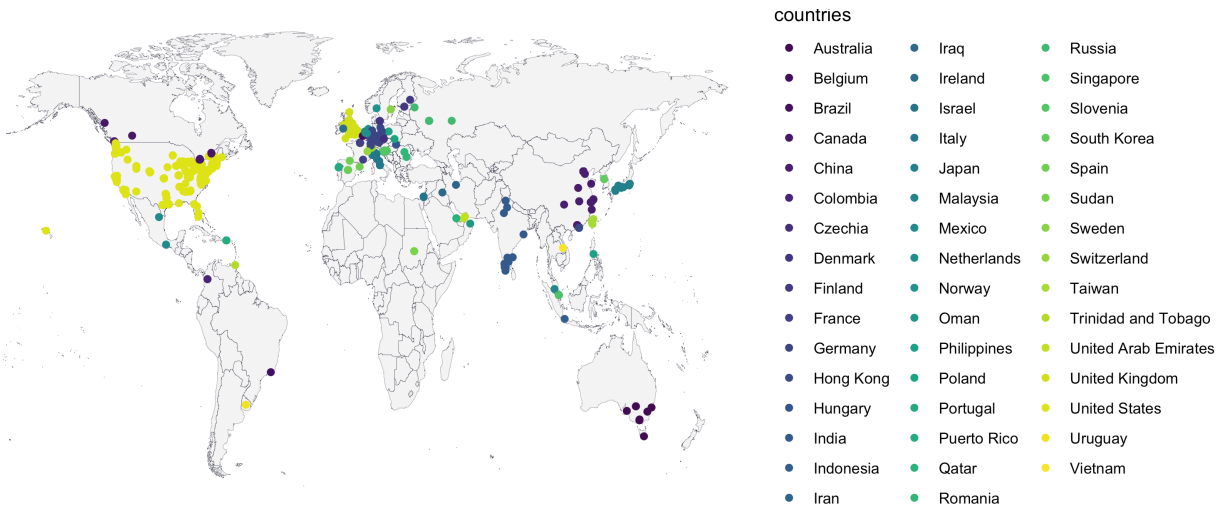
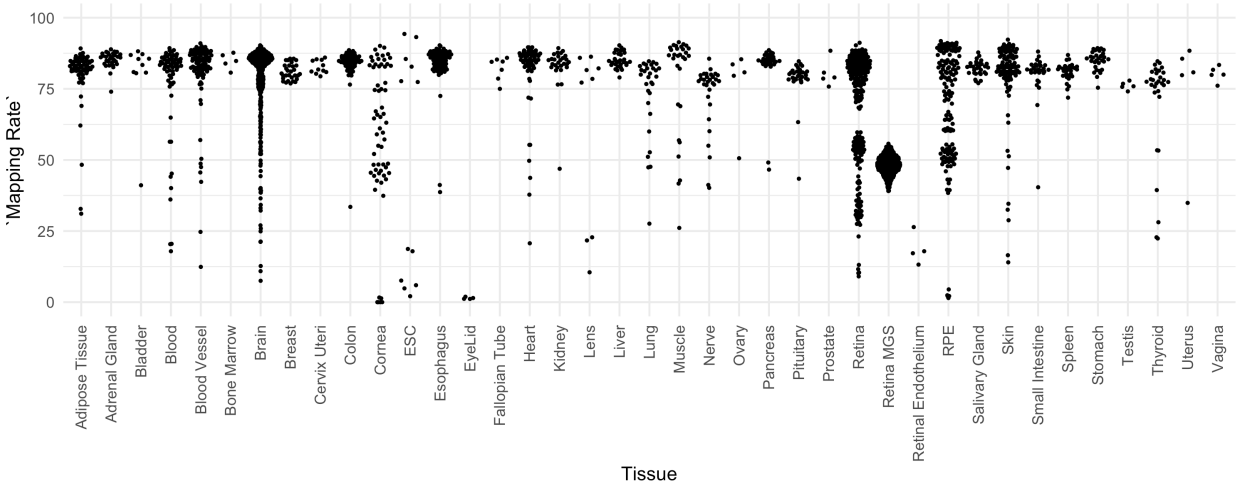
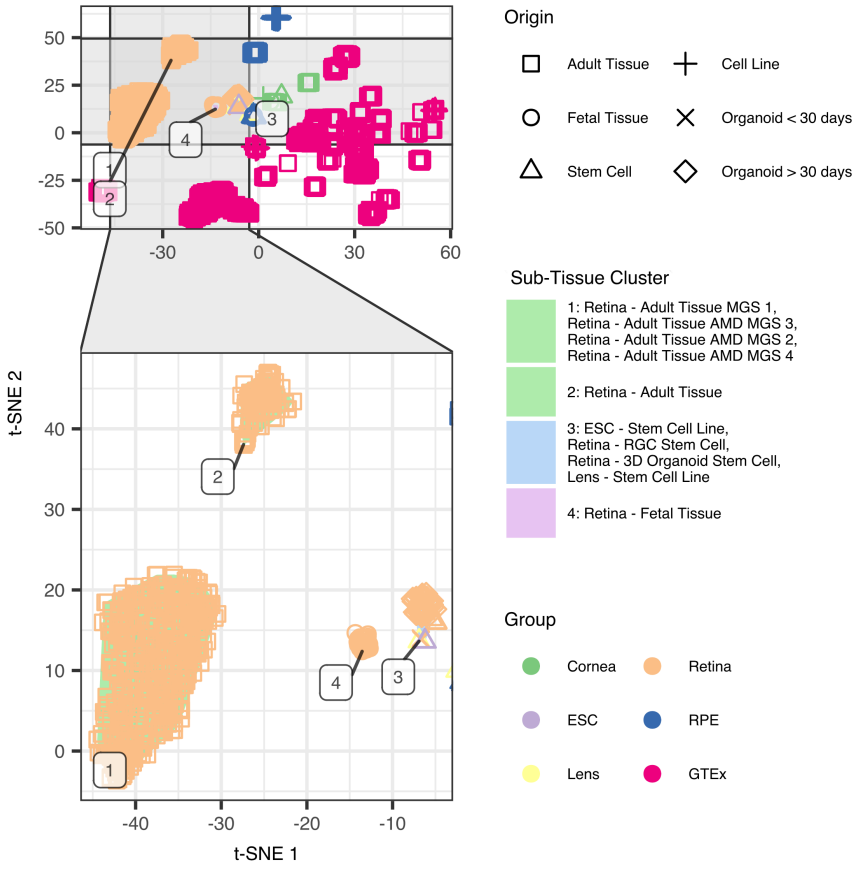
Supplementary Figures and Tables



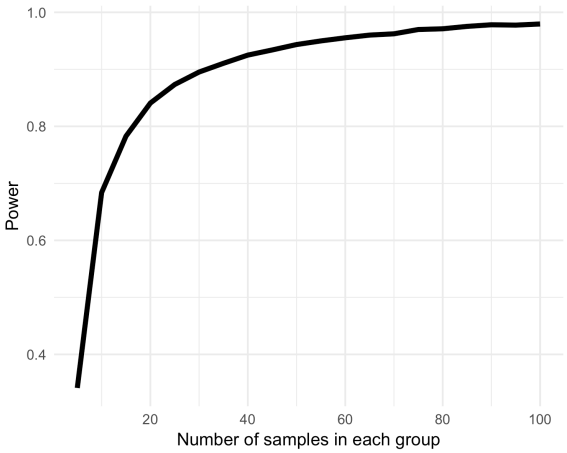
Supplemental Figure 1: As of June 2019, eyeIntegration has had usage across 367 cities and 47 countries.



