

Progress since last Committee Meeting - September 8, 2016**Awards**

- ASHG Reviewer's Choice Abstract Award 2016 & 2017
- Awarded & Renewed F31 Ruth L. Kirschstein NRSA 9/2016-9/2018
- FASEB MARC Travel Award to ASHG 2014 & 2015
- Genetics and Regulation Training Grant 2013-2016

Publications

- Submitted to PeerJ:
- Revising: Mozaffari, S. V., DeCara, J. M., Shah, S. J., Lang, R. M., Nicolae, D. L., and Ober, C. (2017). Parent of Origin Effects on Quantitative Phenotypes in a Founder Population. *bioRxiv*
- Under revision Scientific Reports:
- Gamazon, E. R., Wheeler, H. E., Shah, K. P., Mozaffari, S. V., Aquino-Michaels, K., Carroll, R. J., Eyler, A. E., Denny, J. C., Nicolae, D. L., Cox, N. J., and Im, H. K. (2015). A gene-based association method for mapping traits using reference transcriptome data. *Nature Genetics*, 47(9):1091–1098

Oral Presentations

- GGSB Work in Progress April 2016 & March 2017
Parent of Origin Effects in the Hutterites
- Human Genetics Work in Progress January 20, 2016
Parent of Origin Effects
- Molecular Biosciences Retreat November 5, 2015
Mozaffari SV, DeCara J, Shah S, Herman C, Lang R, Nicolae D, Ober C., *Parent of Origin GWAS with Cardiovascular Disease Associated Traits in the Hutterites*. 2015: Nov 5; Galena, IL.
- ASHG October 9, 2015
Mozaffari SV, DeCara J, Shah S, Herman C, Lang R, Nicolae D, Ober C., *Parent of Origin GWAS of CVD-Associated Phenotypes in the Hutterites* (Abstract Program #310). Presented at the Annual Meeting of The American Society of Human Genetics; 2015: Oct 9; Baltimore, MD.

Posters

- Mozaffari SV, Nicolae D, Ober C. *Opposite Allele Parent of Origin Effects on Cardiovascular and Asthma Associated Traits in the Hutterites*. Poster presented at the Gordon Research Seminar and Conference; 2017: July 8-14; Stowe, VT
- Reviewer's Choice Abstract Award:
Mozaffari SV, Nicolae D, Ober C. *Opposite Allele Parent of Origin Effects on Body Mass Index in the Hutterites*. Poster presented at the Annual Meeting of The American Society of Human Genetics Conference; 2016: Oct 18-22; Vancouver, Canada
- Mozaffari SV, Gamazon E, Aquino-Michaels K, Cox NJ, Im HK. *Quantifying Context Specificity of Gene Regulation using Predicted Gene Expression Levels*. Poster presented at the Annual Meeting of The American Society of Human Genetics Conference; 2014: Oct 18-22; San Diego, CA

Teaching Assistantship Requirements Completed

- HGEN 47000 *Human Genetics* Fall 2015 & 2017
- Graduate Course: Classic and modern approaches to studying cytogenetic, Mendelian, and complex human diseases. Grant proposal writing course.
- Fall 2017: Conducting two-day computational workshop on GWAS
- MGCB 31400 (BIOS 21236) *Genetic Analysis of Model Organisms* Fall 2014
- Graduate & Undergraduate Course: Introduction to genetic tools, experiments, and model organisms

Additional Courses

- STAT 24500 *Statistical Theory & Methods II* Winter 2015
- HGEN 46900 *Human Variation & Disease* Spring 2015
- STAT 35500 *Statistical Genetics* Spring 2015
- myChoice Mini-Course: *Effective Writing in the Biological Sciences* Fall 2015
- HGEN 48600 *Fundamentals of Computational Biology: Models & Inference* Winter 2016
- PBHS 31831 *Genetic & Molecular Epidemiology* Spring 2016

Additional Workshops & Conferences

- ComSciCon Chicago (Communicating Science Conference & Workshop) August 2017
- Gordon Research Seminar & Conference: Human Genetics & Genomics July 2017
Discussion Leader for mentorship session at GRS
- Summer Institute in Statistical Genetics at the University of Washington July 2016
- Master R Developer Workshop taught by Hadley Wickham May 2015

