



www.sbvimprover.com

Lessons Learned from The Species Translation Challenge

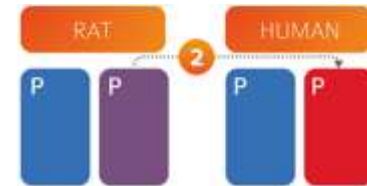
Pablo Meyer
Computational Biology Center
IBM Research

29 October 2013

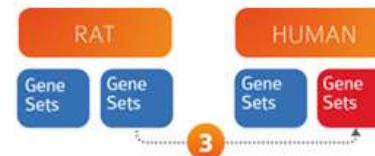
- Similarities between Human and Rat responses
- Splitting the data between training/test set
- A diverse set of methodological approaches
- Results of translating 16 phosphoprotein activities from Rat to Human
- Results of translating 246 gene set activities from Rat to Human
- Conclusions

Central questions of the Species Translation Challenge

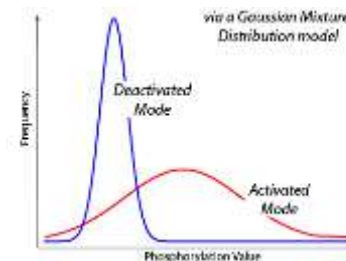
- Can the phosphoproteomic responses of stimuli in human cells be predicted given the responses generated with the same stimuli in rat cells?



- Which gene expression regulatory processes (biological pathways) are translatable and therefore predictable across species, and which are too divergent?



- How does the accuracy/precision of the prediction depend on the nature of the applied perturbation?



Training, 26 stimuli + DME

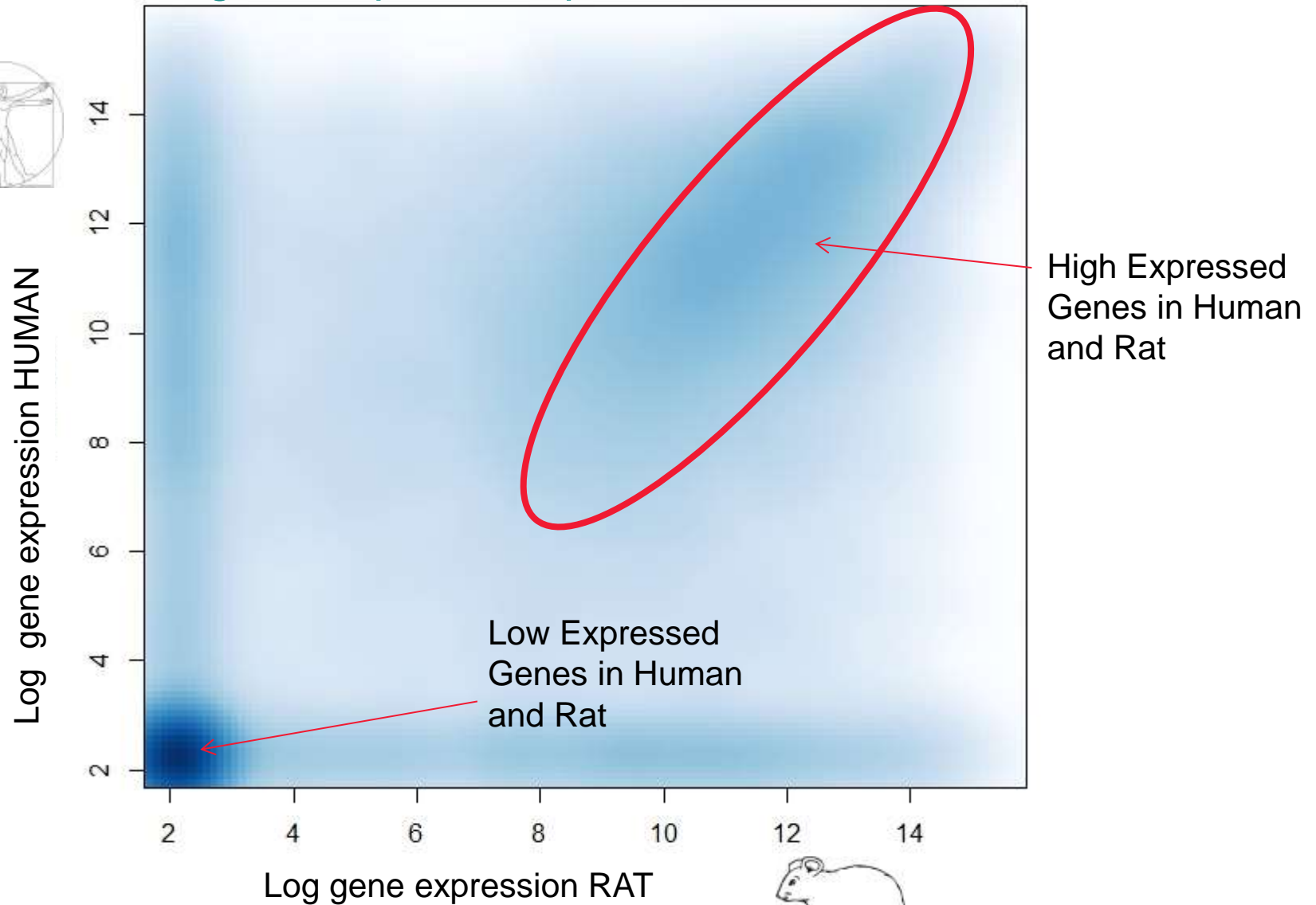
- 5AZA:2
- AMPHIREGULIN:2
- BETAHISTINE:3
- BISACODYL:4
- CHOLESTEROL:4
- CLENBUTEROL:1
- DME:1
- DME:2
- DME:3
- DME:4
- EGF:3
- FLAST:2
- FGF8:3
- FORSKOLIN:2
- HIGHGLU:4
- IFNG:2
- IGFII:1
- IL4:3
- MEPYRAMINE:3
- NORETHINDRONE:2
- ODN2006:4
- PDGFB:3
- PMA:4
- PROKINECITIN2:3
- PROMETHAZINE:2
- SEROTONIN:1
- SHH:1
- TGFA:4
- TNFA:1
- WISP3:1