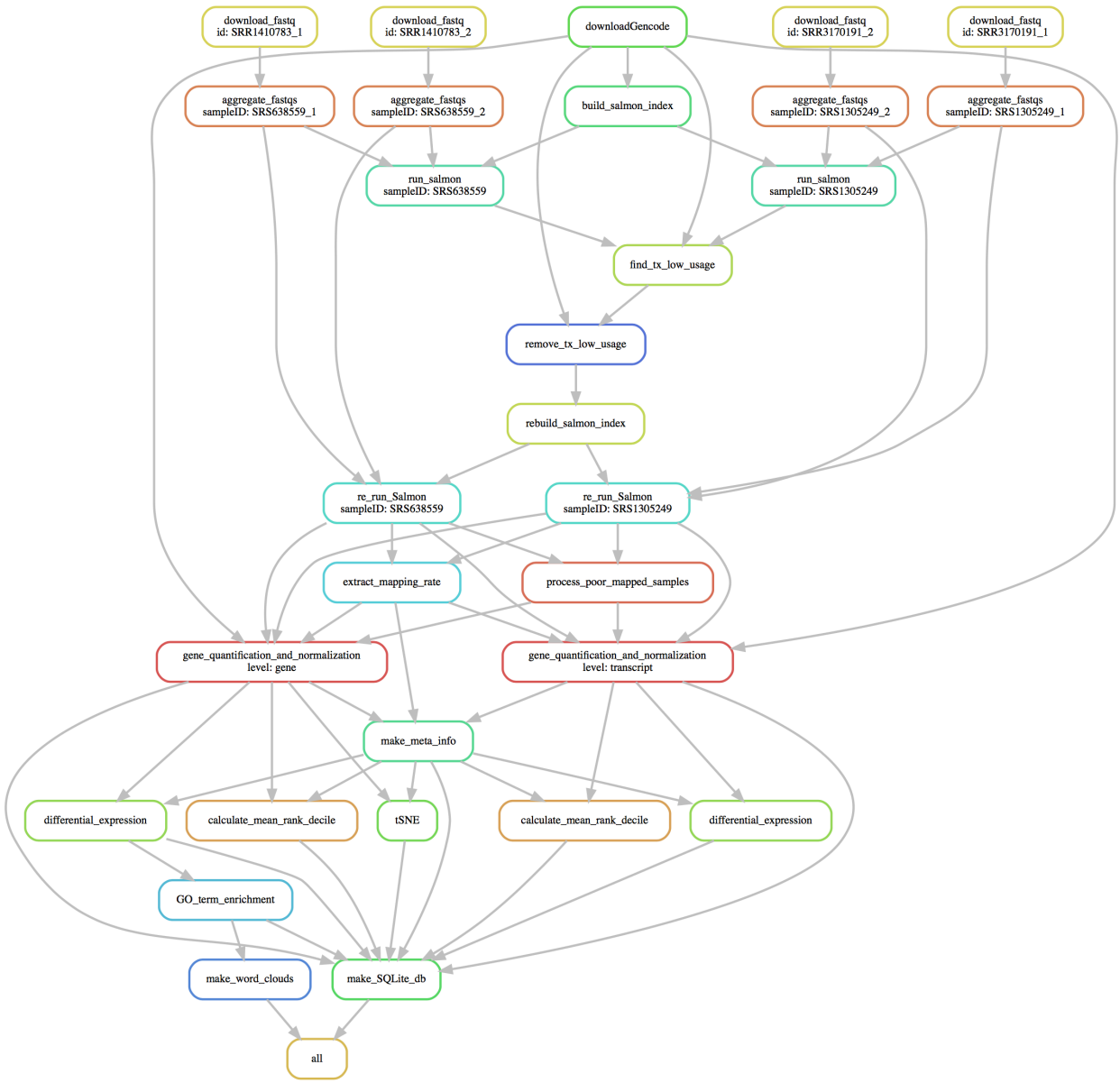
Supplemental Figure 2: Salmon mapping rate for each sample, grouped by tissue type. 1st quartile mapping rate is 51.5%, median is 80.5%, mean is 70.5%, and third quartile is 85.1%.



