Eye in a Disk: eyeIntegration human pan-eye and body transcriptome database version 1.0

Vinay Swamy[1] and David McGaughey[1,✉]

March 15, 2019

PURPOSE: To develop an easily accessible and reliable RNA-seq transcriptome database of healthy human eye tissues and a matching reactive web application to query gene expression in eye and body tissues. METHODS: We downloaded the raw sequnce data for 892 RNA-seq samples across 44 tissues in the GTEx project as a non-eye reference set. We then queried several public repositories to find all healthy, non-perturbed, human eye-related tissue RNA-seq samples. The 1311 total samples we found were sent into a Snakemake-based reproducible pipeline we wrote to quantify all known transcripts and genes, removes samples with poor sequence quality and mislabels, normalizes expression values across each tissue, performs 450 differential expression tests, calculates GO term enrichment, and outputs all as a single SQLite database file: the Eye in a Disk (EiaD) dataset. Furthermore, we rewrote the web application eyeIntegration (https://eyeIntegration.nei.nih.gov) to display EiaD. RESULTS: The new eyeIntegration portal provides quick visualization of human eye-related transcriptomes published to date, by database version, gene/transcript, and 58 tissues. As a test of the value of this unified pan-eye dataset, we showed that fetal and organoid retina are highly similar at a pan-transcriptome level but display distinct differences in certain pathways and gene families like protocadherin and HOXB family members. CONCLUSION: The eyeIntegration v1.0 web app serves the pan-human eye and body transcriptome dataset, EiaD. This offers the eye community a powerful and quick mean to testing hypotheses on human gene and transcript expression across 44 body and 14 eye tissues.

[1] Ophthalmic Genetics and Visual Function Branch, National Eye Institute, National Institutes of Health

[✉] Correspondence: David McGaughey <mcgaugheyd@mail.nih.gov>

Introduction

RNA-seq is the predominant technology for deciphering transcriptomes

From anterior to posterior along the light trajectory, the human eye is composed of the cornea, lens, retina, retinal pigment epithelium (RPE), and choroid. The differentiation, maturation, and function of these tissues is mediated through spatial and temporal specific transcript and gene expression patterns, also known as the transcriptome. Today, RNA-sequencing (RNA-seq) is the predominant technology for quantifying the transcriptome. Analysis of the transcripts' expression across tissue, time, and perturbation allows researchers to decipher the genetic controls of eye development and function. To this end, a wide variety of human tissue sources have been used to assess gene function, including primary tissue (fetal and post-mortem), differentiated stem cells,

immortalized cell lines, and most recently, organoids. These tissue types have been deeply sequenced across the cornea[1–7], lens[8], retina[9–17], and RPE (choroid)[14,17–34].

The GTEx gene expression resource lacks eye-specific tissues

The Genotype-Tissue Expression (GTEx) Project has generated RNA-seq data across dozens of post-mortem human tissues from hundreds of unique donors and presents the gene and transcript level data in a comprehensive and user-friendly web app (https://gtexportal.org/); however eye tissues have not been included[35,36]. Recently Ratnapriya et al. have published a huge set of post-mortem retina, both normal and with varying degrees of age-related macular degeneration and the GTEx project is providing the data as a download link. This data, as of March 2019, is not available in the interactive GTEx visualizations[37]. The Sequence Read Archive (SRA) and European Nucleotide Archive (ENA) are the primary repositories for all raw sequence data and two groups have quantified large portions of the RNA-seq data, including some human eye tissues, from the SRA: recount2 and ARCHS4[38,39]. To date, no curation of the sample level metadata has been done, therefore it is challenging to parse out which eye tissues are present and even more difficult to determine whether any samples were chemically or genetically perturbed. More targeted web resources that allow researchers to quickly assess gene expression in eye tissues include iSYTE, EXPRESS, and retina.Tigem.it[16,40,41]. However iSYTE only includes lens samples, EXPRESS is limited to a subset of mouse lens and retina samples, and retina.Tigem.it is retina only. We thus aimed our efforts at developing an easily accessible and reliable RNA-seq based transcriptome database of healthy human eye tissues and a matching reactive web application to query gene expression in eye and body tissues.

The eyeIntegration app interactivey serves huge GTEx and human eye tissue datasets (EiaD)

The eyeIntegration web resource (https://eyeIntegration.nei.nih), originally released in 2017 at version 0.6, provides the largest set of transcriptomes from hand-curated human eye tissues along with hundreds of GTEx tissue samples[42]. This interactive web app allows for quick transcript and gene comparisons across many eye tissues and dozens of other body tissues. The dataset that the original eyeIntegration web app served was created with a series of scripts, several of which were run interactively to manually assess quality control for the samples. The interactive nature of some of the steps precluded efficient and regular data updates for the data.

To better meet the needs of the eye research community we have re-written the bioinformatic pipeline that creates the eye and body RNA-seq dataset to allow for regular, versioned updates for eyeIntegration. We call this reproducible and versioned transcriptome dataset "Eye in a Disk" (EiaD). The pipeline automates the EiaD creation, ensures full reproducibility of the results, allow for external data comparison, provides consistent sample quality control, and improves efficiency for future sample updates. The 2019 EiaD dataset contains several new tissue types, full gene product quantification, along with over 100 new samples and improved sample labeling. The eyeIntegration web app has also been re-written to provide many new features, including versioned EiaD datasets, custom URL shortcut creation, new visualizations, improved data tables searching, easy download of core datasets, and local install of the entire app with three commands. Additionally, we are prototyping new tools to display single cell RNA-seq (scRNA-seq) data to provide researchers access to cell type specific information about gene expression across murine retinal development.

The EiaD dataset can be used to identify potential avenues to improve retina organoid maturation

Retina organoids are an increasingly popular means to model human retina development. We used our pan-study EiaD dataset to show that, at a pan-transcriptome level, organoids are highly similar to early fetal retina tissue. We also show that important temporal gene expression patterns in the fetal retina tissue are recapitulated in the organoids. As the organoid differentiation methods do not yet produce fully mature retina, we focused on identifying differentially expressed processes between organoid retina and embryonic retina and detected, for example, identifying protocadherin and HOXB family gene expression differences which suggest targetable pathways to improve and benchmark organoid differentiation methods.

Methods

Identification of potential eye samples

We exhaustively searched the SRA with the SRAdb R package for eye related tissues using the query 'cornea|retina|RPE|macula|fovea|choroid|sclera|iris|lens|eye' across all columns and rows in the 'SRA' table[43,44]. We hand selected relevant studies and selected healthy, control or unmodified samples spanning primary adult tissue, primary fetal tissue, induced pluripotent stem cell (iPSC)-derived tissue, stem cell derived organoids, and

immortalized cell lines. In order to compare gene expression in the eye against expression in other body tissues, we obtained samples from 44 different body tissues that contained at least 10 male and 10 female samples from the GTEx project. Using SRA metadata from each study we extracted sample and run accessions, library type, tissue of origin, and sub-tissue of origin. Any of the preceding information missing from the SRA metadata was added by hand, when available. Stem cell-derived tissues and cell lines are marked as sub-tissues of the tissue they model.

Raw data download and quantification

We downloaded the relevant SRA files for each sample directly from the NCBI ftp server using the file transfer software Aspera. SRA files were converted to FASTQ format using the tool fastq-dump from the SRAtoolkit software package[43]. Samples only available in the BAM format were converted to FASTQ format using SAMTools[45]. Sample transcriptomes were quantified using the alignment free quantification software Salmon, using transcriptomic index built from gencode v28 protein coding transcript sequences using the transcriptomic aligner Salmon[46,47]. Using the resulting expression quantification, we identified lowly or unused transcripts within the gencode annotation, and removed transcripts that accounted for 5% or less of the total expression for its parent gene as per Sonneson et al[48]. Samples were re-quantified against a transcriptomic index built on the filtered transcript sequences. The Salmon count values were quantified as (transcript) length scaled Transcripts Per Million (TPM) to the transcript and gene level using tximport[49].

Quality control

We first removed samples with a Salmon calculated mapping rate less than 40%. We further filtered samples by calculating the median gene expression for each sample and only kept samples that had a median count >2 across all genes. We removed lowly expressed genes by calculating the median expression across all samples for each gene and kept genes that had a median count >200 across all samples. To reduce the noise from experimental variability between each study, we normalized samples by sequence library size using the calcNormFactors function from the edgeR r-package, and then quantile smoothed expression data using the R package qsmooth at the tissue level[50,51].

We generated sample level two-dimensional groupings by with the t-Stochastic Neighbor Embedding (t-SNE) algorithm[52]. We then calculated the center of the resulting sub-tissue clusters by finding the average two-dimensional coordinates of all samples of a given sub-tissue type (p1). We used these sub-tissues groupings to

identify outlier samples, or samples that failed to cluster with other samples of the same sub-tissue type, by measuring the euclidean distance of each sample (p2) from its sub-tissue cluster center and removed samples that had distance 4 standard deviations from the average distance with the equation $dist = \sqrt{\sum}(p1 - p2)^2$.

Differential Gene Expression Analysis and GO term enrichment

We used the processed data to determine differential gene expression between different sub-tissue types. First, we generated a synthetic body set to serve as single representative sub-tissue type for pan-body gene expression by randomly sampling GTEx tissues. We used the voom function from the limma R package to convert gene expression to precision weights, and then performed pairwise differential expression tests for all combinations of eye sub-tissues, the synthetic body tissue, and human body tissues using an empirical Bayes test[53,54]. We extracted significant genes (FDR $p < 0.01$) for all 450 comparisons and used these to calculate GO enrichment. The significant gene list for each eye sub-tissue was split into upregulated and down regulated sets and each set f tested for enrichment using the enrichGO function from the clusterProfiler R package (q-value < 0.01)[55].

eyeIntegration web app and R package

The data generated in the above steps is consolidated into a SQLite database, with the original dataset for eyeIntegration and the new 2019 EiaD dataset each getting a separate database file. The code that creates the eyeIntegration web app is written in Shiny and R and has been wrapped into an R package (https://github.com/davemcg/eyeIntegration_app/) that can be deployed on a local computer or a web server (https://eyeIntegration.nei.nih.gov). The app can be deployed on a local computer with 35GB of free disk space by running three commands in R: “devtools::install_github(‘davemcg/eyeIntegration_app’)”, “eyeIntegrationApp::get_eyeIntegration_datasets()”, and “eyeIntegrationApp::run_eyeIntegration()”.

Snakemake reproducible pipeline

While the sample search and metadata parsing in a semi-curated process, the processing from the raw data to the creation of the SQLite EiaD database underlying eyeIntegration is wrapped in a Snakemake pipeline, which ensures full reproducibility of the results[56]. We make the code for the pipeline available at https://github.com/davemcg/EiaD_build.

scRNA-seq processing

The eyeIntegration site, as of January 2019, hosts two large scRNA-seq datasets from Macosko et al. and Clark et al[57,58]. We use the processed gene count data from each group, as well as their cluster assignments which specify what cell type each individual cell is. The count data is mean averaged to the cell type, age, and gene level for the single cell expression section of eyeIntegration. We also display t-SNE and UMAP-based two-dimensional visualizations of the Macosko and Clark data, respectively, in the web app. For detail so the t-SNE processing we did on the Macosko dataset, see the methods of Bryan et al[42].

Manuscript as code

This manuscript's figures, tables, and most numbers, are all created and laid out in a R markdown document that interweaves code and text. The knitr and pandoc program is used to lay out the figures and tables and output a docx file. The code that generates this manuscript can be found at https://github.com/davemcg/eyeIntegration_v1_app_manuscript.