Extracurricular

- myCHOICE Data Science Trek to San Francisco Bay Area November 2017
- Museum of Science & Industry: Science Connections volunteer Fall 2014-current
Introduce genetic concepts to guests in a fun and engaging way, incorporating hands-on activities. Assist in the Fabrication Lab helping guests design and print custom objects using 3D printers and laser cutters.
- UChicago Software Carpentry Helper Fall 2015, 2016, & 2017
- Expanding Your Horizons (EYH) Chicago: volunteer March 2017
Engage, inspire, and empower young girls to pursue STEM careers at a one-day STEM symposium where 300 middle school girls participate in hands-on science, technology, engineering and math led by academic and professional women.
- myCHOICE Internship: Institute of Translational Medicine: writer Summer 2015
Translate complex research into dynamic science stories. Share translational research stories in weekly newsletter, ITM website, and social media platforms

Proposal Updates:

AIM 1a To identify and characterize parent of origin effects on quantitative traits in the Hutterites.
AIM 1b To estimate maternal and paternal heritability measures on quantitative traits in the Hutterites.

AIM 2a To identify and characterize parent of origin effects on gene expression in 430 Hutterites.
AIM 2b To identify and characterize parent of origin and allele specific effects on gene expression in 430 Hutterites.

Since my last Committee Meeting

AIM 1a

Completed with preprint on bioRxiv⁷, working on revising and resubmitting. We are trying to replicate the significant opposite effects of SNPs in birth weight in the Hutterites.

AIM 1b

Currently working on, adapting methods published in Mott *et al.*⁶. Last committee meeting, it was established this would be the last aim I work on .

AIM 2a

I have remapped the LCL RNA-seq data (and whole blood data for replication) using STAR² and corrected sample swaps using verifyBamID.⁵ I used WASP⁸ to remove mapping bias and remove duplicate reads. I then mapped reads to the maternal (Y_m) and paternal (Y_p) haplotypes. I used STAR to measure gene count from reads. I removed lowly expressed genes but did not normalize gene expression or remove covariates since comparing maternal and paternal expression is done in the same sample under the same conditions.

For this aim I am using simple binomial tests to detect patterns of parent of origin effects (i.e. imprinting) using maternal and paternal gene expression but not any SNPs (no POeQTL). (Removed duplicate reads so no need to model overdispersion). First, I can test for directional asymmetry to test if any genes have maternal or paternal effects by generating a binomial Z-score for each subject for each gene. I only use individuals where $Y_p + Y_m = 5$ and $Z_i \neq 0$ to get a statistic of how skewed the gene expression is for each gene using a binomial test.

$$Z_i = \frac{(Y_p^i - Y_m^i)}{\sqrt{(Y_p^i + Y_m^i)}} \quad (1)$$

$$S \sim \text{Bin}(\#subjects, \frac{1}{2}) \quad (2)$$

Second, I test for asymmetry in maternal and paternal gene expression by genes and weighing by library size where $w_i \propto \frac{1}{librarysize_i}$ and $w_i = \frac{10millionreads}{\#ofreads_i}$

$$Z_g = \frac{\sum_{i=1}^n w_i (Y_p^i - Y_m^i)}{\sqrt{(\sum_{i=1}^n w_i^2 (Y_p^i + Y_m^i))}} \quad (3)$$

AIM 2b

I will test for POeQTLs in this aim testing maternally inherited SNPs with the maternal gene expression and paternally inherited SNPs with paternal expression. I will combine this with POeQTL results from before using the sum of gene expression, especially for genes which we don't have maternal or paternal expression.