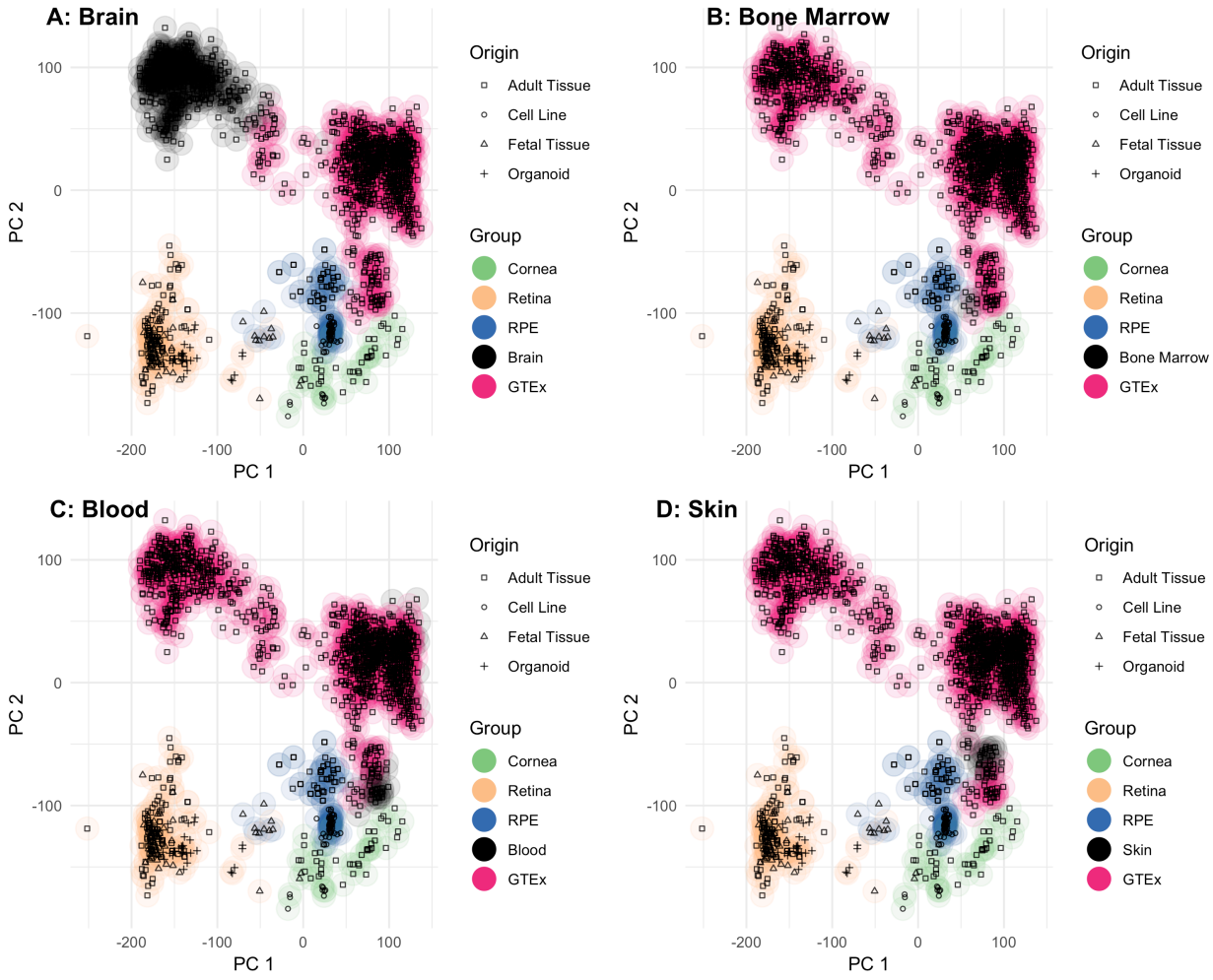
Supplemental Figure 3: Adult retina samples from Ratnapriya et al. (MGS 1 - 4, not-AMD is MGS 1, AMD is MGS 2 through 4) cluster independently from all other adult retina samples collected.