Results

EiaD 2019 contains 44 GTEx body sub-tissue types along with 22 new human eye RNA-seq studies, 191 new samples, 14 eye sub-tissue types

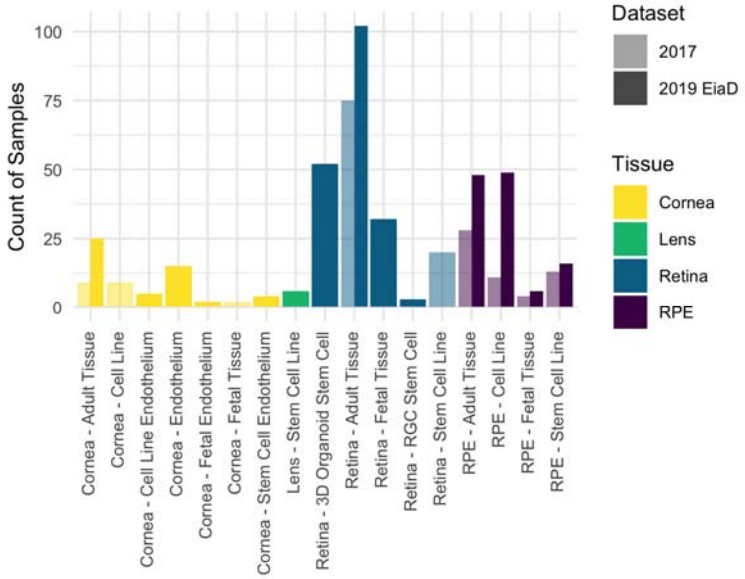


Figure 1: Substantial increase in eye tissue count and type from 2017 (180, lighter color) to 2019 (371, darker color) EiaD. We also improved the metadata labelling, the cornea samples (yellow) now delineates endothelial and

epithelial tissues and the retina samples (teal) distinguish retina organoid and retinal ganglion cell (RGC) from stem cells.

Our query on January 7th 2019 to the SRA found 107 potentially relevant studies. We removed non-pertinent studies and selected healthy or unmodified tissue from each relevant study for a total, including of 44 studies, 28 of which are new to the 2019 EiaD dataset. The 2019 EiaD dataset contains 371 human eye tissue samples and also includes 848 GTEx samples across 44 tissues for easy comparison (Table 1, Supplemental Table 1). The 2019 EiaD contains 6 undifferentiated iPSC, 51 cornea, 6 lens, 189 retina, and 119 RPE (choroid) samples; in total we have added 191 new samples to the 2019 EiaD (Figure 1). We refer to native-tissue extracted RPE as RPE (choroid) because it is not possible to remove the choroid from the RPE without culturing.

Stem cell-derived cornea, stem cell-derived lens, and fetal retina are three new types of sub-tissues that are now available in EiaD. We have also substantially improved the granularity of the cornea tissue metadata, now delineating whether the tissue is from the endothelium or epithelium (Figure 1); previously these had been grouped together as adult tissue. Another substantial addition to the 2019 EiaD are non-protein coding genes; while protein-coding is the most common gene and transcript typse, there are dozens of different non-coding classes. The 2017 version of eyeIntegration only quantified protein coding genes and transcripts. We now quantify expression across 42 gene and 46 transcript types, including protein coding, retained intron, lincRNA, antisense, and pseuodogenes (Supplemental Table 2).

Rigorous quality control and reproducible workflow system ensures high quality transcriptomes that consistently cluster together by tissue type

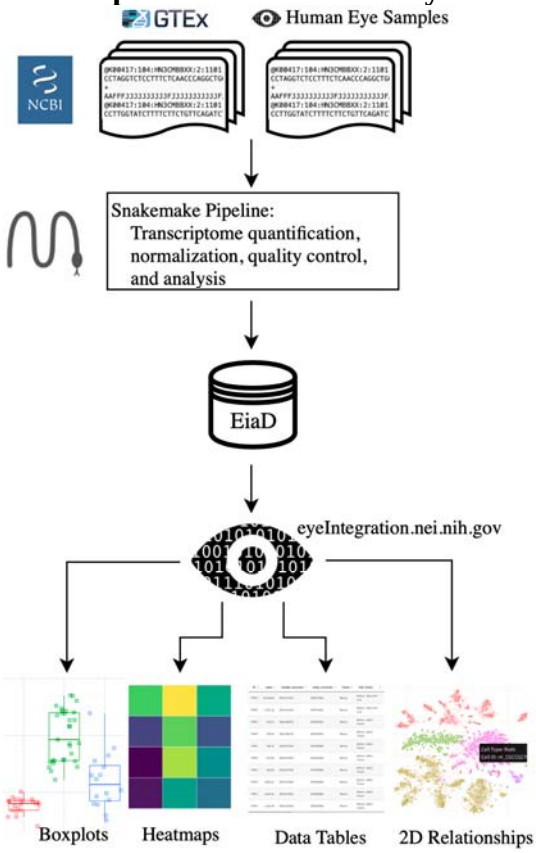


Figure 2: Raw RNA-seq data from the SRA is run through our pipeline to create the EiaD, which is used by eyeIntegration app to serve interactive gene expression visualizations across 63 tissues

We built an automated pipeline for processing and analyzing all data for the web app using the program Snakemake, a python-based workflow management system that allows for efficient parallel execution of the analysis, facilitates reproduction by others, and simplifies long term maintenance of the EiaD data (Figure 2, Supplemental Figure 2)[56]. To create a high quality final dataset across the 1311 initial samples (Supplemental Table 3) and 46,030,412,878 reads we developed a rigorous quality control procedure as part of our analysis, considering a sample's read mapping rate and median count level as well as behavior relative to samples of the same sub-tissue type (see Methods). To identify tube mislabeling or sample extraction issues, we used t-SNE to group samples by sub-tissue type to identify variability within samples of the same sub-tissue and ensure overall consistency in data processing (Figure 3).

After our quality control and processing workflow, we found that samples of the same tissue type and origin cluster well together (Figure 3). For example, in the retina group, primary adult tissue clusters tightly and

distinctly from other cell types, and retinal organoids and fetal retina samples cluster together. Our ability to uniformly cluster data by known biological source independent of study origin demonstrates that our workflow can effectively account for technical variation between studies.

While t-SNE is a powerful algorithm for grouping samples, it is not consistent for determining the relationships between clusters[59]; PCA is more useful in this regard. We ran a PCA dimensionality reduction (Supplemental Figure 3) on all samples, finding that the eye tissues still generally group together and apart from all other human body tissues. Adult retina is most similar to the brain tissue, pituitary is similar to adult RPE, and skin and blood are similar to cornea.

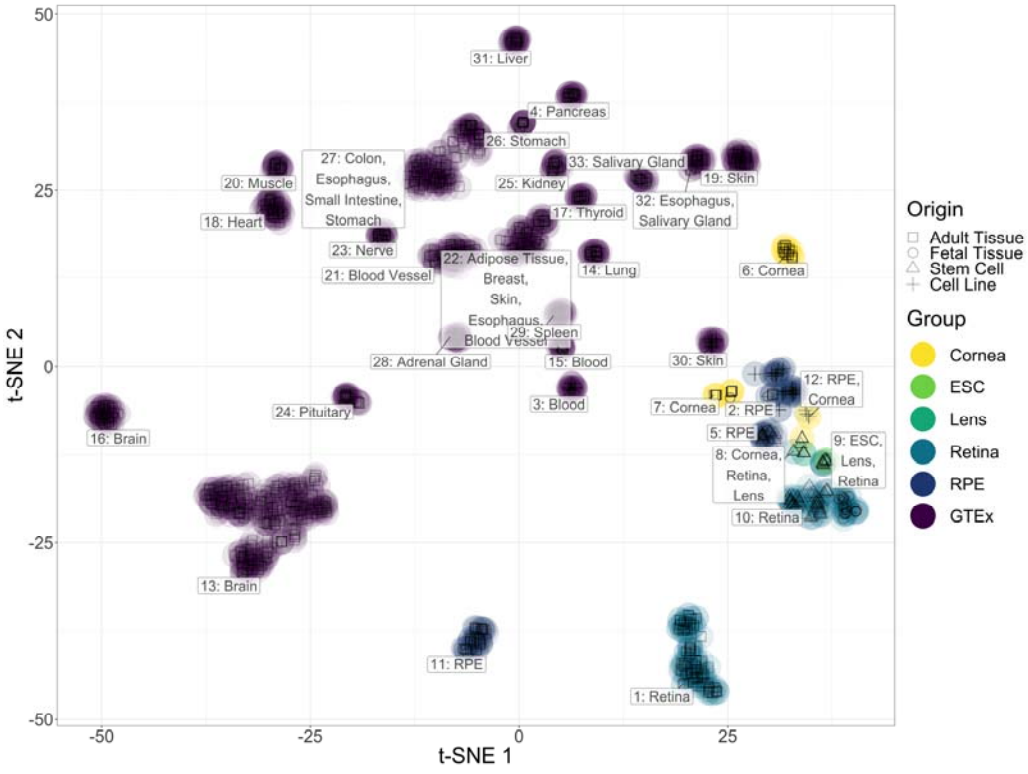


Figure 3: t-SNE two-dimensional transcriptome profiles by sample demonstrate effective quality control and transcriptome processing. Colors match different tissue types and shapes of points define the origin of the tissues.

Tissue	Pre QC Count	Count	Sub-Tissue Types (Count)
Cornea	62	51	Adult Tissue (25), Cell Line Endothelium (5), Endothelium (15), Fetal Endothelium (2), Stem Cell Endothelium (4)
Lens	9	6	Stem Cell Line (6)

Tissue	Pre QC Count	Count	Sub-Tissue Types (Count)
Retina	198	189	3D Organoid Stem Cell (52), Adult Tissue (102), Fetal Tissue (32), RGC Stem Cell (3)
RPE	144	119	Adult Tissue (48), Cell Line (49), Fetal Tissue (6), Stem Cell Line (16)

Table 1: EiaD contain a huge set of diverse eye tissues

The eyeIntegration web app provides interactive visual portal to all data

The EiaD 2019 dataset is used directly by the eyeIntegration web app (https://eyeIntegration.nei.nih.gov). The web app was designed to provide a simple interface that has the same general concept – select specific genes and tissue and view relevant information. The web-app is divided into four general categories: expression, two-dimension sample relationships, gene networks, and data tables.

Custom gene and tissue expression boxplots

The 'Expression' tab of the webpage provides a wealth of information about both gene- and transcript-level expression for eye and body tissues, giving the user the ability to compare the expression of different genes within a single tissue, as well as the expression of genes across multiple tissues (Figure 4A). The user first selects either the 2017 or 2019 gene or transcript EiaD dataset, then genes (or transcripts), then tissues. A boxplot is then generated after hitting the "Re(Draw) Plot" button with overlaid individual data points. On mouse-over, the metadata for the individual sample is displayed. A tabular report is generated based on selected genes and tissues: a table with links to Ensembl, GeneCard, and OMIM for each gene for quick referencing, and a table containing expression levels for each selected gene in each selected tissue. The tables can be arranged or sorted to the user's preference and can be easily downloaded for local use.

Figure 4: Screenshots from eyeIntegration web app. A. Pan-tissue gene expression box plots with accompanying data tables. B. Heatmap visualization. C. CRX mouse retina gene expression heatmap and table information from Clark et al. E11 to P14 scRNA-seq. D. t-SNE visualization of gene expression profiles of individual cells from Macosko et al. E. Data table export view. F. Differential gene expression across different tissues and with GO term enrichment.

Heatmaps based on expression can be drawn for selected genes and tissues and gene expression can be compared across many genes and tissues (Figure 4B). Finally, a session can be saved or shared by building a custom link for the session with the "Build URL Shortcut" button.

Differential expression and gene ontology enrichment tests allow quick comparison of gene differences between groups

We performed multiple differential comparisons at the sub-tissue level within all eye tissues and against a pan-body synthetic set comprised of a stratified sample of all tissues present in our subset of the GTEx dataset, allowing quick identification of eye specific genes across 450 different comparisons. We have expanded the

differential tests in the 2019 EiaD by adding the GTEx tissues as direct comparisons to our eye sub-tissues. The user can view the results selecting 'Differential' under the 'Expression' tab (Figure 4F). As with 'Expression', the user can select which version of the web app to draw data from as well as select for gene- or transcript-level comparisons. The user additionally has the option to select different gene classes to examine, e.g. protein coding, lincRNA.

