# Mutation Analysis for *AgrC* from *Staphylococcus aureus*

Samuel da Silva Guimarães, Danielle Mendes Silva, Mônica Pacheco Silva, Pedro Marcus Vidigal, Andréa de Oliveira Barros Ribon, Sabrina de Azevedo Silveira

*Universidade Federal de Viçosa*

*Staphylococcus aureus* is one of the main pathogens of bovine mastitis. The analysis of SNPs on the genomes of four strains of S. aureus associated with mastitis showed the presence of 6 variations in 5 positions in the sequence of the *agrC* protein. Thus, it was hypothesized that these variations could be related with different manifestations of the disease. We searched for the reference sequence of *agrC* in Protein Data Bank (PDB) and found the entry 4BXI.A, which comprehends just part of the sequence, with 153 residues (278-430). A new search was made in the PDB to find structures similar to 4BXI.A, resulting in a set of 82 different entries, grouped by 40% of sequence similarity. A pairwise structural alignment was performed using the MultiProt to align each of the structures against 4BXI.A. A visual representation for this alignment was generated using the CINEMA color scheme, so that each sequence is represented by a line and each column corresponds to an alignment position. Also, we used the Expectation Maximization (EM) algorithm to group similar sequences, so that these similar sequences appear next to each other, which helps the user to detect trends and exceptions in the data. Next, the interactions were modeled as graphs in which nodes represent residues and edges represent interactions between residues. To calculate the interactions we used the Voronoi diagram followed by the Delaunay triangulation and we labeled nodes as positively charged, negatively charged, aromatic, hydrophobic, donor or acceptor. The edges were labeled according to a distance criteria and the type of edges as hydrogen bond, aromatic stacking, hydrophobic, repulsive and salt bridge. From the set of proteins modeled as graphs, some centrality measures that are commonly used in complex networks were calculated using the iGraph package from R, each of them providing a different perspective of centrality. This strategy enabled us to identify 2 positions on the *agrC* sequence were SNPs can potentially impact on protein structure and should be further studied to evaluate possible connections to bacterial virulence mechanisms.

Financial support: CAPES, CNPq, FAPEMIG.