

DANIELE FERREIRA DA SILVA

**FILOGENIA MOLECULAR E GENÔMICA COMPARATIVA DE BACTÉRIAS
GRAM-POSITIVAS DO TRATO GASTROINTESTINAL**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Microbiologia Agrícola, para obtenção do título de *Doctor Scientiae*.

VIÇOSA
MINAS GERAIS – BRASIL
2007

RESUMO

SILVA, Daniele Ferreira, D.Sc., Universidade Federal de Viçosa, dezembro de 2007. **Filogenia molecular e genômica comparativa de bactérias Gram-positivas do trato gastrointestinal.** Orientadora: Célia Alencar de Moraes. Co-orientadores: Arnaldo Chaer Borges, Cosme Damião Cruz e Marisa Vieira de Queiroz.

A divulgação recente das informações do genoma completo de bactérias do ácido láctico probióticas e do de *Listeria monocytogenes*, juntamente com as da presença de grandes blocos de genes de virulência, denominados de ilhas de patogenicidade (PAIs) em *L. monocytogenes*, trouxe à tona o questionamento sobre a possível existência de seqüências comuns aos dois grupos, probióticas e enteropatogênicas, e de genes similares que somente são ativados durante a passagem pelo trato gastrointestinal. O presente trabalho teve como objetivo avaliar por meio de genômica comparativa as semelhanças e diferenças no conteúdo de genes que codificam proteínas de virulência em *Listeria monocytogenes* EGD-e e os de probiose em *Lactobacillus* probióticos, além de identificar e caracterizar o gene de resistência ao estresse *clp* em *Lactobacillus delbrueckii* UFV H2b20. As seqüências dos genes selecionados foram capturadas no banco de dados do National Center for Biotechnology Information (NCBI) e para cada gene foi feita uma busca de homologia, com o uso do programa BLASTp, e de ortologia com a utilização do programa Clusters of Orthologous Groups (COG). Os genes

homólogos e ortólogos foram selecionados de genomas de *Listeria innocua* CLIP1162, *Lactobacillus acidophilus* NCFM, *Lactobacillus plantarum* WCFS1, *Lactobacillus johnsonii* NC 533, *Lactobacillus gasseri* ATCC33233, *Lactobacillus salivarius* subsp. *salivarius* UCC118, *Lactobacillus reuteri* F275, *Lactobacillus delbrueckii* subsp. *bulgaricus* ATCC BAA-365 e *Lactobacillus delbrueckii* subsp. *bulgaricus* ATCC 11842. A análise das árvores filogenéticas reconstruídas com base nos genes de maior interesse reforça a hipótese de que uma comparação paralela pode ser traçada entre a resposta patogênica e a probiótica. As LABs apresentam genes que são de fundamental importância para a probiose e que estão ausentes no patógeno *L. monocytogenes*, a exemplo da maioria dos genes envolvidos no mecanismo de adesão. Foi confirmada a existência do gene *clp*, em *L. delbrueckii* UFV H2b20, com homólogos referente ao complexo proteolítico Clp apresentando alta porcentagem de identidade e similaridade. Dentre os melhores homólogos está o gene *clpC* presente em vários lactobacilos. Vários fagos positivos foram identificados no banco genômico de *L. delbrueckii* UFV H2b20. Possivelmente, mais de uma subunidade desse complexo proteolítico foi identificada, uma vez que a sonda se hibridizava com várias dessas subunidades. A confirmação da presença e análise de função de outros genes identificados nos genomas das bactérias probióticas, em *L. delbrueckii* UFV H2b20, por estudos de genômica funcional, demandará análise comparativa de funções desses genes em outras estirpes probióticas, para justificar a importância de *L. delbrueckii* UFV H2b20 na saúde e bioprocessamento de alimentos.

ABSTRACT

SILVA, Daniele Ferreira, D.Sc., Universidade Federal de Viçosa, December of 2007. **Molecular phylogeny and comparative genomics of Gram-positive bacteria of the gastrointestinal tract.** Adviser: Célia Alencar de Moraes. Co-advisers: Arnaldo Chaer Borges, Cosme Damião Cruz and Marisa Vieira de Queiroz.

Recent information on the complete genome of probiotic lactic acid bacteria and *Listeria monocytogenes*, and the elucidation on those caused on large clusters of virulence genes, denominated pathogenicity islands (PAIs) present in *L. monocytogenes*, have brought the attention to the likely existence of gene sequences common to both groups, probiotic and enteropathogenic bacteria, as well as of similar genes that are only activated in the gastrointestinal tract. This work aimed to evaluate by means of comparative genomics the similarities and differences among genes encoding virulence proteins in *L. monocytogenes* EGD-e and those encoding probiotic functions in probiotic *Lactobacillus*, and also to isolate and characterize the *clp* stress resistance gene in *Lactobacillus delbrueckii* UFV H2b20. The genes sequences were captured from the databank of the National Center for Biotechnology Information (NCBI) and for each gene a homology search was performed using the program BLASTp, and an orthology search, using the program Clusters of Orthologous Groups (COG). The homologous and orthologous genes were selected from genome sequences of *Listeria innocua* CLIP1162, *Lactobacillus*

acidophilus NCFM, *Lactobacillus plantarum* WCFS1, *Lactobacillus johnsonii* NC 533, *Lactobacillus gasseri* ATCC33233, *Lactobacillus salivarius* subsp. *salivarius* UCC118, *Lactobacillus reuteri* F275, *Lactobacillus delbrueckii* subsp. *bulgaricus* ATCC BAA-365 and *Lactobacillus delbrueckii* subsp. *bulgaricus* ATCC 11842. The analysis of the phylogenetic trees reconstructed based on higher interest genes reinforces the hypothesis that a parallel may be made to some extent between the pathogenic and the probiotic responses. The probiotic LABs present genes fundamentally important for probiosis that are absent in the pathogen *L. monocytogenes*, as are most of the genes involved in the mechanism of adhesion. The existence of the *clp* gene was confirmed in *L. delbrueckii* UFV H2b20, with the homologous genes encoding proteins of the proteolytic complex Clp presenting high percentage of identity and similarity. Among the best homologous genes is the gene *clpC* present in various lactobacili. From a *L. delbrueckii* UFV H2b20 library, several phages were tested positive for *clp*. Possibly, more than one subunit of this proteolytic complex was identified, as the DNA probe hybridized with several of these subunits genes. Confirmation of the presence in *L. delbrueckii* UFV H2b20 and functional analysis of other genes already identified in probiotic LABs genomes will demand a comparative analysis of the functions of such genes in other probiotic strains, to justify the importance of *L. delbrueckii* UFV H2b20 in health and for food bio-processing.