Ocular disease mechanisms elucidated by genetics of human fetal retinal pigment epithelium gene expression

Project Proposal Theme07 - Gene Expression Analysis

Lisa Hu
414264
Bio-Informatica
Hanzehogeschool Groningen, ILST
Marcel Kempenaar
8 March 2022

Short description

The retinal pigment epithelium (RPE) is the pigmented layer of cells in the retina of the eye that serves vital roles in the visual mechanisms. Genetic variants that affect the RPE gene expression could lead to various ocular diseases. By examining genetics of gene expression of cultured human fetal RPE (fRPE) cells under two different metabolic conditions, these genetic variants could be elucidated. Hundreds of expression or splice quantitative loci (e/sQTLs) were discovered, but a particular variant in **RDH5** contributes to both increased risk of age-related macular degeneration (AMD) and decreased myopia risk.

Methods

Genome_build: hg19

1. Samples of human fetal RPE cells were grown and sequenced, resulting in RNA sequences.
2. Raw data de-multiplexed with *bcl2fastq2* from **Illumina** (default parameters).
3. Aligned with **STAR** (against hg19 human reference genome with GENCODE v19 annotations + default parameters).
4. Duplicates marked with **Picard MarkDuplicates**. Duplicates and non-perfect mapping qualities deleted.
5. **HTSeq** for counting reads overlapping gene based on GENCODE v19 annotation (counted on reverse strand).
6. Quantifying RPKM with **RNA-SeQC** (with hg19 human reference genome + GENCODE v19 annotations (flags `-noDoC -strictMode`), else default parameter).
7. Quantified allele-specific expression with **RASQUAL** (with createASVCF.sh + default parameters).
8. **LeafCutter** for splicing quantification:
    1. Converted BAM files to splice junction counts (*bam2junc.sh*)
    2. Cluster introns based on sharing of splice donor or acceptor sites (*leafcutter_cluster.py*)

Results

A total of 837 protein coding and lncRNA genes showed evidence of significant differential (upregulated) expression. Three of the top ten genes are involved in lipid metabolism. using GSEA, the top two upregulated pathways in galactose medium are cholesterol homeostasis and mTORC1 signaling. By forcing the cells to rely on oxidation of glutamine for ATP generation increases expressions of genes that promote the lipid synthesis.

Sample group

24 fetal RPE cell lines were processed under glucose and galactose metabolic conditions, giving 48 samples in total.

Project plan

I will be re-doing part of the analysis: Transcriptomic differences across two metabolic conditions.

Stated packages and softwares

- DESeq2
- Git repository: https://github.com/boxiangliu/rpe

Example data

```
##                   V1   V2
## 1 ENSG00000000003.10 3596
## 2  ENSG00000000005.5    0
## 3  ENSG00000000419.8  668
## 4  ENSG00000000457.9  567
## 5 ENSG00000000460.12  189
```