The results of differential expression are presented in a tabular format showing $\log_2$ fold change, average expression, and p-values. Depending on the comparison, there are 967 to 29422 differentially expressed genes (Supplementary Files). The table can be easily searched for any given gene, viewed and ordered to the user's preference, and downloaded in CSV format. Differential expression can be visualized through fold change bar graphs with the 'Pan-tissue plots' selection under 'Expression'. Additionally, we performed GO enrichment for all differential comparisons. Enriched GO terms are presented first as a word cloud, for quick comparison of GO enrichment. We provide tables, with similar viewing options as the differential expression table, for enriched GO terms in each class of a given differential comparison.

Murine scRNA-seq enables testing of retina cell type specific expression

We incorporated scRNA-seq data from murine retina across two studies[57,58]. This allows researchers to quickly examine gene expression across individual cell types in the retina. Single cell gene expression data is visualized through a heatmap showing the expression of a gene across multiple retinal cell types and different developmental time points, from embryonic day (E)11 to postnatal day (P)14 (when available), and a table of expression values is generated containing the expression data used to draw the heatmap (Figure 4C). We also provide t-SNE/UMAP based clustering using cell type specific labeling created by the publishing authors (Figure 4D, see Methods). The plots show all cell types present at a given developmental stage, and highlights cells expressing a gene above a user-selected given level.

EiaD 2019 suggests that iPSC-derived organoids and fetal retina have closely related transcriptomes

There are, currently, two major approaches to studying developing human retina: post-mortem fetal tissue and stem-cell derived organoids. We looked at whether at how well these approaches to studying developing retina

compare at a transcriptomic level, both for tissue - organoid relationships and how well they correlate across early development.

To evaluate how the tissues and organoids compare at a transcriptome level, we looked at the same t-SNE plot from Figure 3 and focused in on the three types of retina tissue (adult, fetal, and organoid) (Figure 5A). Here we saw three distinct groupings: adult retina (1), developing fetal retina and stem cell-derived organoid (2), and undifferentiated and early differentiating stem cells (3). We identified several organoid samples in cluster 3, but these share one important difference from the rest of the organoid samples in cluster 3: they have been differentiating for less than 30 days (shape 'X'). All of the organoid retina samples in cluster 2 are older than 50 days.

To assess how similarly the fetal and organoid retina develop through time, we plotted expression of retinal progenitors, photoreceptors, and retinal ganglion markers by time in days (Figure 5B). Each row is a gene marker of either retinal progenitor, photoreceptor, or RGC. The rows are hierarchially clustered to put more similar expression patterns closer together, as denoted by the height of the dendrogram. We split the organoid tissues into three groups: Kaewkhaw et al. GFP+ and GFP- samples, and Eldred et al. samples[12,60]. The Kaewkhaw samples are flow sorted for a GFP marker (GFP+) under the control of the *CRX* promoter, an important regulator of photoreceptor development. GFP+ cells wouldbe enriched in photoreceptor populations. We saw that the retinal progenitor, photoreceptor, and RGC groups are largely clustered together, with patterns of expression consistent across the fetal retinal and organoid groups.

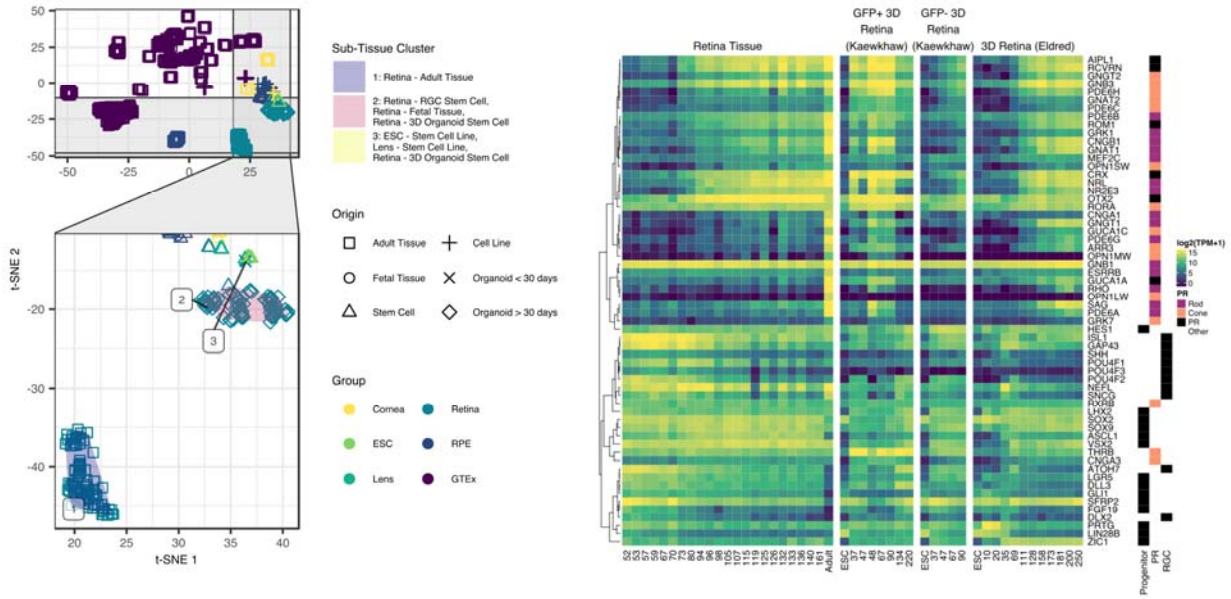


Figure 5: Organoid retina, stem cell retina, and fetal retina tissue have highly similar transcriptomes. The zoom inset (A) shows the retina samples. The "Sub-Tissue Cluster" shading shows the cluster membership of the three major groups. The shapes of the points show the different origin types - notable types include the square for adult, the 'X' for organoid under 30 days of differentiation, and the diamond for organoid over 30 days of differentiation. Major markers of retina progenitor, photoreceptors (cone and rod), and retinal ganglia cells (RGC) have simliar gene expression patterns across development in retina fetal tissue and organoids.

Differential gene expression of organoid retina versus fetal tissue identifies sets of genes relating to patterning (HOXB family), cell adhesion (protocadherin family), and RGC identity (BRN3/POU4F, NEFL, GAP43, SNCG)

To identify specific changes between retinal organoid and fetal retina tissue, we performed differential gene expression and GO term enrichment analyses. The GO term enrichment identified cell adhesion (protocadherins) and patterning (HOXB family) as enriched gene sets in retinal organoids As there is some evidence suggesting that protocadherins influence RGC viability CITATION and we noticed that several RGC markers appeared to have lower expression in the organoids compared to the fetal tissue Figure 5B we looked more closely into RGC marker expression[61].

We plotted HOXB family, protocadherin family and RGC genes in a heatmap visualization, with columns as age in days of fetal or organoid retina. Rows are genes, split by the three different groups of genes and are internally clustered by how similar the expression patterns are. We observed that there are strong, consistent gene expression differences in these three groups of genes between fetal retina and the organoid samples (Supplemental

Figure 4). We also plotted the differential expression values between all organoids and all fetal retina samples; all genes across all three groups are significantly differentially expressed with an FDR corrected p value < 0.01.

Data accessibility

Individual data files for gene expression and sample metadata can be downloaded from the 'Data' tab on the web app. All data and code used to generate the web app can be installed from the R command line by running devtools::install_github('davidmcg/eyeIntegration_app'). The code for the EiaD data processing pipeline can be found at https://github.com/davidmcg/Eyeintegration_autobuild.

Discussion

EiaD 2019 contains a large set of carefully curated, reproducibly processed human eye RNA-seq datasets alongside a human body tissue comparison set from the GTEx project. It is available for local install as an R package at https://www.github.com/davemcg/eyeIntegration_app and it is served via a web app, eyeIntegration at https://eyeIntegration.nei.nih.gov. The web app offers a wide range of user-driven visualizations to compare expression of genes across dozens of human body and eye tissues. Furthermore, murine scRNA-seq datasets have been incorporated, allowing for examination of retina cell type-specific gene expression. Several human and non-human primate studies have been posted in the past year on the pre-print server bioRxiv and as the raw data becomes publicly available, we will be updating this section of eyeIntegration[62–64].

As human fetal tissue is difficult to obtain and thus not very amenable for chemical or genetic modification, it is crucial for organoid-based models to be developed. Our merging of these datasets and analysis at the transcriptome level (as compared to cross-analyzing using a limited number of known marker genes) indicates that these two approaches successfully recapitulate fetal retina tissue, to a first approximation, at the whole transcriptome level. However as organoids do not develop to full function, it is important to look at how gene expression differs between retinal organoid and fetal tissue so as to suggest areas for improvement.

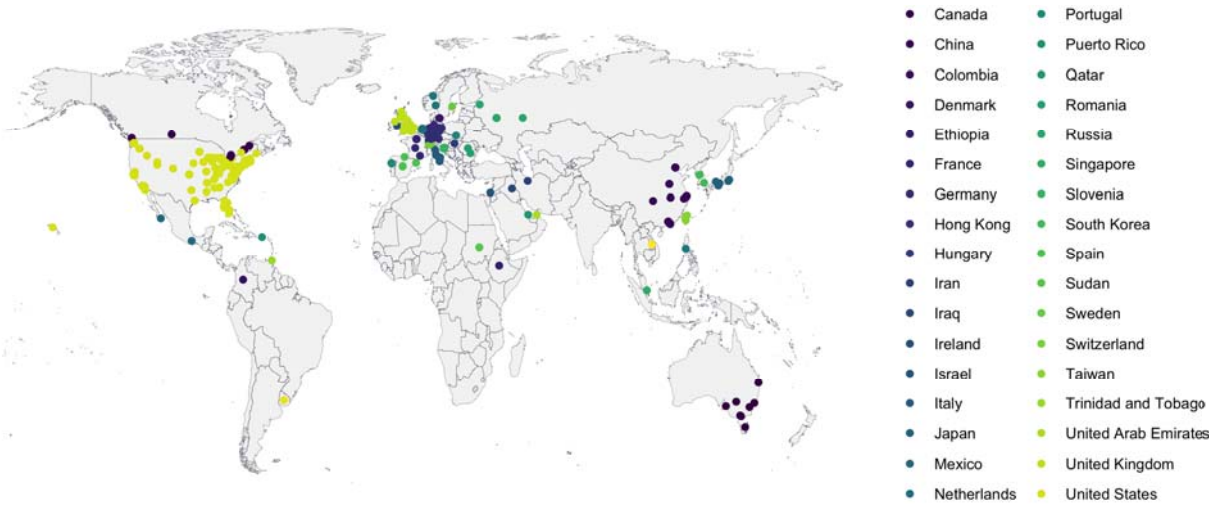
We used our large dataset to narrow in on three core processes which differ significantly and substantially between retinal organoids and fetal retina. First we showed that the HOXB family is overexpressed in the organoids. The homeobox family is well known to initiate polarity of the embryo during early development63[65]. Retinoic acid

317 is applied at about day 20 in culture to help differentiate stem cell to organoids and is also known to activate genes
318 members of the HOXB family. The lack of HOXB expression at any age in fetal retina and the broad chromatin and
319 gene expression changes HOXB family members can mediate suggests that HOXB activity may be unwanted for
320 organoid maturation.

