



HAL
open science

Analysis of allele specific expression in mouse liver by RNA-Seq

Pierre-François Roux, Sandrine Lagarrigue, L. Martin, F. Hormozdiari, A. van Nas, Olivier Demeure, A. Ghazalpour, E. Eskin, A.J. Lusis

► **To cite this version:**

Pierre-François Roux, Sandrine Lagarrigue, L. Martin, F. Hormozdiari, A. van Nas, et al.. Analysis of allele specific expression in mouse liver by RNA-Seq: Marked differences compared to cis-eQTL identified using genetic linkage. 63. Annual Meeting of the American Society of Human Genetics (ASHG), Oct 2013, Boston, United States. , 2013, Poster book. hal-01210510

HAL Id: hal-01210510

<https://hal.science/hal-01210510>

Submitted on 3 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

American Society of Human Genetics 63rd Annual Meeting October 22–26, 2013 Boston

POSTER ABSTRACTS

The program and abstract/poster board number next to each listing is followed by a **W** (Wednesday), or **T** (Thursday), or **F** (Friday) to indicate the day on which authors must be present at their poster boards. Posters will remain on the boards for all three days (Wednesday through Friday).

Session Topic/Title	Abstract/Poster Board Numbers		Session Topic/Title	Abstract/Poster Board Numbers	
	Start #	End #		Start #	End #
Epigenetics	412	511	Genetics/Genomics Education	2407	2424
Genome Structure, Variation and Function	512	694	Ethical, Legal, Social and Policy Issues in Genetics	2425	2476
Pharmacogenetics	695	743	Genetic Counseling	2477	2500
Complex Traits and Polygenic Disorders	744	1180	Health Services Research	2501	2521
Psychiatric Genetics, Neurogenetics and Neurodegeneration	1181	1430	Clinical Genetic Testing	2522	2639
Bioinformatics and Genomic Technology	1431	1693	Clinical Genetics and Dysmorphology	2640	2816
Statistical Genetics and Genetic Epidemiology	1694	1942	Prenatal, Perinatal and Reproductive Genetics	2817	2884
Evolutionary and Population Genetics	1943	2105	Molecular Basis of Mendelian Disorders	2885	3160
Cardiovascular Genetics	2106	2260	Development	3161	3187
Therapy for Genetic Disorders	2261	2312	Cytogenetics	3188	3257
Metabolic Disorders	2313	2406	Cancer Genetics	3258	3507

Posters should remain on the board for all three days (Wednesday through Friday)

POSTER AUTHOR SCHEDULE

The program and abstract/poster board number next to each listing is followed by a **W** (Wednesday), **T** (Thursday), or **F** (Friday) to indicate the day on which authors must be present at their poster boards. Refer to the schedule below for presentation times and for the poster mounting/removal schedule. **Posters should remain on the boards for all three days.**

Wednesday

10:00 am–10:30 am
10:00 am–6:00 pm
10:30 am–12:30 pm

Authors place posters on boards
Posters open for viewing

Poster Session I (W)

10:30 am–11:30 am (*odd poster board numbers; author must be present*)
11:30 am–12:30 pm (*even poster board numbers; author must be present*)

Thursday

10:00 am–4:30 pm
10:30 am–12:30 pm

Posters open for viewing

Poster Session II (T)

10:30 am–11:30 am (*odd poster board numbers; author must be present*)
11:30 am–12:30 pm (*even poster board numbers; author must be present*)

Friday

10:00 am–2:00 pm
10:30 am–12:30 pm

Posters open for viewing

Poster Session III (F)

10:30 am–11:30 am (*odd poster board numbers; author must be present*)
11:30 am–12:30 pm (*even poster board numbers; author must be present*)

2:00 pm–2:30 pm
2:30 pm

Authors must remove posters
Exhibit Hall and Posters closed

657T

Chromatin enhancer maps in early differentiating T helper cells reveal regulatory SNPs associated with autoimmune diseases. *D. Hawkins¹, A. Larjo², S. Tripathi³, U. Wagner⁴, Y. Luu⁴, S. Raghav³, L. Lee⁴, S. Raghav³, R. Lund³, B. Ren⁴, H. Lahdesmaki², R. Lahesmaa³.* 1) Medicine and Genome Sciences, University of Washington, Seattle, WA; 2) Department of Information and Computer Science, Aalto University, Aalto, FI; 3) Turku Centre for Biotechnology, Turku, FI; 4) Ludwig Institute for Cancer Research, Department of Molecular and Cellular Biology, UCSD, San Diego, CA.

Naïve CD4⁺ T cells can differentiate into specific helper and regulatory T cell lineages in order to combat infection and disease. The correct response to cytokines and a controlled balance of these populations is critical for the immune system and the avoidance of autoimmune disorders. To investigate how early cell fate commitment is regulated, we generated the first human genome-wide maps of histone modifications that reveal enhancer elements after 72 hrs of *in vitro* polarization toward T helper-1 (Th1) and T helper-2 (Th2) cell lineages. Our analysis indicated that at this very early time point, cell-specific gene regulation and enhancers are at work directing lineage commitment, and likely opposed a role in cell maintenance as in fully differentiated cells. We determined a unique set of enhancer elements relative to recent ENCODE data in fully differentiated cells. Nucleosome-free regions were determined within enhancer chromatin structures. Examination of enhancers at nucleosome resolution identified transcription factor binding site (TFBS) motifs for expressed TFs with known and unknown T cell roles as putative drivers of lineage-specific gene expression.

