



CENTRE OF BIOLOGICAL ENGINEERING
UNIVERSIDADE DO MINHO



INSTITUTE FOR BIOTECHNOLOGY AND BIOENGINEERING

Exploring current knowledge on *Escherichia coli* stringent response via text mining

Anália Lourenço, Sónia Carneiro, Isabel Rocha, Eugénio C. Ferreira

IBB-Institute for Biotechnology and Bioengineering, Centre of Biological Engineering, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, Portugal

Automated extraction systems have become a time saving necessity in Systems Biology [1]. The goals of the present work were set into the development of automated extraction system prototype for the compilation of knowledge about the *E. coli* stringent response and the evaluation of its potential value in augmenting curation throughput and efficiency.

An analysis based on the statistical co-occurrence of biological concepts on textual annotations highlighted the entities that are most likely to be biologically engaged during this response [2]. The co-occurrence of the (p)ppGpp nucleotides with several gene products suggested that besides enzymes RelA and SpoT, which control the basal levels of (p)ppGpp nucleotides, many other proteins have a key role in this response. In turn, the identification of less annotated concepts revealed that although some (p)ppGpp-induced functional activities are still overlooked, they are involved in this stress response as well. Additional functional enrichment analysis drew out the involvement of specific stress responses during the stringent response and sustained a retrospective analysis of the experimental techniques supporting such studies.

Although preliminary, these results showed that it is possible to scale-up conventional manual model curation to cope with the ever-increasing publication rate and, at the same time, provide automatic means of identifying and contextualising entities of interest. Beyond the accomplishments of the prototype on the study of *E. coli* stringent response, its extension to the analysis of other stress responses and organisms is fairly easy and interesting.

References

- [1] Kowald A, Schmeier S, "Text mining for systems modelling", *Methods Mol Biol* (2011), 696: 305-318.
- [2] Carneiro S, Lourenço A, Ferreira EC, Rocha I (2011) Stringent response of *Escherichia coli*: revisiting the bibliome using Literature Mining. *Microbial Informatics and Experimentation* 1:14 (accepted).