$$Y_m \sim X_{SNP} + PO + PO \cdot X_{SNP} \quad (4)$$

$$Y_p \sim X_{SNP} + PO + PO \cdot X_{SNP} \quad (5)$$

Gene	Imprinted	Gene pvalue	weighted Zscore	weighted pvalue
<i>ZDBF2</i>	yes	$7e^{-37}$	45.57	0
<i>NHP2L1</i>	no, Docherty et al. ³	$7e^{-31}$	18.98	$2e^{-80}$
<i>L3MBTL1</i>	yes	$2e^{-29}$	35.73	$1e^{-279}$
<i>PEG10</i>	yes	$2e^{-29}$	51.34	0
<i>SNHG14</i>	no, Baran et. al ¹	$8e^{-27}$	41.09	0
<i>ZNF331</i>	no, Baran et. al ¹	$3e^{-23}$	30.28	$2e^{-201}$
<i>KCNQ1</i>	yes	$4e^{-21}$	-28.94	$4e^{-184}$
<i>LPAR6</i>	no, Baran et. al ¹	$7e^{-21}$	33.50	$4e^{-246}$
<i>FAM50B</i>	yes	$4e^{-20}$	28.89	$2e^{-183}$
<i>PXDC1</i>	no, neighboring gene of FAM50B	$3e^{-14}$	16.05	$6e^{-58}$
<i>PWAR6</i>	no, Prader Willi/Angelman Region 6	$1e^{-09}$	25.59	$2e^{-144}$
<i>NAP1L5</i>	yes	$3e^{-07}$	21.05	$3e^{-98}$
<i>ATP6V0D1</i>	no	$1e^{-06}$	-5.40	$7e^{-8}$

Table 1: Significant genes from sign test (p-values in column 3) with weighted Z score with corresponding p-values in column 4 and 5, respectively. Imprinted as defined by geneimprint.com database.

References

- [1] Baran, Y., Subramaniam, M., Biton, A., Tukiainen, T., Tsang, E. K., Rivas, M. A., Pirinen, M., Gutierrez-Arcelus, M., Smith, K. S., Kukurba, K. R., Zhang, R., Eng, C., Torgerson, D. G., Urbanek, C., GTEx Consortium, Li, J. B., Rodriguez-Santana, J. R., Burchard, E. G., Seibold, M. A., MacArthur, D. G., Montgomery, S. B., Zaitlen, N. A., and Lappalainen, T. (2015). The landscape of genomic imprinting across diverse adult human tissues. *Genome Research*, 25(7):927–936.
- [2] Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T. R. (2012). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29(1):15–21.
- [3] Docherty, L. E., Rezwan, F. I., Poole, R. L., Jagoe, H., Lake, H., Lockett, G. A., Arshad, H., Wilson, D. I., Holloway, J. W., Temple, I. K., and Mackay, D. J. G. (2014). Genome-wide DNA methylation analysis of patients with imprinting disorders identifies differentially methylated regions associated with novel candidate imprinted genes. *Journal of medical genetics*, 51(4):229–238.
- [4] Gamazon, E. R., Wheeler, H. E., Shah, K. P., Mozaffari, S. V., Aquino-Michaels, K., Carroll, R. J., Eyler, A. E., Denny, J. C., Nicolae, D. L., Cox, N. J., and Im, H. K. (2015). A gene-based association method for mapping traits using reference transcriptome data. *Nature Genetics*, 47(9):1091–1098.
- [5] Jun, G., Flickinger, M., Hetrick, K. N., Romm, J. M., Doheny, K. F., Abecasis, G. R., Boehnke, M., and Kang, H. M. (2012). Detecting and Estimating Contamination of Human DNA Samples in Sequencing and Array-Based Genotype Data. *The American Journal of Human Genetics*, 91(5):839–848.
- [6] Mott, R., Yuan, W., Kaisaki, P., Gan, X., Cleak, J., Edwards, A., Baud, A., and Flint, J. (2014). The Architecture of Parent-of-Origin Effects in Mice. *Cell*, 156(1-2):332–342.
- [7] Mozaffari, S. V., DeCara, J. M., Shah, S. J., Lang, R. M., Nicolae, D. L., and Ober, C. (2017). Parent of Origin Effects on Quantitative Phenotypes in a Founder Population. *bioRxiv*.
- [8] van de Geijn, B., McVicker, G., Gilad, Y., and Pritchard, J. K. (2015). WASP: allele-specific software for robust molecular quantitative trait locus discovery. *Nature Methods*, 12(11):1061–1063.