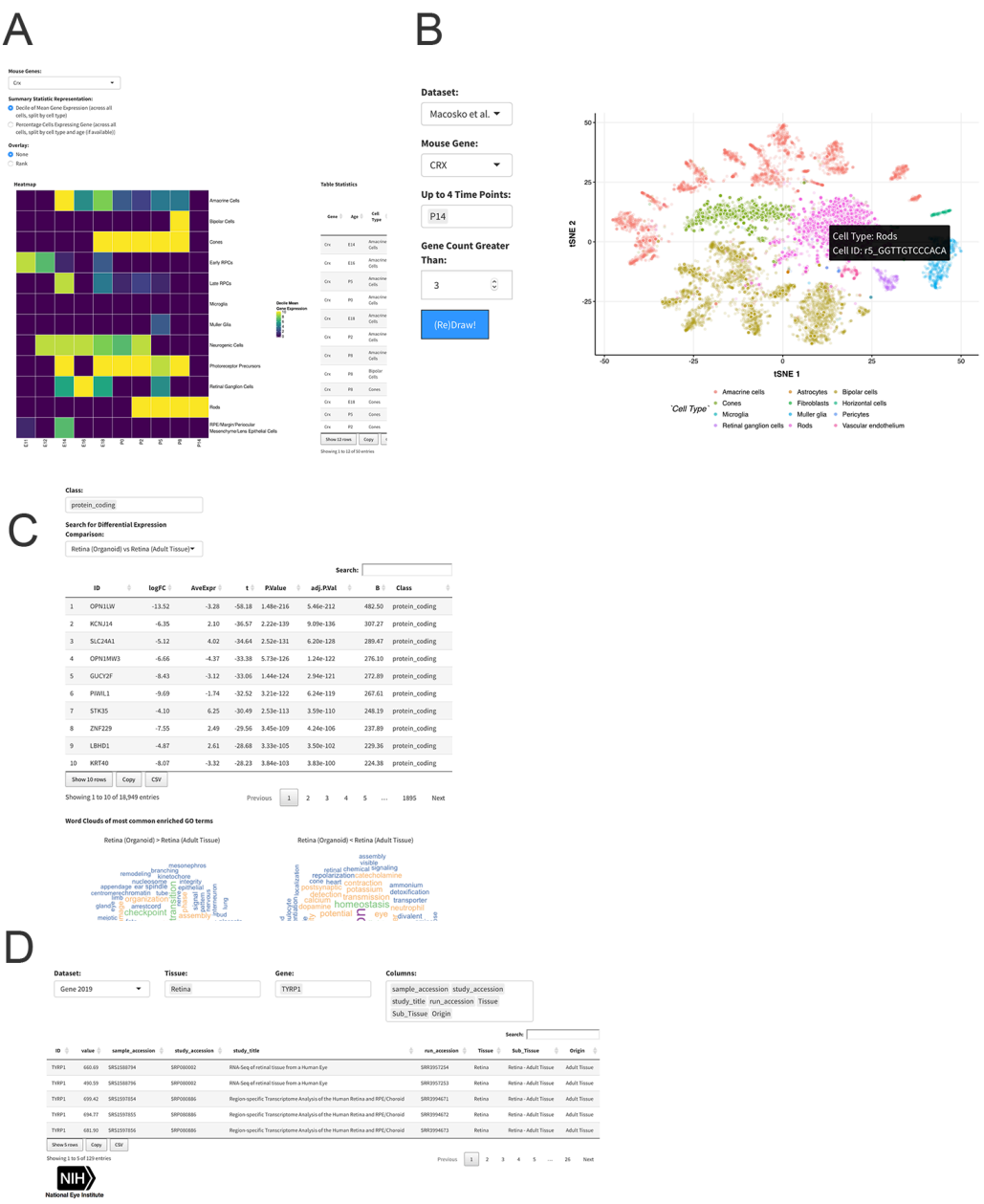
Supplemental Figure 4: Power curve to assess ability to detect >= 1 log2(FoldChange) in gene expression between n samples (x-axis) in each group.