Test, 26 stimuli

- BENPERIDOL:1
- BETAXOLOL:2
- C12IE:1
- CARBACHOL:4
- CCL3L:1
- CISAPRIDE:4
- CLOMIPRAMINE:2
- CNTF:3
- DEXAMETHASONE:3
- DIMETHYLOXALYLGLYCINE:2
- EPIGALLOCATECHIN:2
- FORMALDEHYDE:4
- FSL1:2
- HBEGF:1
- HEXESTROL:2
- IL11:1
- IL1B:3
- INS:3
- NAACL:4
- NISOXETINE:1
- NT3:2
- POLYIC:4
- RAPAMYCIN:3
- RILUZOLE:4
- TAUROCHOLIC_ACID:3
- TOLMETIN:1

Batches are indicated by: 1, 2, 3 or 4

Density plot showing Species Similarity between Human and Rat gene expression profiles for TGF α



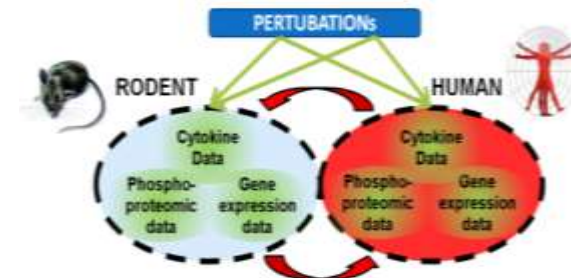
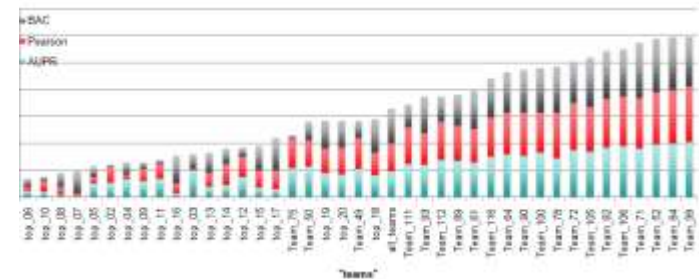
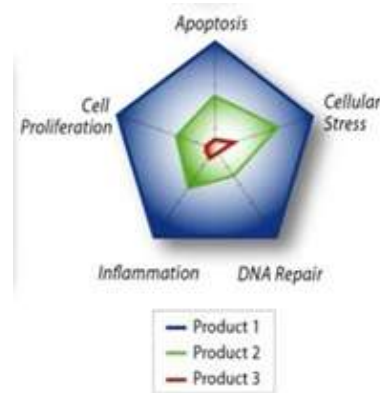
A variety of classification algorithms were used across all sub-challenges

- Support Vector Machine 7 participants
- Regression based methods 14 participants
- Decision Trees/Random Forest 8 participants
- Neural Networks 4 participants
- Bayesian approach 3 participants

Setting up the species translation problem as a challenge allowed us to:



- Determine Stimuli Predictability
- Determine Biological translatability at different molecular levels (pathways, functions, processes)
- Determine Best approach/ Combined approaches
- Evaluate the experimental approach used in the context of translatability.



Conclusions of the Species Translation Challenge

- Participants were able to predict consistently better than what expected from species similarity
- Participants were better at predicting the activities of pathways/processes than stimuli activations
- Participants were better at predicting certain pathways/processes
- Participants were better at predicting the effect of certain stimuli

Thank you for your attention

The sbv IMPROVER project and www.sbvimprover.com are part of a collaboration designed to enable scientists to learn about and contribute to the development of a new crowd sourcing method for verification of scientific data and results. The project team includes scientists from Philip Morris International's (PMI) Research and Development department and IBM's Thomas J. Watson Research Center. The project is funded by PMI.