No. 19

9th annual meeting of whole-cell project of *Thermus thermophilus* HB8
« Towards system Biology »,
Riken Institute Hyogo, Japan

Proteogenomics of bacteria: from genome annotation to physiology

Jean Armengaud

CEA, DSV, iBEB, Lab Biochim System Perturb, Bagnols-sur-Cèze, F-30207, France
jean.armengaud@cea.fr

High-quality genome annotation is a crucial starting point for systems biology. Proteogenomics consists of high-throughput identification and characterization of proteins by extra-large shotgun MS/MS approaches and the integration of these data with genomics data. We applied such strategy on *Deinococcus deserti* VCD115, a bacterium from the *Thermus-Deinococcus* phylum that was isolated from surface sands from Sahara. This bacterium exhibits an extraordinary ability to withstand desiccation and ionizing radiations. We sequenced and annotated its genome by proteogenomics. We systematically investigated the existence of remaining unannotated genes, defined translational start sites, and listed signal peptide processing events and post-translational modifications. Surprisingly, experimental evidences were obtained indicating that DnaA (the protein involved in the DNA replication initiation process) and RpsL (the S12 ribosomal conserved protein) translation is initiated in *Deinococcaceae* from non-canonical codons (ATC and CTG, respectively). Our proteomic results demonstrate that predicting translation initiation codons is still difficult for some bacteria and that proteomic-based refinement of genome annotations may be helpful in such cases. Advantageously, the same experimental proteomic data sets may be used to characterize the specific metabolic traits of the organism under study. We analyzed the proteome dynamics of *D. deserti* after exposure to 3 kGy γ radiations and during rehydration following a 3 weeks period of desiccation. Identification of thousands of peptides corresponding to hundreds of proteins was systematically done by mass spectrometry on tens of samples. Proteins were semi-quantified by spectral counting. Several proteins were found overproduced in the first hours after stress and could result important for DNA repair.

Références :

(1) Armengaud J. (2010) Proteogenomics and systems biology: quest for the ultimate missing parts. *Expert Rev. Proteomics* 7(1), *65-77*.
(2) Baudet M, Ortet P, Gaillard J-C, Fernandez B, Guérin P, Enjalbal C, Subra G, de Groot A, Barakat M, Dedieu A, Armengaud J (2010) Proteomic-based refinement of *Deinococcus deserti* genome annotation reveals an unwonted use of non-canonical translation initiation codons. *Mol. Cell. Proteomics* 9(2), 415-426.
(3) Armengaud J. (2009) A perfect genome annotation is within reach with the ProteOMICs and GenOMICs alliance. *Curr. Opin. Microbiol.* 12*, 292-300*.
(4) De Groot A, Dulermo R, Ortet P, Blanchard L, Guérin P, Fernandez B, Vacherie B, Dossat C, Jolivet E, Siguier P, Chandler M, Barakat M, Dedieu A, Barbe V, Heulin T, Sommer S, Achouak W, Armengaud J. (2009) Alliance of proteomics and genomics to unravel adaptation of Sahara bacterium *Deinococcus deserti* to harsh environmental conditions. *PLOS Genetics* 5, e10000434.