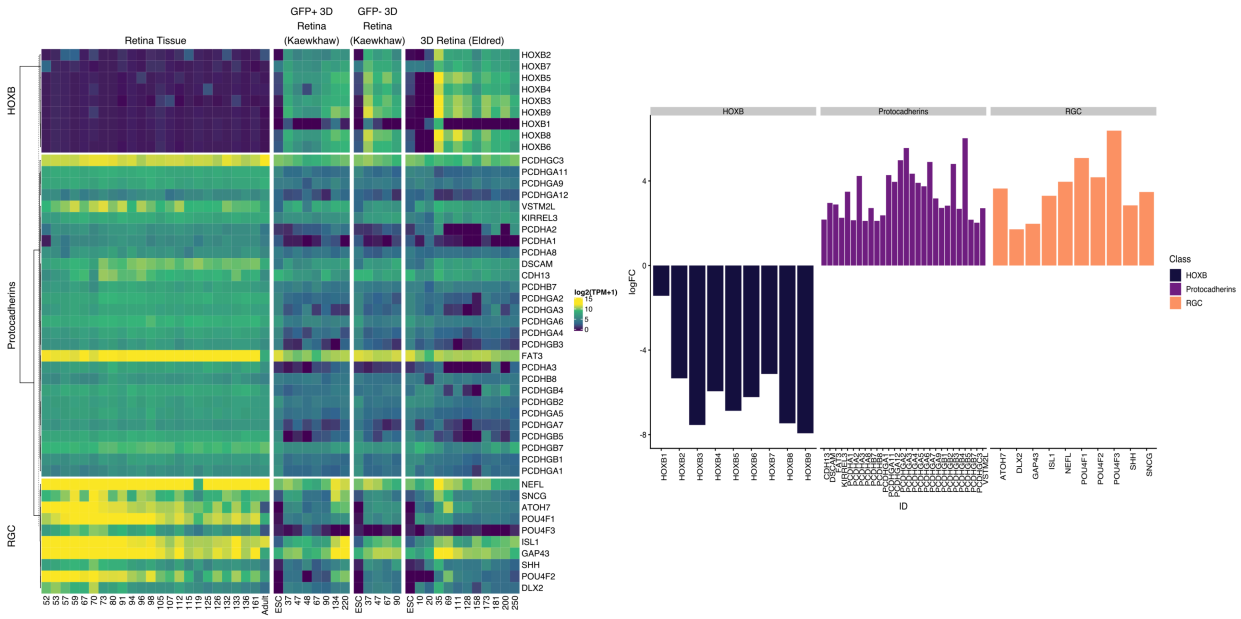
Supplemental Figure 5: Snakemake pipeline to create EiaD 2019 consists of small modular compute sections to ensure sample tracking through the full pipeline



Supplemental Figure 6: PCA plot of all samples suggests that the non-eye tissue most similar to adult retina is the brain, RPE and cornea are most similar to bone marrow, blood, and skin



Supplemental Figure 7: A. CRX mouse retina gene expression heatmap and table information from Clark et al. E11 to P14 scRNA-seq. B. t-SNE visualization of gene expression profiles of individual cells from Macosko et al. C. Data table export view. D. Differential gene expression across different tissues and with GO term enrichment.



Supplemental Figure 8: 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	Version	Pre QC Count	Count	Tissue Types (Count)
GTEx	2017	847	847	Adipose Tissue (39), Adrenal Gland (20), Blood (26), Blood Vessel (57), Brain (255), Breast (20), Colon (40), Esophagus (59), Heart (37), Kidney (17), Liver (20), Lung (20), Muscle (19), Nerve (20), Pancreas (20), Pituitary (20), Salivary Gland (20), Skin (59), Small Intestine (20), Spleen (20), Stomach (19), Thyroid (20)
GTEx	2019	1375	1314	Adipose Tissue (54), Adrenal Gland (30), Bladder (9), Blood (49), Blood Vessel (88), Bone Marrow (5), Brain (372), Breast (30), Cervix Uteri (11), Colon (58), Esophagus (87), Fallopian Tube (7), Heart (58), Kidney (27), Liver (30), Lung (29), Muscle (29), Nerve (30), Ovary (4), Pancreas (30), Pituitary (29), Prostate (5), Salivary Gland (30), Skin (83), Small Intestine (29), Spleen (30), Stomach (30), Testis (5), Thyroid (27), Uterus (4), Vagina (5)

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

Biotype	Gene Count	Transcript Count	Definition
bidirectional_promoter_lncRNA	68	122	A non-coding locus that originates from within the promoter region of a protein-coding gene, with transcription proceeding in the opposite direction on the other strand.