An integrative analysis of immunopathogenic associated single nucleotide polymorphisms (SNPs) suggests a role for distal regulatory elements in disease etiology. We found SNPs overlapping enhancers associated with asthma, Crohn's disease, multiple sclerosis (MS), psoriasis, rheumatoid arthritis (RA), type 1 diabetes (T1D) and ulcerative colitis. Guided by the motif analysis, we used DNA Affinity Precipitation Assays (DAPA) to determine if autoimmune disease-associated SNPs overlapping TFBS motifs could disrupt TF binding. We successfully found that individual SNPs associated with rheumatoid arthritis, type 1 diabetes and ulcerative colitis are enough to disrupt TF binding at enhancer sequences.

Lastly, we are using computational approaches to predict target genes of enhancers that overlap autoimmune disease-associated SNPs. We are also mapping histone modifications to find novel or alternative promoters that may also overlap associated SNPs. Collectively, our approach sheds new light on the functionality of non-genic disease-associated SNPs and begins to provide novel insight on the etiology autoimmune diseases.

658F

Analysis of allele specific expression in mouse liver by RNA-Seq: marked differences compared to cis-eQTL identified using genetic linkage. *P.-F. Roux^{1, 2, 3}, S. Lagarrigue^{1, 2, 3}, L. Martin⁴, F. Hormozdiari^{5, 6}, A. van Nas⁶, O. Demeure^{1, 2, 3}, A. Ghazalpour⁶, E. Eskin^{5, 6}, A.J. Lusis^{4, 6, 7}.* 1) INRA, UMR1348 Pegase, Rennes, France; 2) Agrocampus Ouest, UMR1348 Pegase, Rennes, France; 3) Université Européenne de Bretagne, France; 4) Department of Medicine/Division of Cardiology, University of California, Los Angeles, United States of America; 5) Department of Computer Sciences, University of California, Los Angeles, United States of America; 6) Department of Human Genetics, University of California, Los Angeles, United States of America; 7) Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, United States of America.

We report an analysis of allele specific expression [ASE] and parent-of-origin expression in adult mouse liver using next generation sequencing (RNA-Seq) of reciprocal crosses of heterozygous F1 mice from the parental strains C57BL/6J and DBA/2J. The genes exhibiting ASE differed markedly from the putative cis-acting expression quantitative trait loci (cis-eQTL) identified in an intercross between the same strains. While about 60% of the ASE, mapped by RNA-Seq, were found in the eQTL gene set, only a small fraction of the eQTL, mapped by linkage analysis, were found in the ASE gene set. We discuss the various biological and technical factors that contribute to these differences, in particular strengths of the two approaches in making a distinction between local and cis eQTL.

We also identify genes exhibiting parental imprinting and complex expression patterns. Our study demonstrates the importance of biological replicates, which is not currently the norm, to limit the number of false positives with such RNA-Seq data.

659W

A large-scale transcriptome study in the Sardinian population. *M. Pala^{1,2,6}, M. Marongiu², Z. Zappala¹, A. Mulas², R. Cusano², F. Crobu², F. Reinier³, R. Berutti^{3,6}, M.G. Piras², C. Jones³, D. Schlessinger⁵, G. Abecasis⁴, A. Angius², S. Sanna², F. Cucca^{2,6,7}, S. Montgomery^{1,7}.* 1) Pathology and Genetics Dept, Stanford University, Stanford, CA; 2) Istituto di Ricerca Genetica e Biomedica (IRGB), CNR, Monserrato, 09042, Italy; 3) CRS4, Advanced Genomic Computing Technology, Pula, Italy; 4) Center for Statistical Genetics, University of Michigan, Ann Arbor, MI; 5) Laboratory of Genetics, National Institute on Aging, Baltimore, Maryland, USA; 6) Dipartimento di Scienze Biomediche, Università di Sassari, Sassari, Italy; 7) co-senior authors.

The interpretation of genome-wide association studies (GWAS) is very challenging since most candidate loci fall in non-coding regions and are difficult to interpret. Given the important role of regulatory variation in phenotypic complexity, expression quantitative trait loci (eQTLs) have been proposed as an informative intermediate phenotype between genetic variation and human disease. To study the role of transcriptional mechanisms on regulating eQTL targets, we sequenced the polyA RNA fraction of peripheral blood mononuclear cells (PBMCs) isolated from 624 related individuals (259 families). These individuals had been subject to whole genome sequencing and characterized for more than 800 quantitative traits (including > 250 immune cell traits) by the SardiNIA project. To assess the impact of eQTLs on lincRNAs, we also sequenced the RNA of a subset of 80 individuals after ribosomal RNA depletion. Our preliminary results show that the expression levels of 10,389 genes are associated with a genetic variant (top SNP with p-value $\leq 10^{-8}$ uncorrected), 40 of which are present in the GWAS catalog. We also assessed splicing events and identified 27,904 donor splice sites that are associated with a genetic variant (top SNP with p-value $\leq 10^{-8}$ uncorrected), of which only 4 are present in the GWAS catalog. Using a subset of 68 individuals, we identified 9,427 heterozygous sites which show allele-specific expression (p-value ≤ 0.05) 39 of which show high allelic imbalance (a mean of 0.80 for the most expressed allele) and low variability between individuals. These sites are enriched for missense variants and located near genes implicated in disease. In addition, we developed a statistical method to identify genes that are specifically and significantly over or under expressed within single families compared to the majority of the population. We identified 39 genes that have family specific expression (p-value ≤ 0.01 ; FDR 0.462). Interestingly, these genes are also enriched for rare variants within 70kb upstream of their transcriptional start site, and we hypothesize that they may have a role in rare diseases. We are now extending these analyses to the entire cohort and will present these analyses. Our goal is to identify associations between genotype and gene expression in order to enhance our overall understanding of gene regulation and correlate these molecular events with clinical data in order to characterize their role in phenotypic complexity and disease.