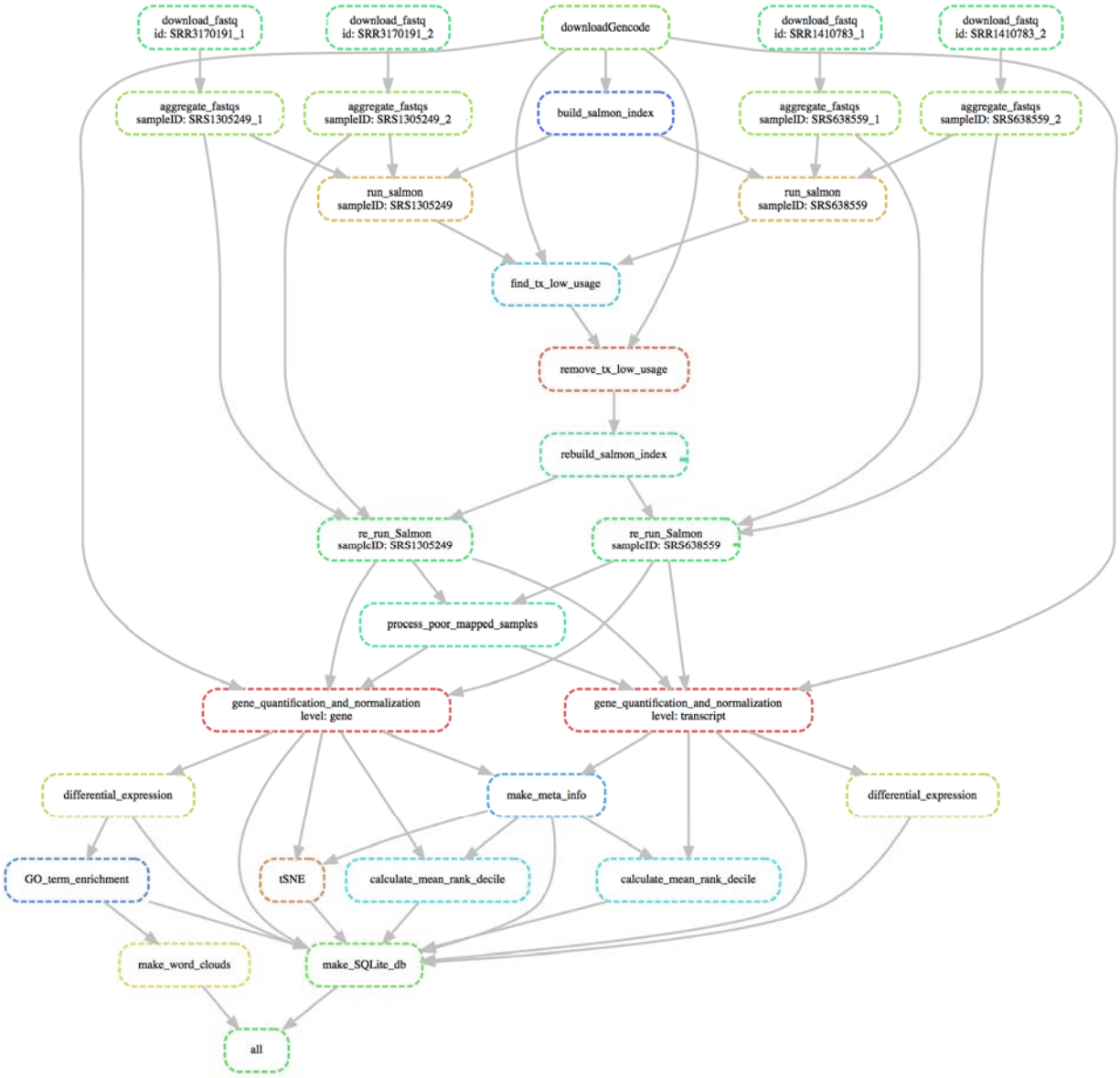
321 Next, we detected several protocadherins more highly expressed in the fetal tissue, relative to the
322 organoids. Protocadherins mediate cell to cell connections and, in the developing mouse, are shown to be important
323 for spinal internneurons and RGC survival[61,66]. We would predict that decreased protocadherin expression reduces
324 the number and maturation of RGC. Indeed we observed that many canonical RGC markers, while present in
325 detectable levels in the organoids, are signficantly underexpressed relative to fetal tissue. This result suggests that
326 modifying culture conditions to promote protocadherin expression may result in higher RGC yield and survival.

327 We built the Eye in a Disk dataset and the accompanying web app, eyeIntegration in the hopes that easily
328 accessible gene expression across tissue space and time will be a useful tool for hypothesis generation and
329 refinement in eye research. Wrapping all of the data processing steps in a Snakemake pipeline has several important
330 advantages for the community: our code is publicly available for review, our analyses are reproducible, future
331 sample updates can be streamlined in with less effort, and because all the processing is in modular pieces it is easier
332 to add new analysis steps. In the future, we plan on regularly adding new samples to EiaD, offering *de novo* eye
333 tissue transcriptomes, expanding the single cell RNA-seq expression tooling, adding non-human eye samples, and
334 epigenetic datasets.

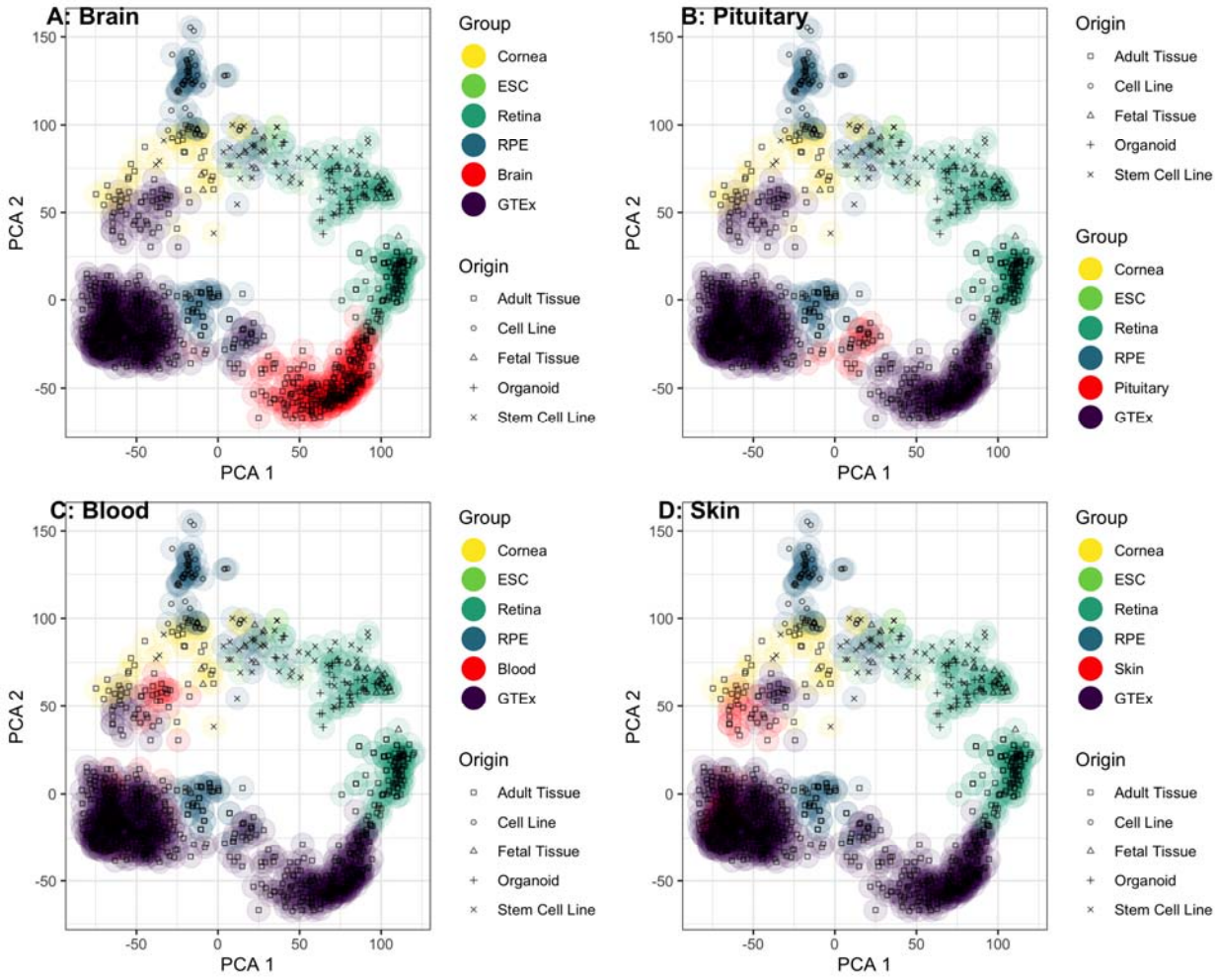
Supplementary Figures and Tables



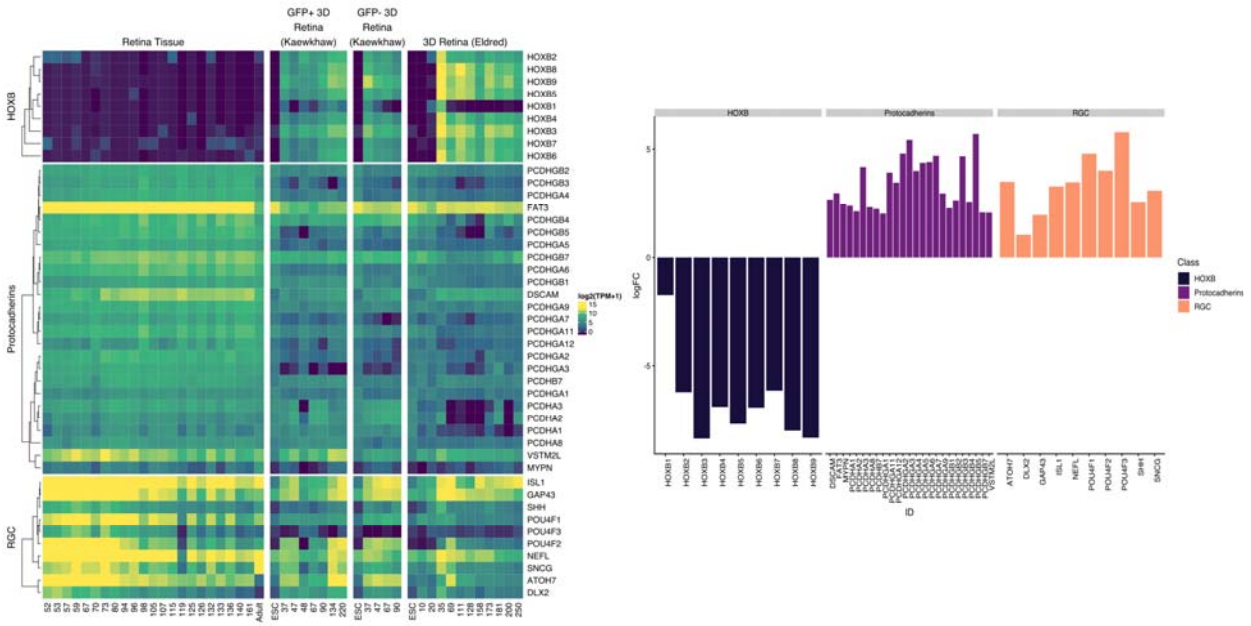
Supplemental Figure 1: As of January 2019, eyeIntegration has had usage across 285 cities and 40 countries.



Supplemental Figure 2: Snakemake pipeline to create EiaD 2019 consists of small modular compute sections to ensure sample tracking through the full pipeline



Supplemental Figure 3: PCA plot of all samples suggests that the non-eye tissue most similar to adult retina is the brain, pituitary is the most similar to adult RPE, and skin and blood are similar to cornea.



Supplemental Figure 4: Heatmap of three sets of genes across fetal retina and organoids divided by age in days (A). Bar plot of differential expression from retina to organoid, where positive values are genes that are higher expressed in fetal tissue than organoid (B). All logFC expression values have FDR corrected p value < 0.01.