unitary_pseudogene	22	21	A species-specific unprocessed pseudogene without a parent gene, as it has an active orthologue in another species.
retained_intron		6552	Alternatively spliced transcript believed to contain intronic sequence relative to other, coding, variants.
protein_coding	19012	34969	Contains an open reading frame (ORF).
processed_transcript	533	4635	Doesn't contain an ORF.
antisense	4014	5656	Has transcripts that overlap the genomic span (i.e. exon or introns) of a protein-coding locus on the opposite strand.

Biotype	Gene Count	Transcript Count	Definition
pseudogene	11	19	Have homology to proteins but generally suffer from a disrupted coding sequence and an active homologous gene can be found at another locus. Sometimes these entries have an intact coding sequence or an open but truncated ORF, in which case there is other evidence used (for example genomic polyA stretches at the 3' end) to classify them as a pseudogene. Can be further classified as one of the following.
nonsense_mediated_decay		3402	If the coding sequence (following the appropriate reference) of a transcript finishes >50bp from a downstream splice site then it is tagged as NMD. If the variant does not cover the full reference coding sequence then it is annotated as NMD if NMD is unavoidable i.e. no matter what the exon structure of the missing portion is the transcript will be subject to NMD.
IG_C_gene	14	17	Immunoglobulin (Ig) variable chain and T-cell receptor (TcR) genes imported or annotated according to the IMGT.
IG_J_gene	2	2	Immunoglobulin (Ig) variable chain and T-cell receptor (TcR) genes imported or annotated according to the IMGT.
IG_V_gene	127	127	Immunoglobulin (Ig) variable chain and T-cell receptor (TcR) genes imported or annotated according to the IMGT.
TR_C_gene	6	6	Immunoglobulin (Ig) variable chain and T-cell receptor (TcR) genes imported or annotated according to the IMGT.
TR_J_gene	6	6	Immunoglobulin (Ig) variable chain and T-cell receptor (TcR) genes imported or annotated according to the IMGT.
TR_V_gene	41	36	Immunoglobulin (Ig) variable chain and T-cell receptor (TcR) genes imported or annotated according to the IMGT.
IG_C_pseudogene	3	3	Inactivated immunoglobulin gene.
IG_J_pseudogene	1	1	Inactivated immunoglobulin gene.
IG_V_pseudogene	21	18	Inactivated immunoglobulin gene.
TR_V_pseudogene	4	4	Inactivated immunoglobulin gene.
sense_intronic	659	664	Long non-coding transcript in introns of a coding gene that does not overlap any exons.
sense_overlapping	153	180	Long non-coding transcript that contains a coding gene in its intron on the same strand.
lincRNA	4323	5881	Long, intervening noncoding (linc) RNA that can be found in evolutionarily conserved, intergenic regions.
rRNA_pseudogene	24	24	Non-coding RNA predicted to be pseudogene by the Ensembl pipeline
miRNA	130	130	Non-coding RNA predicted using sequences from Rfam and miRBase
misc_RNA	260	260	Non-coding RNA predicted using sequences from Rfam and miRBase
Mt_rRNA	2	2	Non-coding RNA predicted using sequences from Rfam and miRBase
Mt_tRNA	22	22	Non-coding RNA predicted using sequences from Rfam and miRBase
ribozyme	2	2	Non-coding RNA predicted using sequences from Rfam and miRBase
rRNA	5	5	Non-coding RNA predicted using sequences from Rfam and miRBase
scaRNA	15	15	Non-coding RNA predicted using sequences from Rfam and miRBase
scRNA	1	1	Non-coding RNA predicted using sequences from Rfam and miRBase
snoRNA	238	242	Non-coding RNA predicted using sequences from Rfam and miRBase
snRNA	104	106	Non-coding RNA predicted using sequences from Rfam and miRBase
polymorphic_pseudogene	19	22	Pseudogene owing to a SNP/DIP but in other individuals/haplotypes/strains the gene is translated.
