Novel Meta-Analyses of Microarray Data First Step Toward Malaria Disease Signature Development

Mohamad Omar Hadied, BS[1], Osama El-Sayed, BS[2], Bilal Zaidi, BS[3], Jihad Al-Jabban, BS[4], Shuaib Raza, BS[5], Imad Al-Jabban, BS[4], James Pan, BA[6], Tej Azad, BA[6] and Dexter Hadley, MD/PhD[7]

(1)Wayne State University School of Medicine, Detroit, MI, (2)College of Medicine, University of Illinois at Chicago, Chicago, IL, (3)University of Michigan, Ann Arbor, MI, (4)Department of Immunology, Harvard University, Boston, MA, (5)Yale School of Medicine, Yale University, New Haven, CT, (6)Stanford University, Stanford, CA, (7)University of California San Francisco, San Francisco, CA

University of California San Francisco

School of Medicine

Mohamad Omar Hadied
1128 Martha Reeves Dr
Detroit, Michigan
mhadied@med.wayne.edu
(248) 330-8959

Abstract

Background:
Malaria accounts for a considerable amount of mortality around the world, especially among young children. Artemisinin Combination Therapies (ACTs) remain the most potent treatment for uncomplicated cases of malaria. Early signs of drug resistance have emerged in the form of delayed clearance of the parasite in Southeast Asia. There is a high demand for newer drugs and vaccines to help control the worldwide malaria burden pending complete drug resistance to ACTs. The aim of this study is to characterize the disease signature of a malaria infection using meta-analysis of microarray data, and use this data to facilitate the pursuit of newer drugs and vaccines to fight malaria worldwide.

Methods:
Through a unique collaboration between medical students and computer scientists in systems biology, Search Tag & Analyze Resource (STAR) is a platform that was created that allows us to easily analyze gene signatures through meta-analyses of series in the Gene Expression Omnibus (GEO). Experiment samples within the database are annotated with tags and run through meta-analysis.

Results:
Using STAR, we generated one preliminary gene signature using samples on GEO: whole blood samples from malaria infections versus blood samples from health controls. We have found LGR4 and p70-S6k to be the top upstream regulators (p=0.002 for both). GBA and C15orf48 were the top significantly up-regulated genes. LEF1, PDCD4, and CNST were the top significantly down-regulated genes. GBA, LEF1, and the p70-S6k pathway have all been reported before as associated with malaria disease expression, the other genes have not.

Conclusion:
These preliminary results are proof-of-concept to the potential of STAR. With the collaboration fully functional, thousands of samples will be added to improve the gene signature of malaria infection. The study will also include parasite gene expression run with the same meta-analysis. We report here a first step towards novel transitional opportunities for better biomarkers and drugs for infectious diseases like malaria and beyond.

Objectives

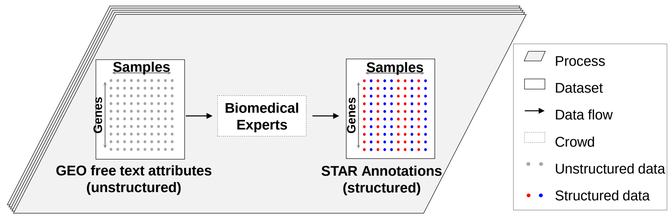
- Test the viability of the STAR system
- Use STAR to start building a gene expression signature of malaria infection

Background

Malaria accounts for a considerable amount of mortality around the world, especially among young children. Artemisinin Combination Therapies (ACTs) remain the most potent treatment for uncomplicated cases of malaria. Early signs of drug resistance have emerged in the form of delayed clearance of the parasite in Southeast Asia. There is a high demand for newer drugs and vaccines to help control the worldwide malaria burden pending complete drug resistance to ACTs. The aim of this study is to characterize the disease signature of a malaria infection using meta-analysis of microarray data, and use this data to facilitate the pursuit of newer drugs and vaccines to fight malaria worldwide.

Methods

Search Tag & Analyze Resource (STAR) is a platform created to allow more efficient analysis of microarray data in the National Center for Biotechnology Information's (NCBI) Gene Expression Omnibus (GEO). STAR allows users to search and tag similar experiments. These annotations are then run through meta-analysis to consolidate the results of all experiments tagged and therefore highlights genes that we significant across many similar experiments. A graphic of this process is provided below (Figure 1):



The STAR platform was utilized for gene expression during malaria infection. Whole blood samples from malaria infected individuals was tagged against whole blood samples from health controls in the same experiments. A total of 9 series and 590 samples were included in the analysis.