Group	Pre QC Count	Count	Tissue Types (Count)
GTEx	878	848	Adipose Tissue (38), Adrenal Gland (20), Blood (33), Blood Vessel (58), Brain (252), Breast (20), Colon (40), Esophagus (60), Heart (38), Kidney (17), Liver (20), Lung (19), Muscle (19), Nerve (19), Pancreas (20), Pituitary (20), Salivary Gland (20), Skin (57), Small Intestine (20), Spleen (20), Stomach (20), Thyroid (18)

Supplemental Table 1: EiaD holds hundreds of GTEx tissues to provide a broad comparison set

Class	Count	Type
protein_coding	34836	Transcript
protein_coding	18968	Gene
retained_intron	6402	Transcript
lincRNA	6352	Transcript
antisense	6057	Transcript
lincRNA	4981	Gene
processed_transcript	4741	Transcript
antisense	4408	Gene

Class	Count	Type
nonsense_mediated_decay	3526	Transcript
processed_pseudogene	3050	Gene
processed_pseudogene	2604	Transcript
TEC	816	Transcript
TEC	809	Gene
sense_intronic	750	Gene
sense_intronic	734	Transcript
unprocessed_pseudogene	708	Gene
transcribed_unprocessed_pseudogene	690	Gene
unprocessed_pseudogene	590	Transcript
misc_RNA	557	Gene
misc_RNA	553	Transcript
processed_transcript	547	Gene
transcribed_processed_pseudogene	349	Gene
snoRNA	310	Transcript
transcribed_unprocessed_pseudogene	299	Transcript
snoRNA	295	Gene
miRNA	280	Gene
miRNA	279	Transcript
sense_overlapping	194	Transcript
snRNA	192	Transcript
snRNA	192	Gene
transcribed_processed_pseudogene	164	Transcript
sense_overlapping	160	Gene
IG_V_gene	133	Gene
IG_V_gene	132	Transcript
bidirectional_promoter_lncRNA	121	Transcript
	107	Transcript
transcribed_unitary_pseudogene	105	Gene
bidirectional_promoter_lncRNA	68	Gene

Class	Count	Type
TR_V_gene	55	Gene
rRNA_pseudogene	44	Transcript
rRNA_pseudogene	44	Gene
TR_V_gene	42	Transcript
transcribed_unitary_pseudogene	42	Transcript
IG_V_pseudogene	39	Transcript
IG_V_pseudogene	39	Gene
3prime_overlapping_ncRNA	30	Transcript
3prime_overlapping_ncRNA	27	Gene
unitary_pseudogene	23	Gene
Mt_tRNA	22	Transcript
Mt_tRNA	22	Gene
unitary_pseudogene	21	Transcript
polymorphic_pseudogene	20	Transcript
pseudogene	19	Transcript
polymorphic_pseudogene	19	Gene
scaRNA	18	Gene
IG_C_gene	17	Transcript
scaRNA	17	Transcript
IG_J_gene	14	Transcript
IG_C_gene	14	Gene
IG_J_gene	14	Gene
pseudogene	12	Gene
TR_J_gene	10	Gene
non_stop_decay	9	Transcript
TR_J_gene	8	Transcript
TR_C_gene	6	Transcript
TR_C_gene	6	Gene
rRNA	5	Transcript
rRNA	5	Gene
IG_C_pseudogene	3	Transcript
TR_V_pseudogene	3	Transcript
IG_C_pseudogene	3	Gene
TR_V_pseudogene	3	Gene

Class	Count	Type
Mt_rRNA	2	Transcript
non_coding	2	Transcript
ribozyme	2	Transcript
translated_processed_pseudogene	2	Transcript
Mt_rRNA	2	Gene
non_coding	2	Gene
ribozyme	2	Gene
translated_processed_pseudogene	2	Gene
IG_J_pseudogene	1	Transcript
macro_lncRNA	1	Transcript
scRNA	1	Transcript
vaultRNA	1	Transcript
IG_J_pseudogene	1	Gene
macro_lncRNA	1	Gene
scRNA	1	Gene
vaultRNA	1	Gene

349 *Supplemental Table 2: Dozens of different types of gene and transcript types quantified*

sample_accession	study_accession	study_title	study_abstract	sample_attribute	Tissue	Sub_Tissue	Origin	Age_Days	Kept
SRS1146040	SRP064956	mammalian cell-lines Transcriptome or Gene expression	Gene expression studies of two mammalian cell lines that display primary cilia. Primary cilia are microtubule-based organelles that are hubs for receiving and intergrating signalling cascades during embryonic development and in maintaining tissue homeostasis. Defects in the structure of function of cilia can cause a group of comparitively common human inherited conditions known as ciliopathies. The cell-lines that have been profiled are standard model systems in the ciliary biology field, and gene expression profiles are important to understand the processes of forming, maintaining and resorbing cilia.	isolate: normal || age: adult || biomaterial_provider: ATCC CRL-4000 || sex: female || tissue: retina, eye; pigmented epithelium || cell_line: hTERT RPE-1 || cell_type: epithelial cells immortalized with hTERT || disease: normal || phenotype: epithelial-like || sample_type: ciliated serum-starved hTERT-RPE1 cells, passage 8, sample 1 || BioSampleModel: Human	RPE	RPE - Cell Line	Cell Line	.	removed
SRS3729714	SRP159246	Thyroid hormone signaling specifies cone subtypes in human retinal organoids		Day_20_1	Retina	Retina - 3D Organoid Stem Cell	Stem Cell	20	Kept

sample_accession	study_accession	study_title	study_abstract	sample_attribute	Tissue	Sub_Tissue	Origin	Age_Days	Kept
SRS2582162	SRP119766	Molecular anatomy of the developing human retina		D96P	Retina	Retina - Fetal Tissue	Fetal Tissue	96	Kept
SRS2305609	SRP110135	Characterization of RPE lincRNA	Long intervening non-coding RNAs (lincRNAs) are increasingly being implicated as important factors in many aspects of cellular development, function, and disease, but remain poorly understood. In this study, we examine the human retinal pigment epithelium (RPE) lincRNA transcriptome using RNA-Seq data generated from human fetal RPE (fRPE), RPE derived from human induced pluripotent stem cells (iPS-RPE), and undifferentiated iPS (iPS). In addition, we determine the suitability of iPS-RPE, from a transcriptome standpoint, as a model for use in future studies of lincRNA structure and function. A comparison of gene and isoform expression across the whole transcriptome shows only minimal differences between all sample types, though fRPE and iPS-RPE show higher concordance than either shows with iPS. Notably, RPE signature genes show the highest degree of fRPE to iPS-RPE concordance, indicating that iPS-RPE cells provide a suitable model for use in future studies. An analysis of lincRNAs demonstrates high concordance between fRPE and iPS-RPE, but low concordance between either RPE and iPS. While most lincRNAs are expressed at low levels (RPKM < 10), there is a high degree of concordance among replicates within each sample type, suggesting the expression is consistent, even at levels subject to high variability. Finally, we identified and annotated 180 putative novel genes in the fRPE samples, a majority of which are also expressed in the iPS-RPE. Overall, this study represents the first characterization of lincRNA expression in the human RPE, and provides a model for studying the role lincRNAs play in RPE development, function, and disease.	isolate: iPS cell || age: 90 days || biomaterial_provider: Chad Cowan, Harvard University, Sherman Fairchild 153, 7 Divinity Ave, Cambridge, MA 02138 || sex: not determined || tissue: iPS cell || BioSampleModel: Human	RPE	RPE - Stem Cell Line	Stem Cell	.	Kept
SRS1572428	SRP079002	Enriched retinal ganglion cells derived from human embryonic stem cells (RNA-seq)	We performed the whole transcriptome analysis in human H9 ES cells and in retinal ganglion cells (RGC) derived from H9 ES cells. Overall design: RNA-seq in H9 human ESCs and RGC derived from H9 ESCs	source_name: RGC from H9 ES cells || cell type: Differentiated ES cells	Retina	Retina - RGC Stem Cell	Stem Cell	.	Kept
SRS5238111	SRP034875	Comprehensive Analysis of Gene Expression in Human Retina and Supporting Tissues	RNA-Seq of retina and RPE/choroid/sclera from 8 normal human eyes	tissue_type: Peripheral RPE Choroid Sclera_3	RPE	RPE - Adult Tissue	Adult Tissue	.	Kept
SRS1305245	SRP070148	Transcriptome profiling of human keratoconus corneas through RNA sequencing identifies collagen synthesis disruption and downregulation of core elements of TGF-ÃƒÂ„, Hippo, and Wnt pathways	To understand better the factors contributing to keratoconus (KTCN), we used RNA sequencing to perform a transcriptome profile of human KTCN corneas. Over 82% of the genes and almost 75% of the transcripts detected as differentially expressed in KTCN and non-KTCN corneas were confirmed in the replication study using another set of samples. We used these differentially expressed genes to generate a network of KTCN-deregulated genes. We found an extensive disruption of collagen synthesis and maturation pathways, as well as downregulation of the core elements of the TGF-ÃƒÂ„, Hippo, and Wnt signaling pathways influencing corneal organization. We identified long noncoding RNAs (lncRNAs) and conducted a computational analysis of their potential functions, and found that lncRNAs regulated the processing and expression of the aforementioned genes. This first comprehensive transcriptome profiling of human KTCN corneas points further to a complex etiology of KTCN. Overall design: Transcription profiling of 25 KTCN and 25 non-KTCN corneas using RNA-Seq	source_name: Cornea || disease state: non-KTCN || study: replication || tissue: cornea	Cornea	Cornea - Adult Tissue	Adult Tissue	.	Kept

