

# A team's story in the IMPROVER Species Translation Challenge

Adi L. Tarca<sup>1,2</sup> and Roberto Romero<sup>2</sup>

1. Wayne State University, Detroit, MI, USA;
2. Perinatology Research Branch, Perinatology Research Branch, Eunice Kennedy Shriver National Institute of Child Health and Human Development, NICHD/NIH, USA;

**Keywords:** gene expression, protein phosphorylation, pathway analysis, crowdsourcing, machine learning

The IMPROVER [1] Species Translation Challenge (STC) [2], organized by IBM Research and Philip Morris International, tested the limits of the fundamental assumption that underpins the use of animal models for gaining insight into human biology, i.e., that there is conservation in the nature of the responses to injury and therapy. Gene expression and protein phosphorylation data measured in rat and human cells after treatment with 26 stimuli were made available at <https://www.sbvimprover.com>. International teams were challenged to 1) use rat gene expression to predict rat protein phosphorylation, 2) use rat gene expression and protein phosphorylation to predict human protein phosphorylation, and 3) use rat gene expression to predict pathway activity in human. Teams were ranked based on performance on a test dataset generated from other 26 stimuli.

We participated in these challenges relying on good old friend *R* and *Bioconductor* packages to process data and build prediction models to address challenges 1 and 2, and derive a data driven homology between rat and human genes to tackle challenge 3. Our team's rank in these challenges was first, second, and third for sub-challenges 1, 3, and 2 respectively.

In this talk we will present the methods we have used in this crowdsourcing effort and the R code that we have developed to implement these methods. Lessons learned about the translatability of findings from rat to human and about best strategies to modeling sparse outcomes in STC will also be discussed.

1. Meyer P et al. (2011) Verification of systems biology research in the age of collaborative competition. *Nat.Biotechnol.* 29(9):811-815.
2. Rhrissorrakrai, K. et al. (2014) Understanding the limits of animal models as predictors of human biology: lessons learned from the sbv improver species translation challenge. *Bioinformatics*, in press.