which could be linked to host specificity, were defined (Fig. 3, roman numerals). Two human isolates (RIVM IV and K2-02) do not belong to any of these three clusters. The majority (84%) of the strains isolated from humans were found in clusters I and III. Both clusters also contained four strains of bovine origin (strains 125 and 130 in cluster I; strains 23 and 57 in cluster III). One of us previously reported that strains 125 and 130, isolated in Japan by Takeuchi et al., harbor a distinct agr locus. The receptor-encoding genes (agrC) of strains 125 and 130 are highly similar to those of the agr group 1 strains, whereas their propeptide-encoding genes (agrD) are highly similar to those of groups 2 and 3, respectively. Both strains are classified in the agr group 1 by the restriction map analysis method but either in group 2 (strain 125) or in group 3 (strain 130) by the multiplex PCR method. Since the propeptide and its receptor belong to different interference groups, these two strains should be impaired in the activation of RNA III by the agr system. These strains are probably in the process of evolutionary development, moving from agr group 1 to groups 2 and 3, respectively (2). Cluster II contains 92% of the strains of bovine origin and also includes six human isolates.

Identical to the agr typing results, hierarchical clustering of binary types confirms that the strain populations of bovine origin and those of human origin are genetically different. Bacterial host specificity was also described by other studies using different typing techniques to analyze the genetic background of S. aureus populations (8, 31). Diverse studies showed divergent data for prevalence of S. aureus strains within the different agr interference groups (5, 12, 15, 20–22), while we showed that the agr interference group I has spread at the expense of the other agr groups, both in the population of strains isolated from cows as well as in the population of strains isolated from humans. When allelic variations within agr groups were analyzed, a more precise scheme emerged: human isolates with the allelic variant R I’-A1 and R IV-A9 and bovine isolates with the allelic variant R III-A1 were predominant (Fig. 2). If the allelic variants R I’-A1 and R III-A1 belong to agr group 1 and are not very divergent (they were differentiated by the presence or absence of one restriction site), the RIV-A9 variant belongs to agr interference group 2 (2). This indicates that in the human population, not only has agr interference group 1 spread at the expense of the other agr groups, but also one particular allelic variant of the agr group 2 has spread at the expense of other groups. BT, detecting DNA elements with a relative high evolutionary clock speed, reflects changes arising in other parts of the genome of each of these agr restriction types. Both methods produced results con-
FIG. 3. Dendrogram showing the clustering of 50 unrelated human S. aureus isolates and 69 unrelated bovine S. aureus isolates on the basis of hybridization scores after BT with 12 DNA probes. Isolates numbers, agr types, agr groups, binary codes, and origins of strains are shown for all isolates. Roman numerals indicate the three main clusters of strains defined at a genetic similarity of 83% (arrow). The scale indicates the level of genetic relatedness within this collection of strains.
sistent with the host specificity concept among S. aureus strains and showed that the transfer of strains between humans and cows is a possible but infrequent event. We believe that all strains are putatively able to infect one or the other host, but our typing results indicate that most of them are much more adapted for the infection of one host versus the other. At this moment, the underlying genetic basis for the variation among S. aureus strains from both populations is unknown, but at least two hypotheses can be conceived: (i) the agr allele of a strain is directly implicated in the regulation of the expression of genes implicated in the tropism of the strain for a particular host; (ii) the agr allele is not directly involved in host specificity, but it reveals the overall genetic evolution of S. aureus after its adaptation (host-adapted lineage[s]) to one or the other host by the acquisition (or the loss) of some virulence genes and/or pathogenicity islands. The University of Minnesota is sequencing the entire genome of a common bovine S. aureus strain. A comparative genomic analysis of a shotgun sequence of around 10% of the genome of this bovine mastitis-associated clone and the genomes of the recently sequenced human-associated strains Mu50 and N315 has already identified numerous unique sequences (3, 11). A recent study also identified a novel putative pathogenicity island, SaPIbov, in the genome of this bovine isolate (1). For these reasons it can be expected that the identification of lineages dominant in a particular host or disease, and further comparative genomic analysis will provide important insights regarding disease pathogenesis and lead to the development of new prophylactic and diagnostic tools.

This work was supported by a grant (AIP P00060, P00223) from the French association Bureau des Ressources Génétiques. We thank Martine Braibant for critical reading of the manuscript.

REFERENCES