sample_accession	study_accession	study_title	study_abstract	sample_attribute	Tissue	Sub_Tissue	Origin	Age_Days	Kept
SRS846899	SRP055101	Transcriptomic analysis of cultured corneal endothelial cells as a validation for their use in cell-replacement therapy	The corneal endothelium plays a primary role in maintaining corneal homeostasis and clarity, and must be surgically replaced with allogenic donor corneal endothelium in the event of visually significant dysfunction. However, a worldwide shortage of donor corneal tissue has led to a search for alternative sources of transplantable tissue. Cultured human corneal endothelial cells (HCEnC) have been shown to restore corneal clarity in experimental models of corneal endothelial dysfunction in animal models, but characterization of cultured HCEnC remains incomplete. To this end, we utilized next-generation RNA sequencing technology to compare the transcriptomic profile of ex vivo human corneal endothelium (evHCEnC) with that of primary HCEnC and HCEnC lines, and to determine the utility of cultured and immortalized corneal endothelial cells as models of in vivo corneal endothelium. Multidimensional analyses of the transcriptome datasets demonstrated that primary HCEnC have a closer relationship to evHCEnC than do immortalized HCEnC. Subsequent analyses showed that the majority of the genes specifically expressed in HCEnC (not expressed in ex vivo corneal epithelium or fibroblasts) demonstrated a marked variability of expression in cultured cells compared with evHCEnC. In addition, genes associated with either corneal endothelial cell function or corneal endothelial dystrophies were investigated. Significant differences in gene expression and protein levels were observed in the cultured cells compared with evHCEnC for each of the genes tested except for AGBL1 and LOXHD1, which were not detected by RNA-seq or qPCR. Our transcriptomic analysis suggests that at a molecular level primary HCEnC most closely resemble evHCEC and thus represent a viable therapeutic option for managing corneal endothelial dysfunction. Our findings also suggest that investigators should perform an assessment of the entire transcriptome of cultured HCEnC prior to determination of the potential clinical utility of the cultured HCEnC for the management of corneal endothelial cell failure. Overall design: Transcriptomes from ex vivo corneal endothelium, primary cultures and three cell lines were compared. Three samples of each endothelial cell group were submitted for RNA sequencing for a total of 15 samples. The transcriptome for the ex vivo corneal endothelium was used as the reference (i.e., proxy for in vivo corneal endothelium). Transcript abundances for a subset of genes associated with corneal endothelial cell function or disease were validated with qPCR and western blot. Samples of ex vivo endothelium used for validation were independent replicates not used for RNA-sequencing.	source_name: primary corneal endothelial cells || cell type: endothelial || imortalization: none || culture medium: F99	Cornea	Cornea - Endothelium	Adult Tissue	.	Kept
SRS1747748	SRP091605	Transcriptomes during lens differentiation of human embryonic stem cells	In order to examine the fundamental mechanisms governing lens cells differentiation, we analyzed the transcriptome changes during the differentiation of human embryonic stem cells (hESCs) into lens cells and lentoid bodies. The differentiation of hESCs was induced by a sequential treatments of growth factors. In briefly,Noggin (100ng/ul) was applied from day0 to day6; then a combination ofbFGF (100ng/ul) and BMP4/7 (20 ng/ml)was addedfrom day 7 to day 18, followed by bFGF (100ng/ul) and Wnt3a(20 ng/ml) from day19 to day32. Cells at day 0, day 6, day 18 and day 32 were collected for analysis of paired-end RNA sequencing using Illuminia Hiseq 2500. The results revealed dynamic transcription network during lentoid bodies differentiation. We observed differential expression of genes involved insignaling pathways, which were considered to be necessary for lens development. These results provide a valuable resource for studying the mechanisms regulating in vitro lentoid body differentiation of hESCs and getting a glimpse of signaling pathway network in lens embryonic development.	isolate: human embryonic stem cells treated with 100ng/ul noggin for 6 days, followed by 100ng/ul bFGF, 20 ng/ml BMP4, and 20 ng/ml BMP7 from day6 to day18 || age: not applicable || biomaterial_provider: Zhongshan Ophthalmic Centre, 57th Xianlie Road, Guangzhou, China || sex: not applicable || tissue: not applicable || cell_line: H9 human embryonic stem cell || cell_type: human embryonic stem cell || BioSampleModel: Human	Lens	Lens - Stem Cell Line	Stem Cell	.	low salmon mapping rate
SRS3117114	SRP012585	Identification of miRNA signatures during the differentiation of hESCs into retinal pigment epithelial cells	Retinal pigment epithelium (RPE) cells can be obtained through in vitro differentiation of both embryonic stem cell (ESC) and induced pluripotent stem cells (iPSC) for cell replacement therapy. We have previously identified 87 signature genes relevant to RPE cell differentiation and function through transcriptome analysis of both human ESC- and iPSC-derived RPE as well as normal fetal RPE. Here, we profiled miRNA expression through small RNA-seq in human ESCs and their RPE derivatives. Much like conclusions drawn from our previous transcriptome analysis, we found that the overall miRNA landscape in RPE is distinct from ESCs and other differentiated somatic tissues. We also profiled miRNA expression during intermediate stages of RPE differentiation and identified unique subsets of miRNAs that are gradually up- or downregulated, suggesting dynamic regulation of these miRNAs is associated with the RPE differentiation process. Indeed, the down-regulation of a subset of miRNAs during RPE differentiation is associated with up-regulation of RPE-specific genes, such as RPE65, which is exclusively expressed in RPE. We conclude that miRNA signatures can be used to classify different degrees of in vitro differentiation of RPE from human pluripotent stem cells. We suggest that RPE-specific miRNAs likely contribute to the functional maturation of RPE in vitro, similar to the regulation of RPE-specific mRNA expression. Overall design: Study miRNA in ESC-derived RPE	source_name: ES-RPE || cell type: ES-RPE || parental cell line: hiPS2	RPE	RPE - Stem Cell Line	Stem Cell	.	low salmon mapping rate

sample_accession	study_accession	study_title	study_abstract	sample_attribute	Tissue	Sub_Tissue	Origin	Age_Days	Kept
SRS1955495	SRP098761	Comprehensive analysis of gene expression in human retina and supporting tissues	Purpose: To identify the differences of gene expression between the macula and periphery of the eye. Methods: RNA Seq was performed on 8 normal postmortem eyes. Results: Significant differential expression was found between the layers of the posterior part of the eye and also between locations of a tissue layer. Conclusions: These results potentially provide a path to more rapidly elucidate not only the genetic basis of eye disease but also the impact of gene expression on molecular networks that in turn induce variations in disease associated traits. Overall design: Retinal and RPE/choroid/sclera RNA profiles from macula and periphery of each eye were generated by deep sequencing on an Illumina Hi-Seq. 32 biological samples were analyzed.	source_name: Peripheral Retina || tissue: Peripheral Retina || age: 88 || Sex: female	Retina	Retina - Adult Tissue	Adult Tissue	.	Kept
SRS1955499	SRP098761	Comprehensive analysis of gene expression in human retina and supporting tissues	Purpose: To identify the differences of gene expression between the macula and periphery of the eye. Methods: RNA Seq was performed on 8 normal postmortem eyes. Results: Significant differential expression was found between the layers of the posterior part of the eye and also between locations of a tissue layer. Conclusions: These results potentially provide a path to more rapidly elucidate not only the genetic basis of eye disease but also the impact of gene expression on molecular networks that in turn induce variations in disease associated traits. Overall design: Retinal and RPE/choroid/sclera RNA profiles from macula and periphery of each eye were generated by deep sequencing on an Illumina Hi-Seq. 32 biological samples were analyzed.	source_name: Peripheral Retina || tissue: Peripheral Retina || age: 95 || Sex: female	Retina	Retina - Adult Tissue	Adult Tissue	.	Kept
SRS523806	SRP034875	Comprehensive Analysis of Gene Expression in Human Retina and Supporting Tissues	RNA-Seq of retina and RPE/choroid/sclera from 8 normal human eyes	tissue_type: Macular Retina_2	Retina	Retina - Adult Tissue	Adult Tissue	.	Kept
SRS1597880	SRP080886	Region-specific Transcriptome Analysis of the Human Retina and RPE/Choroid	Proper spatial differentiation of retinal cell types is necessary for normal human vision. Many retinal diseases, such as Best disease and male germ cell associated kinase (MAK)-associated retinitis pigmentosa, preferentially affect distinct topographic regions of the retina. While much is known about the distribution of cell types in the retina, the distribution of molecular components across the posterior pole of the eye has not been well-studied. To investigate regional difference in molecular composition of ocular tissues, we assessed differential gene expression across the temporal, macular, and nasal retina and retinal pigment epithelium (RPE)/choroid of human eyes using RNA-Seq. RNA from temporal, macular, and nasal retina and RPE/choroid from four human donor eyes was extracted, poly-A selected, fragmented, and sequenced as 100 bp read pairs. Digital read files were mapped to the human genome and analyzed for differential expression using the Tuxedo... (for more see dbGaP study page.)	gap_accession: phs001151 || submitter handle: NEI_RetinaRPE_Choroid || biospecimen repository: NEI_RetinaRPE_Choroid || study name: Region-specific Transcriptome Analysis of the Human Retina and RPE/Choroid || study design: Control Set || biospecimen repository sample id: S27 || submitted sample id: S27 || submitted subject id: 5 || gap_sample_id: 1975640 || gap_subject_id: 1622646 || sex: female || body site: eye || histological type: neural retina || analyte type: RNA || is tumor: No || subject is affected: No || molecular data type: RNA Seq (NGS) || gap_consent_code: 1 || gap_consent_short_name: GRU-NPU	Retina	Retina - Adult Tissue	Adult Tissue	.	Kept
E-MTAB-4377.RNA30	E-MTAB-4377	RNAseq 50 Normal Human Retina	RNA-seq of post-mort retina donor without clinically relevant visual impairment. Ploy-A enriched. 75-nt paired-end. Short time lapse between tissue sampling and cDNA generation.		Retina	Retina - Adult Tissue	Adult Tissue	.	Kept
SRS3729729	SRP159246	Thyroid hormone signaling specifies cone subtypes in human retinal organoids		Day_158_2	Retina	Retina - 3D Organoid Stem Cell	Stem Cell	158	Kept

sample_accession	study_accession	study_title	study_abstract	sample_attribute	Tissue	Sub_Tissue	Origin	Age_Days	Kept
SRS1955480	SRP098761	Comprehensive analysis of gene expression in human retina and supporting tissues	Purpose: To identify the differences of gene expression between the macula and periphery of the eye. Methods: RNA Seq was performed on 8 normal postmortem eyes. Results: Significant differential expression was found between the layers of the posterior part of the eye and also between locations of a tissue layer. Conclusions: These results potentially provide a path to more rapidly elucidate not only the genetic basis of eye disease but also the impact of gene expression on molecular networks that in turn induce variations in disease associated traits. Overall design: Retinal and RPE/choroid/sclera RNA profiles from macula and periphery of each eye were generated by deep sequencing on an Illumina Hi-Seq. 32 biological samples were analyzed.	source_name: Macular Retina || tissue: Macular Retina || age: 79 || Sex: male	Retina	Retina - Adult Tissue	Adult Tissue	.	Kept
SRS3729720	SRP159246	Thyroid hormone signaling specifies cone subtypes in human retinal organoids		Day_69_2	Retina	Retina - 3D Organoid Stem Cell	Stem Cell	69	Kept
SRS1597871	SRP080886	Region-specific Transcriptome Analysis of the Human Retina and RPE/Choroid	Proper spatial differentiation of retinal cell types is necessary for normal human vision. Many retinal diseases, such as Best disease and male germ cell associated kinase (MAK)-associated retinitis pigmentosa, preferentially affect distinct topographic regions of the retina. While much is known about the distribution of cell types in the retina, the distribution of molecular components across the posterior pole of the eye has not been well-studied. To investigate regional difference in molecular composition of ocular tissues, we assessed differential gene expression across the temporal, macular, and nasal retina and retinal pigment epithelium (RPE)/choroid of human eyes using RNA-Seq. RNA from temporal, macular, and nasal retina and RPE/choroid from four human donor eyes was extracted, poly-A selected, fragmented, and sequenced as 100 bp read pairs. Digital read files were mapped to the human genome and analyzed for differential expression using the Tuxedo... (for more see dbGaP study page.)	gap_accession: phs001151 || submitter handle: NEI_RetinaRPE_Choroid || biospecimen repository: NEI_RetinaRPE_Choroid || study name: Region-specific Transcriptome Analysis of the Human Retina and RPE/Choroid || study design: Control Set || biospecimen repository sample id: S18 || submitted sample id: S18 || submitted subject id: 3 || gap_sample_id: 1975630 || gap_subject_id: 1622644 || sex: female || body site: eye || histological type: RPE/choroid || analyte type: RNA || is tumor: No || subject is affected: No || molecular data type: RNA Seq (NGS) || gap_consent_code: 1 || gap_consent_short_name: GRU-NPU	RPE	RPE - Adult Tissue	Adult Tissue	.	Kept
SRS1955491	SRP098761	Comprehensive analysis of gene expression in human retina and supporting tissues	Purpose: To identify the differences of gene expression between the macula and periphery of the eye. Methods: RNA Seq was performed on 8 normal postmortem eyes. Results: Significant differential expression was found between the layers of the posterior part of the eye and also between locations of a tissue layer. Conclusions: These results potentially provide a path to more rapidly elucidate not only the genetic basis of eye disease but also the impact of gene expression on molecular networks that in turn induce variations in disease associated traits. Overall design: Retinal and RPE/choroid/sclera RNA profiles from macula and periphery of each eye were generated by deep sequencing on an Illumina Hi-Seq. 32 biological samples were analyzed.	source_name: Peripheral Retina || tissue: Peripheral Retina || age: 62 || Sex: male	Retina	Retina - Adult Tissue	Adult Tissue	.	Kept

Supplemental Table 3: Full metadata for 20 random eye samples. Full metadata available as supplementary file.