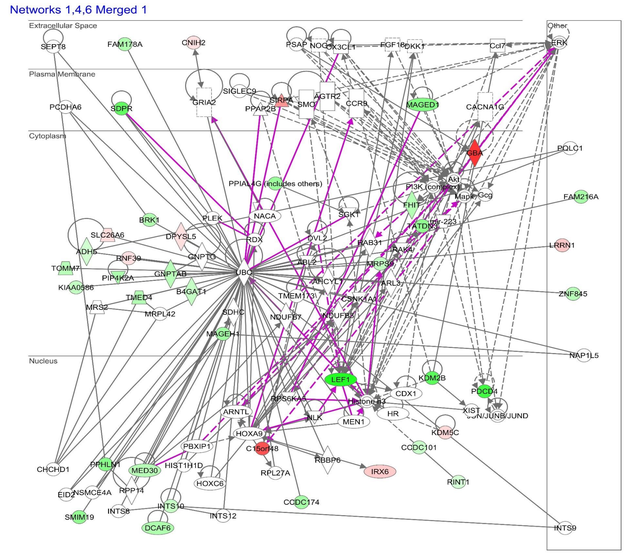
Results

We have found LGR4 and p70-S6k to be the top upstream regulators (p=0.002 for both). GBA and C15orf48 were the top significantly up-regulated genes. LEF1, PDCD4, and CNST were the top significantly down-regulated genes. GBA, LEF1, and the p70-S6k pathway have all been reported before as associated with malaria disease expression, the other genes have not. Series tags and sample numbers are shown below in Table 1.

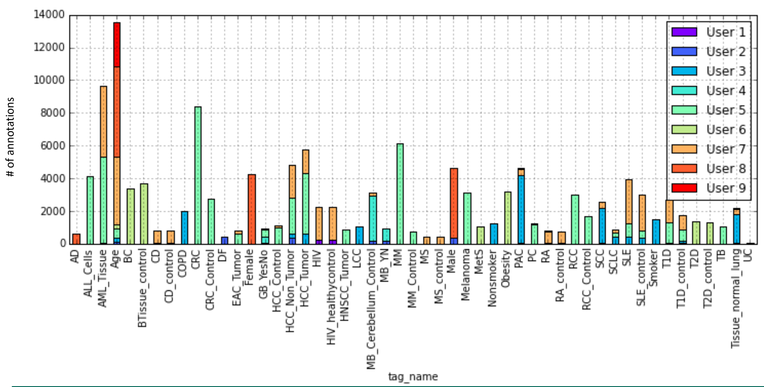
ID	Series	Platform	Tag	Samples
429	GSE15221	GPL6102	malaria	28
431	GSE16463	GPL6102	malaria	24
436	GSE34404	GPL10558	malaria	155
438	GSE7586	GPL570	malaria	20
2550	GSE5418	GPL96	malaria	71
2552	GSE18323	GPL570	malaria	39
2553	GSE18323	GPL571	malaria	215
2557	GSE24849	GPL570	malaria	28
430	GSE15221	GPL6102	malaria_control	28
432	GSE16463	GPL6102	malaria_control	24
434	GSE33811	GPL6244	malaria_control	10
435	GSE34404	GPL10558	malaria_control	155
437	GSE7586	GPL570	malaria_control	20
2551	GSE5418	GPL96	malaria_control	71
2554	GSE18323	GPL570	malaria_control	39
2555	GSE18323	GPL571	malaria_control	215
2556	GSE24849	GPL570	malaria_control	28

Results

Below (Figure 2) is an illustration of three gene networks and the over lap between them. Red represents upregulated, green is downregulated, and the pink lines represent overlap among the networks.



The STAR platform has been applied to many disease states. Multiple users contribute to making annotations. Figure 2 below shows number of samples per tag.



Conclusions

- Our data are proof-of-concept to the potential of STAR.
- Our data set for malaria infection is relatively small, more series and samples need to be analyzed to strengthen the expression signature.
- We anticipate further studies on not only human gene expression for malaria infection, but also using parasite gene expression at corresponding stages of the parasite's life-cycle.