TITLE: Preliminary study of the absolute quantitation of *Escherichia coli* (EHEC) O157:H7 proteome

AUTHORS: Silva, W.M1; Amigo, N1.; Chen, Z3, Zhang, Q3, Amadio, A2.; Larzabal, M1.; Bei J. 3.; Cataldi, A1.

INSTITUTION: 1 - Institute of Biotechnology, CICVyA, National Institute of Agricultural Technology. Hurlingham, Buenos Aires, Argentina. 2 - Rafaela Experimental Station, National Institute of Agricultural Technology. Rafaela, Santa Fe, Argentina. 3 - AGRO-Biological Gene Research Center, Guangdong `Academy of Agricultural Sciences (GDAAS), Guangzhou, China.

ABSTRACT

Enterohemorrhagic *Escherichia coli* (EHEC) O157:H7 is a human pathogen responsible for diarrhea, hemorrhagic colitis and hemolytic uremic syndrome (HUS). The infection by EHEC is widely distributed in the world and numerous outbreaks of enterohemorrhagic disease were reported. In industrialized countries HUS is reported as a disease of low incidence. On the other hand, Argentina is considered the country with the highest HUS incidence in the world. The primary reservoirs of EHEC are cattle. Proteomic studies show that absolute quantitation of proteins present within a complex protein mixture is extremely important for understating physiological adaptations in response to biological demands promoted by environmental changes. Thus in this study, to promote the more comprehensive insight into the molecular basis of EHEC O157:H7 physiology and pathogenesis, we applied high-throughput proteomics, using iTRAQ-based quantitative proteomic analysis to estimate the absolute abundance of EHEC proteome. EHEC O157:H7 strains Rafaela II (clade 8) and Anguil 7.1 (clade 6) isolated from cattle in Argentina and the standard EDL933 (clade 3) strain recovered from an hamburguer in USA were grown in Dulbecco's modified Eagle's medium to exponential phase (OD_{600nm} = 0.6) at 37°C under a 5% CO_2 atmosphere. Two biological replicates of whole bacterial lysates of each strain were enzymatically digested with trypsin, pooled and posteriorly labeled with 6-plex iTRAQ reagents. Finally the tryptic fragments generated were subjected to proteomic analysis by 2D-LC MS/MS analysis. From this proteomic analysis, we identified 2,234 non-redundant protein of EHEC O157:H7. When compared this result with *in silico* data of EHEC O157:H7 genome was characterized approximately 40% of the proteome predicted this pathogen. Kegg enrichment analysis revealed that Carbon metabolism, Purine metabolism and Microbial metabolism in diverse environments were the three most enriched. To determine the abundance of the identified proteins, we used the emPAI approach. According this analysis the more abundant proteins are related to metabolism cellular, translation, motility and stress response. On the other hand the less abundant are related to DNA process, cell respiration and prophage. Among the proteins that composed the Type III Secretion System the more abundant protein was EspA protein. The result obtained to date show important proteins that contribute to physiology and pathogenicity of EHEC O157:H7.

Keywords: EHEC, iTRAq, Absolute protein quantitation, bacterial proteomic