ID	Description	GeneRatio	pvalue	qvalue	geneID

ID	Description	GeneRatio	pvalue	qvalue	geneID
GO:0030198	extracellular matrix organization	55/741	6.68e-18	0.000	P4HA1 LOX PLOD2 VCAM1 FOXC1 HAS2 ICAM1 NID2 WT1 POSTN DDR2 COL1A2 SULF1 PDPN COL3A1 COL14A1 COL27A1 COL5A1 FOXF1 DPP4 KLK4 ECM2 SMOC2 COL4A6 FN1 DCN ADAMTS5 ITGA8 TGFBI COL1A1 ITGB6 ITGB3 CDH1 ICAM5 SERPINE1 TNFRSF11B FMOD ITGB4 MMP7 COL6A3 MMP9 COL8A1 MMP1 NPNT CCDC80 FBLN5 CTGF HPSE2 GFAP SPINT1 FGG TTR BGN THBS1 ELF3
GO:0061448	connective tissue development	39/741	3.12e-12	0.000	HOXB3 HOXC4 HOXA3 BMPR1B HOXA5 ALX1 WT1 WNT11 PLA2G16 GDF7 MSX2 SNAI2 MGP SULF1 FOXA1 STC1 COL14A1 COL27A1 NKX3-2 PRRX1 COL5A1 EGR1 OSR2 TRPV4 SNAI1 BMP5 TGFBI HOXD3 RUNX2 COL1A1 EFEMP1 MATN2 HAND1 SIX2 HAND2 COL6A3 FRZB ACTA2 CTGF
GO:0001503	ossification	46/741	3.68e-11	0.000	HGF EGR2 GPC3 IFITM1 HOXA3 BMPR1B DHRS3 IGFBP5 FOXC1 WNT11 KLF10 DDR2 MSX2 FHL2 SNAI2 MGP STC1 PHOSPHO1 GPNMB CEBPB KREMEN2 IGFBP3 ASPN ALOX15 PTGS2 OSR2 PTN FGFR2 SNAI1 ISG15 DLX5 BMP5 HOXA2 RUNX2 COL1A1 AREG IGF2 ALPL WNT4 SIX2 HAND2 MMP9 RANBP3L NPNT IL6 CTGF
GO:0007156	homophilic cell adhesion via plasma membrane adhesion molecules	25/536	3.90e-11	0.000	PCDHGA6 PCDHGA5 PCDHGB5 PCDHGB3 PCDHGA7 PCDHGA2 PCDHGA3 PCDHGA4 PCDHGA11 PCDHGA9 PCDHGB2 PCDHA3 FAT3 PCDHGA12 PCDHGB4 PCDHB7 PCDHGB7 PCDHGA1 PCDHGB1 DSCAM PCDHA8 VSTM2L PCDHA1 PCDHA2 MYPN
GO:0050878	regulation of body fluid levels	51/741	8.00e-10	0.000	FZD6 CYP26B1 GUCA1B HAS2 ADM HK2 HEG1 TFPI2 SLC7A11 LYN XDH COL1A2 PDPN COL3A1 DOCK6 PRTN3 ENTPD2 PLAT PLAU CLDN4 EDNRB AQP4 F2RL1 TRPV4 SCNN1A CLDN1 PROCR CD9 FOXA2 COL1A1 FOXB1 APOE ITGB3 CAV1 NFE2 GATA5 EHD2 SERPINE1 SERPINB2 KDF1 SCNN1B PRKCG GATA4 GRHL3 CEACAM1 IL6 AGR2 FGG AQP1 THBS1 SERPINA1
GO:0031589	cell-substrate adhesion	40/741	1.49e-09	0.000	VCAM1 GBP1 HAS2 NID2 POSTN ACTN3 PDPN COL3A1 TNFRSF12A RHOD FOXF1 ALOX15 HOXA7 ECM2 EPHA1 PLAU PTN SMOC2 WNT1 FN1 ITGA8 HOXD3 ANGPT1 COL1A1 ITGB6 ITGB3 SKAP1 WNT4 SERPINE1 ITGB4 LGALS1 COL8A1 NPNT CCDC80 FBLN5 CEACAM6 CTGF AGR2 FGG THBS1

ID	Description	GeneRatio	pvalue	qvalue	geneID
GO:0030335	positive regulation of cell migration	50/741	1.50e-09	0.000	HGF LGALS3 EGF IGFBP5 HAS2 ICAM1 VEGFD WNT11 POSTN LYN ANXA1 DDR2 SEMA3C SNAI2 PDPN TMEM102 RHOD GPNMB FOXF1 PTGS2 EPHA1 PLAU FOXC2 GLIPR2 TNFAIP6 F2RL1 SMOC2 TRPV4 SNAI1 FN1 IL1R1 ANGPT1 COL1A1 FAM110C ITGB3 ENPP2 CAV1 EDN3 LRRC15 SERPINE1 ANXA3 PTP4A1 SRPX2 MMP9 GRB7 S100A14 IL6 CEACAM6 CXCL8 THBS1
GO:0050678	regulation of epithelial cell proliferation	37/741	1.21e-08	0.000	GPC3 HOXA5 BAX PAX2 HAS2 SIX4 SIX1 KLF9 VEGFD XDH SNAI2 SULF1 HTR2B DLX6 EDNRB OSR2 PTN FGFR2 PPARG DLX5 BMP5 ALDH1A2 APOE ITGB3 CAV1 EGR3 ODAM KDF1 CCL2 ZFP36 EGFL7 ESRP2 SCG2 LIMS2 CEACAM1 THBS1 NR4A1
GO:0030099	myeloid cell differentiation	30/536	5.10e-06	0.001	POU4F1 LDB1 POU4F2 SIGLEC15 HIST1H3C HIST1H4L HIST1H4F HIST1H3F HIST1H3B HIST1H3I PRKCB TREM2 HIST1H3J HIST1H4C TRIM58 C1QC LTBR FES TYROBP HIST1H4A GPR68 HIST1H4J HIST1H3A CSF1R CCL3 SPI1 EVI2B HIST1H4B HIST1H4D LILRB4
GO:0030099	myeloid cell differentiation	34/741	7.56e-05	0.001	HOXB8 HIST1H4K PIR HOXB7 GPC3 LGALS3 HOXA5 TLR3 LYN KLF10 UBD EPO PRTN3 INHBA CEBPB MT1G HOXA7 RELB F2RL1 LIF B2M PPARG ISG15 HOXA9 CA2 NFE2 ZFP36 MMP9 TESC MB IRF4 CEACAM1 FOS THBS1
GO:0050769	positive regulation of neurogenesis	28/536	2.64e-04	0.015	POU4F2 MYB TRIM67 ASCL1 ISLR2 HEYL PAX6 P2RY12 VWC2L PLXNA2 NEFL STMN2 CUX2 DSCAM RIT2 CX3CR1 SHH ALKAL2 FES RELN NEUROG3 ADRA2B GDF6 OLIG2 RIMS1 CPNE6 IRX3 NEUROD2

Supplemental Table 4: Top GO terms enriched between fetal retina and organoid retina, after paring of redundant terms with REVIGO

Acknowledgements

We would like the thank the dozens of groups who provided the raw data required to create this project. We keep a running list of the projects and associated citations at https://github.com/davemcg/eyeIntegration_app/blob/master/inst/citations.md and strongly encourage anyone who uses EiaD and eyeIntegration to cite relevant projects. We would also like to thank Kapil Bharti, Robert Hufnagel,

and Brian Brooks for their continuous set of critiques and suggestions in the development of eyeIntegration app over the past two years. Tiziana Cogliati was especially helpful in the editing of this manuscript. Finally, this work utilized the computational resources of the NIH HPC Biowulf cluster (http://hpc.nih.gov).

Funding

This research was supported by the Intramural Research Program of the National Eye Institute, National Institutes of Health.

Bibliography

1. Chen Y, Huang K, Nakatsu MN, Xue Z, Deng SX, Fan G. Identification of novel molecular markers through transcriptomic analysis in human fetal and adult corneal endothelial cells. *Human Molecular Genetics*. 2013;22(7):1271-1279. doi:10.1093/hmg/dds527

2. Chng Z, Peh GSL, Herath WB, et al. High Throughput Gene Expression Analysis Identifies Reliable Expression Markers of Human Corneal Endothelial Cells. *PLOS ONE*. 2013;8(7):e67546. doi:10.1371/journal.pone.0067546

3. Chung DD, Frausto RF, Lin BR, Hanser EM, Cohen Z, Aldave AJ. Transcriptomic Profiling of Posterior Polymorphous Corneal Dystrophy. *Investigative Ophthalmology & Visual Science*. 2017;58(7):3202-3214. doi:10.1167/iovs.17-21423

4. Frausto RF, Le DJ, Aldave AJ. Transcriptomic Analysis of Cultured Corneal Endothelial Cells as a Validation for Their Use in Cell Replacement Therapy. *Cell Transplantation*. 2016;25(6):1159-1176. doi:10.3727/096368915X688948

5. Kabza M, Karolak JA, Rydzanicz M, et al. Collagen synthesis disruption and downregulation of core elements of TGF-β, Hippo, and Wnt pathways in keratoconus corneas. *European Journal of Human Genetics*. 2017;25(5):582-590. doi:10.1038/ejhg.2017.4

6. Ouyang H, Xue Y, Lin Y, et al. WNT7A and PAX6 define corneal epithelium homeostasis and pathogenesis. *Nature*. 2014;511(7509):358-361. doi:10.1038/nature13465

7. Song Q, Yuan S, An Q, et al. Directed differentiation of human embryonic stem cells to corneal endothelial cell-like cells: A transcriptomic analysis. *Experimental Eye Research*. 2016;151:107-114. doi:10.1016/j.exer.2016.08.004

8. Han C, Li J, Wang C, et al. Wnt5a Contributes to the Differentiation of Human Embryonic Stem Cells into Lentoid Bodies Through the Noncanonical Wnt/JNK Signaling Pathway. *Investigative Ophthalmology & Visual Science*. 2018;59(8):3449-3460. doi:10.1167/iovs.18-23902

9. Aldiri I, Xu B, Wang L, et al. The Dynamic Epigenetic Landscape of the Retina During Development, Reprogramming, and Tumorigenesis. *Neuron*. 2017;94(3):550-568.e10. doi:10.1016/j.neuron.2017.04.022

10. Farkas MH, Grant GR, White JA, Sousa ME, Consugar MB, Pierce EA. Transcriptome analyses of the human retina identify unprecedented transcript diversity and 3.5 Mb of novel transcribed sequence via significant alternative splicing and novel genes. *BMC Genomics*. 2013;14(1):486. doi:10.1186/1471-2164-14-486

11. Hoshino A, Ratnapriya R, Brooks MJ, et al. Molecular Anatomy of the Developing Human Retina. *Developmental Cell*. 2017;43(6):763-779.e4. doi:10.1016/j.devcel.2017.10.029

12. Kaewkhaw R, Kaya KD, Brooks M, et al. Transcriptome Dynamics of Developing Photoreceptors in Three-Dimensional Retina Cultures Recapitulates Temporal Sequence of Human Cone and Rod Differentiation Revealing Cell Surface Markers and Gene Networks. *Stem Cells (Dayton, Ohio)*. 2015;33(12):3504-3518. doi:10.1002/stem.2122

13. Kaewkhaw R, Swaroop M, Homma K, et al. Treatment Paradigms for Retinal and Macular Diseases Using 3-D Retina Cultures Derived From Human Reporter Pluripotent Stem Cell Lines. *Investigative Ophthalmology & Visual Science*. 2016;57(5):ORSFl1-ORSFl11. doi:10.1167/iovs.15-17639

14. Li M, Jia C, Kazmierkiewicz KL, et al. Comprehensive analysis of gene expression in human retina and supporting tissues. *Human Molecular Genetics*. 2014;23(15):4001-4014. doi:10.1093/hmg/ddu114

403 15. Mustafi D, Kevany BM, Bai X, et al. Transcriptome analysis reveals rod/cone photoreceptor specific signatures
404 across mammalian retinas. *Human Molecular Genetics*. 2016;25(20):4376-4388. doi:10.1093/hmg/ddw268

405 16. Pinelli M, Carissimo A, Cutillo L, et al. An atlas of gene expression and gene co-regulation in the human retina.
406 *Nucleic Acids Research*. 2016;44(12):5773-5784. doi:10.1093/nar/gkw486

407 17. Whitmore SS, Wagner AH, DeLuca AP, et al. Transcriptomic analysis across nasal, temporal, and macular
408 regions of human neural retina and RPE/choroid by RNA-Seq. *Experimental Eye Research*. 2014;129:93-106.
409 doi:10.1016/j.exer.2014.11.001

