CALIFORNIA STATE SCIENCE FAIR
2014 PROJECT SUMMARY

Name(s)

Swetha Revanur

Project Number

34695

Project Title

Factomics: A Cloud-enabled Web Portal Incorporating Gene Expression and GWAS Facilitating Disease Causation Analysis

Abstract

Objectives/Goals

Researchers and medical professionals are inundated with tremendous amounts of genomic #big data#. However, there is a lack of tools that are capable of offering insightful interpretations of this raw data. To address this void, I developed Factomics, a cloud-enabled web portal, which provides a built-in workflow for performing disease correlation/causation analysis based on integrative genomics data. Results from my workflow can drive new hypotheses.

Methods/Materials

The Factomics workflow is divided into 4 distinct phases # Launch, Discovery, Correlation and Causation. These phases take the user from an initial set of diseases to their candidate causal single nucleotide polymorphisms (SNPs), genes, and pathways based on gene expression data and industry standard analytical modules.

Factomics is organized in a 3-tier architecture. The front-end is written in Google Apps Script (GAS; server-side JavaScript). The middle-tier is integrated with the analytical modules and with public data repositories including NCBI databases. Analytical findings are stored on Google Drive, the third tier.

Results

I demonstrated a use-case of Factomics with Alzheimer#s disease, Type 2 Diabetes Mellitus, Ovarian Cancer, and Pancreatic Cancer. Gene expression data used for each of these diseases enabled the identification of missing gene to disease and gene to pathway associations. Linkage disequilibrium analysis identified non-synonymous (deleterious) and regulatory candidate causal SNPs for Alzheimer#s Disease and Ovarian Cancer. A multi-dimensional view of these diseases showed several overlapping upregulated genes and pathways. Some of the findings were corroborated by literature, and others were novel. This information can be used to optimize drug development, drug repositioning and diagnostic tools.

Conclusions/Discussion

My portal bridges the gap between current tools, and is capable of elucidating the biological mechanism for any disease. Moreover it is scalable across a wide range of fields (bench biology, bioinformatics, pharmaceutics, and in clinical settings), while providing cloud-capabilities and a user-friendly interface. To the best of my knowledge, there is no existing tool offering similar functionality.

Summary Statement

I developed a novel cloud-enabled web portal that can computationally identify candidate causal SNPs, genes and pathways for any disease.

Help Received

Dr. Susan Lato of Codexis reviewed my work. Mr. David Walz is my teacher sponsor.