GALAHAD – A WEB SERVER FOR GENE EXPRESSION DATA ANALYSIS IN SUPPORT OF DRUG DEVELOPMENT

Griet Laenen[1,2,*], *Amin Ardeshirdavani*[1,2], *Yves Moreau*[1,2] & *Lieven Thorrez*[1,3].

Dept. of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven[1] ; *iMinds Future Health Dept., KU Leuven*[2] ; *Dept. of Development and Regeneration @ Kulak, KU Leuven*[3].

[*]*griet.laenen@esat.kuleuven.be*

With many candidate drugs failing in the late clinical stages, drug development is a high-risk business. In order to reduce the high rates of attrition, a better knowledge of a drug's mechanism of action is required. Hence we have developed Galahad, a web server for the analysis of gene expression data following drug treatment, aimed at predicting a drug's molecular targets and physiological effects.

INTRODUCTION

The pharmaceutical industry is facing unprecedented productivity challenges. Attrition rates have risen sharply, especially in late-phase clinical trials. With safety and efficacy being the main bottlenecks, a better knowledge of a candidate drug's mode of action and its off-target effects could be of substantial value to drug development. DNA microarray technology enables a genome-wide analysis on the transcriptional response to a compound treatment, and thus can provide valuable information for identifying the compound-protein interactions and resulting effects prior to clinical trials. In addition, this information may also be useful for already marketed drugs, in the light of drug repositioning.

METHODS

We have developed a new, easy-to-use web server called Galahad, for the in-depth exploration of a drug's mode of effect based on gene expression changes following treatment. Our software provides the main tools needed for gaining new insights into the biological effects of a drug by combining

- **preprocessing** of gene expression data obtained from different Affymetrix array types;
- **quality assessment and exploratory analysis** of these data to ascertain data quality, uncover experimental issues or sample mix-us, and help in deciding whether certain arrays need to be considered as outlying;
- **differential expression analysis** to determine the significance of gene up- or down-regulation following drug treatment by fitting a linear model to the expression data for each gene;
- genome-wide **drug target prioritization** by means of an in-house developed algorithm for network neighborhood analysis integrating the expression data with functional protein association information[1];
- prediction of **Reactome pathways** involved in the drug's mode of effect;
- identification of associated **disease phenotypes** from the Human Phenotype Ontology enabling side effect prediction and drug repositioning.

RESULTS & DISCUSSION

All of the above functionalities can be demonstrated on gene expression data for treatment with drugs exhibiting a well-defined mechanism of action. One such drug is infliximab, a tumor necrosis factor (TNF)-binding monoclonal antibody marketed under the brand name Remicade and used in the treatment of several autoimmune diseases. Infliximab target prioritization based on the gene expression profiles from eleven Crohn's colitis patients treated with this drug[2] ranks the target TNF in the top 1%, although not differentially expressed following treatment. Enrichment analysis on the significantly up- and down-regulated genes returned several immune pathways, as well as links to disease phenotypes reported in literature either as a known indication, a side effect, or a possibility for repositioning. By application to a larger set of well-characterized chemical drugs, Galahad will now be further optimized.

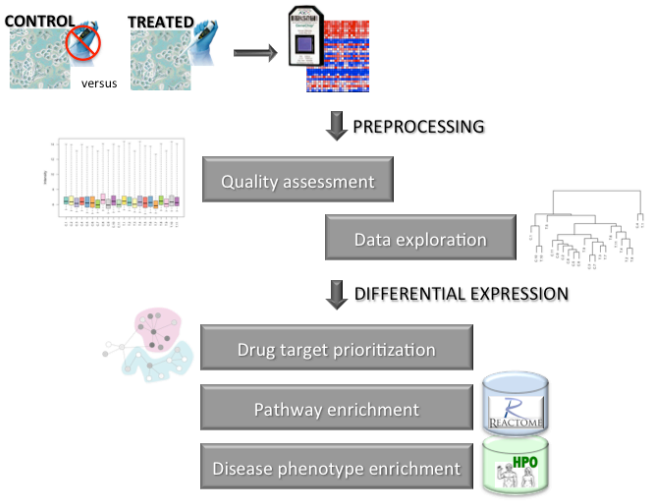

***FIGURE 1**. Overview of analyses provided by Galahad.*

REFERENCES

1. Laenen G. *et al. Mol BioSyst* **9**, 1676-1685 (2013).
2. Arijs I. *et al. Am J Gastroenterol* **106**, 748-761 (2011).