410 18. Darrow EM, Huntley MH, Dudchenko O, et al. Deletion of DXZ4 on the human inactive X chromosome alters
411 higher-order genome architecture. *Proceedings of the National Academy of Sciences*. 2016;113(31):E4504-E4512.
412 doi:10.1073/pnas.1609643113

413 19. Harenza JL, Diamond MA, Adams RN, et al. Transcriptomic profiling of 39 commonly-used neuroblastoma cell
414 lines. *Scientific Data*. 2017;4:170033. doi:10.1038/sdata.2017.33

415 20. Hu G, Huang K, Yu J, et al. Identification of miRNA Signatures during the Differentiation of hESCs into Retinal
416 Pigment Epithelial Cells. *PLOS ONE*. 2012;7(7):e37224. doi:10.1371/journal.pone.0037224

417 21. Nozawa R-S, Boteva L, Soares DC, et al. SAF-A Regulates Interphase Chromosome Structure through
418 Oligomerization with Chromatin-Associated RNAs. *Cell*. 2017;169(7):1214-1227.e18.
419 doi:10.1016/j.cell.2017.05.029

420 22. Oberstein A, Shenk T. Cellular responses to human cytomegalovirus infection: Induction of a mesenchymal-to-
421 epithelial transition (MET) phenotype. *Proceedings of the National Academy of Sciences*. 2017;114(39):E8244-
422 E8253. doi:10.1073/pnas.1710799114

423 23. Peng S, Gan G, Qiu C, et al. Engineering a Blood-Retinal Barrier With Human Embryonic Stem Cell-Derived
424 Retinal Pigment Epithelium: Transcriptome and Functional Analysis. *STEM CELLS Translational Medicine*.
425 2013;2(7):534-544. doi:10.5966/sctm.2012-0134

426 24. Radeke MJ, Radeke CM, Shih Y-H, et al. Restoration of mesenchymal retinal pigmented epithelial cells by
427 TGFβ pathway inhibitors: Implications for age-related macular degeneration. *Genome Medicine*. 2015;7(1):58.
428 doi:10.1186/s13073-015-0183-x

429 25. Saini JS, Corneo B, Miller JD, et al. Nicotinamide Ameliorates Disease Phenotypes in a Human iPSC Model of
430 Age-Related Macular Degeneration. *Cell Stem Cell*. 2017;20(5):635-647.e7. doi:10.1016/j.stem.2016.12.015

431 26. Samuel W, Jaworski C, Postnikova OA, et al. Appropriately differentiated ARPE-19 cells regain phenotype and
432 gene expression profiles similar to those of native RPE cells. *Molecular Vision*. 2017;23:60-89.

433 27. Santaguida S, Vasile E, White E, Amon A. Aneuploidy-induced cellular stresses limit autophagic degradation.
434 *Genes & Development*. 2015;29(19):2010-2021. doi:10.1101/gad.269118.115

435 28. Shao Z, Wang H, Zhou X, et al. Spontaneous generation of a novel foetal human retinal pigment epithelium
436 (RPE) cell line available for investigation on phagocytosis and morphogenesis. *Cell Proliferation*.
437 2017;50(6):e12386. doi:10.1111/cpr.12386

438 29. Shih Y-H, Radeke MJ, Radeke CM, Coffey PJ. Restoration of Mesenchymal RPE by Transcription Factor-
439 Mediated Reprogramming. *Investigative Ophthalmology & Visual Science*. 2017;58(1):430-441.
440 doi:10.1167/iovs.16-20018

441 30. Smith JR, Todd S, Ashander LM, et al. Retinal Pigment Epithelial Cells are a Potential Reservoir for Ebola Virus
442 in the Human Eye. *Translational Vision Science & Technology*. 2017;6(4):12-12. doi:10.1167/tvst.6.4.12

443 31. Stevenson NL, Bergen DJM, Skinner REH, et al. Giantin-knockout models reveal a feedback loop between
444 Golgi function and glycosyltransferase expression. *J Cell Sci*. 2017;130(24):4132-4143. doi:10.1242/jcs.212308

445 32. Tresini M, Warmerdam DO, Kolovos P, et al. The core spliceosome as target and effector of non-canonical ATM
446 signalling. *Nature*. 2015;523(7558):53-58. doi:10.1038/nature14512

447 33. Wheway G, Schmidts M, Mans DA, et al. An siRNA-based functional genomics screen for the identification of
448 regulators of ciliogenesis and ciliopathy genes. *Nature Cell Biology*. 2015;17(8):1074-1087. doi:10.1038/ncb3201

449 34. Au ED, Fernandez-Godino R, Kaczynksi TJ, Sousa ME, Farkas MH. Characterization of lincRNA expression in
450 the human retinal pigment epithelium and differentiated induced pluripotent stem cells. *PLOS ONE*.
451 2017;12(8):e0183939. doi:10.1371/journal.pone.0183939

452 35. Carithers LJ, Ardlie K, Barcus M, et al. A Novel Approach to High-Quality Postmortem Tissue Procurement:
453 The GTEx Project. *Biopreservation and Biobanking*. 2015;13(5):311-319. doi:10.1089/bio.2015.0032

454 36. GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature*. 2017;550(7675):204-213.
455 doi:10.1038/nature24277

456 37. Ratnapriya R, Sosina OA, Starostik MR, et al. Retinal transcriptome and eQTL analyses identify genes
457 associated with age-related macular degeneration. *Nature Genetics*. February 2019. doi:10.1038/s41588-019-0351-9

458 38. Collado-Torres L, Nellore A, Kammers K, et al. Reproducible RNA-seq analysis using *Recount2*. *Nature*
459 *Biotechnology*. 2017;35:319-321. doi:10.1038/nbt.3838

460 39. Lachmann A, Torre D, Keenan AB, et al. Massive mining of publicly available RNA-seq data from human and
461 mouse. *Nature Communications*. 2018;9(1):1366. doi:10.1038/s41467-018-03751-6

462 40. Budak G, Dash S, Srivastava R, Lachke SA, Janga SC. Express: A database of transcriptome profiles
463 encompassing known and novel transcripts across multiple development stages in eye tissues. *Experimental Eye*
464 *Research*. 2018;168:57-68. doi:10.1016/j.exer.2018.01.009

465 41. Kakrana A, Yang A, Anand D, et al. iSyTE 2.0: A database for expression-based gene discovery in the eye.
466 *Nucleic Acids Research*. 2018;46(D1):D875-D885. doi:10.1093/nar/gkx837

467 42. Bryan JM, Fufa TD, Bharti K, Brooks BP, Hufnagel RB, McGaughey DM. Identifying core biological processes
468 distinguishing human eye tissues with precise systems-level gene expression analyses and weighted correlation
469 networks. *Human Molecular Genetics*. June 2018. doi:10.1093/hmg/ddy239

470 43. Leinonen R, Sugawara H, Shumway M. The Sequence Read Archive. *Nucleic Acids Research*. 2011;39(Database
471 issue):D19-D21. doi:10.1093/nar/gkq1019

472 44. Zhu Y, Stephens RM, Meltzer PS, Davis SR. SRAdb: Query and use public next-generation sequencing data
473 from within R. *BMC Bioinformatics*. 2013;14:19. doi:10.1186/1471-2105-14-19

474 45. Li H, Handsaker B, Wysoker A, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*
475 *(Oxford, England)*. 2009;25(16):2078-2079. doi:10.1093/bioinformatics/btp352

476 46. Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: The reference human genome annotation for The
477 ENCODE Project. *Genome Research*. 2012;22(9):1760-1774. doi:10.1101/gr.135350.111

478 47. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of
479 transcript expression. *Nature Methods*. 2017;14(4):417-419. doi:10.1038/nmeth.4197

480 48. Soneson C, Matthes KL, Nowicka M, Law CW, Robinson MD. Isoform prefiltering improves performance of
481 count-based methods for analysis of differential transcript usage. *Genome Biology*. 2016;17:12. doi:10.1186/s13059-
482 015-0862-3

483 49. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: Transcript-level estimates improve
484 gene-level inferences. *F1000Research*. 2016;4:1521. doi:10.12688/f1000research.7563.2

485 50. Hicks SC, Okrah K, Paulson JN, Quackenbush J, Irizarry RA, Bravo HC. Smooth quantile normalization.
486 *Biostatistics*. 2018;19(2):185-198. doi:10.1093/biostatistics/kxx028

487 51. Robinson MD, McCarthy DJ, Smyth GK. edgeR: A Bioconductor package for differential expression analysis of
488 digital gene expression data. *Bioinformatics*. 2010;26(1):139-140. doi:10.1093/bioinformatics/btp616

489 52. Maaten L van der. Accelerating t-SNE using Tree-Based Algorithms. *Journal of Machine Learning Research*.
490 2014;15:3221-3245.

491 53. Ritchie ME, Phipson B, Wu D, et al. Limma powers differential expression analyses for RNA-sequencing and
492 microarray studies. *Nucleic Acids Research*. 2015;43(7):e47. doi:10.1093/nar/gkv007

493 54. Law CW, Chen Y, Shi W, Smyth GK. Voom: Precision weights unlock linear model analysis tools for RNA-seq
494 read counts. *Genome Biology*. 2014;15:R29. doi:10.1186/gb-2014-15-2-r29

495 55. Yu G, Wang L-G, Han Y, He Q-Y. clusterProfiler: An R Package for Comparing Biological Themes Among Gene
496 Clusters. *OMICS : a Journal of Integrative Biology*. 2012;16(5):284-287. doi:10.1089/omi.2011.0118

497 56. Köster J, Rahmann S. Snakemakea scalable bioinformatics workflow engine. *Bioinformatics*. 2012;28(19):2520-
498 2522. doi:10.1093/bioinformatics/bts480

499 57. Macosko EZ, Basu A, Satija R, et al. Highly Parallel Genome-wide Expression Profiling of Individual Cells
500 Using Nanoliter Droplets. *Cell*. 2015;161(5):1202-1214. doi:10.1016/j.cell.2015.05.002

501 58. Clark B, Stein-O'Brien G, Shiau F, et al. Comprehensive analysis of retinal development at single cell resolution
502 identifies NFI factors as essential for mitotic exit and specification of late-born cells. *bioRxiv*. July 2018:378950.
503 doi:10.1101/378950

504 59. Wattenberg M, Viégas F, Johnson I. How to Use t-SNE Effectively. *Distill*. 2016;1(10):e2.
505 doi:10.23915/distill.00002

506 60. Eldred KC, Hadyniak SE, Hussey KA, et al. Thyroid hormone signaling specifies cone subtypes in human
507 retinal organoids. *Science*. 2018;362(6411):eaau6348. doi:10.1126/science.aau6348

508 61. Lefebvre JL, Zhang Y, Meister M, Wang X, Sanes JR. γ-Protocadherins regulate neuronal survival but are
509 dispensable for circuit formation in retina. *Development*. 2008;135(24):4141-4151. doi:10.1242/dev.027912

510 62. Lukowski S, Lo C, Sharov A, et al. Generation of human neural retina transcriptome atlas by single cell RNA
511 sequencing. *bioRxiv*. September 2018:425223. doi:10.1101/425223

512 63. Lu Y, Yi W, Wu Q, et al. Single-cell RNA-seq analysis maps the development of human fetal retina. *bioRxiv*.
513 September 2018:423830. doi:10.1101/423830

514 64. Peng Y-R, Shekhar K, Yan W, et al. Molecular Classification and Comparative Taxonomics of Foveal and
515 Peripheral Cells in Primate Retina. *bioRxiv*. September 2018:428110. doi:10.1101/428110

516 65. Philippidou P, Dasen JS. Hox Genes: Choreographers in Neural Development, Architects of Circuit
517 Organization. *Neuron*. 2013;80(1). doi:10.1016/j.neuron.2013.09.020

518 66. Hayashi S, Takeichi M. Emerging roles of protocadherins: From self-avoidance to enhancement of motility. *J*

519 *Cell Sci*. 2015;128(8):1455-1464. doi:10.1242/jcs.166306