unprocessed_pseudogene	593	585	Pseudogene that can contain introns since produced by gene duplication.
translated_processed_pseudogene	2	2	Pseudogene that has mass spec data suggesting that it is also translated.
processed_pseudogene	2537	2435	Pseudogene that lack introns and is thought to arise from reverse transcription of mRNA followed by reinsertion of DNA into the genome.
transcribed_processed_pseudogene	338	169	Pseudogene where protein homology or genomic structure indicates a pseudogene, but the presence of locus-specific transcripts indicates expression.
transcribed_unitary_pseudogene	101	38	Pseudogene where protein homology or genomic structure indicates a pseudogene, but the presence of locus-specific transcripts indicates expression.

Biotype	Gene Count	Transcript Count	Definition
transcribed_unprocessed_pseudogene	680	327	Pseudogene where protein homology or genomic structure indicates a pseudogene, but the presence of locus-specific transcripts indicates expression.
TEC	755	803	To be Experimentally Confirmed. This is used for non-spliced EST clusters that have polyA features. This category has been specifically created for the ENCODE project to highlight regions that could indicate the presence of protein coding genes that require experimental validation, either by 5' RACE or RT-PCR to extend the transcripts, or by confirming expression of the putatively-encoded peptide with specific antibodies.
non_stop_decay		10	Transcript that has polyA features (including signal) without a prior stop codon in the CDS, i.e. a non-genomic polyA tail attached directly to the CDS without 3' UTR. These transcripts are subject to degradation.
3prime_overlapping_ncRNA	25	28	Transcript where ditag and/or published experimental data strongly supports the existence of short non-coding transcripts transcribed from the 3'UTR.
non_coding	2	2	Transcript which is known from the literature to not be protein coding.
macro_lncRNA	1	1	Unspliced lncRNA that is several kb in size.

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

sample accession	Tissue	Sub Tissue	Origin	Age Days	Kept	study title	sample attribute	study abstract	study accession	mapping rate
SRS846894	Cornea	Cornea - Cell Line Endothelium	Cell Line	.	Kept	Transcriptomic analysis of cultured corneal endothelial cells as a validation for their use in cell-replacement therapy	source_name: corneal endothelial cell line || cell type: endothelial || imortalization: telemorase || culture medium: F99	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.	SRP055101	83.700
SRS3493151	Retina	Retina - Adult Tissue AMD MGS 3	Adult Tissue	.	Kept	Integrated analysis of genetic variants regulating retinal transcriptome (GREx) identifies genes underlying age-related macular degeneration	id: GSM3191271: R42015-399pf_382-IR_L6; Homo sapiens; RNA-Seq | sample: SRS3493151 | run: SRR7461134 | patient_number: 382 | r_id: 382_3 | donor: 07-1068 | A69S_rs10490924: G/G | Y402H_rs1061170: C/C | os_od: OD | age: 95 | sex: F | mgs_level: 3 | cause_of_death: acute cardiac event | death_category: Cardiovascular | cataracts: NA | cataract_timing: NA | ocular_history: IOL pseudophakic | compiled_medical_history: arthritis, HTN, high chol | smoking_history: NA | cholesterol: yes | heart_disease: NA | hypertension: yes | postmortem_interval_hrs: 13.77 | rna_isolation_date: 7/6/15 | rna_isolation_batch: isobatch2 | rin: 8 | library_sequenced_date: 1/29/16	Age-related macular degeneration (AMD) is a complex multifactorial disease with at least 34 loci contributing to genetic susceptibility. To gain functional understanding of AMD genetics, we generated transcriptional profiles of retina from 453 individuals including both controls and cases at distinct stages of AMD. We integrated retinal transcriptomes, covering 13,662 protein-coding and 1,462 noncoding genes, with genotypes at over 9 million common single nucleotide polymorphisms (SNPs) for expression quantitative trait loci (eQTL) analysis of a tissue not included in Genotype-Tissue Expression (GTEx) and other large datasets. Cis-eQTL analysis revealed 10,474 genes under genetic regulation, including 4,541 eQTLs detected only in the retina. We then integrated the AMD-genome-wide association studies (GWAS) data with eQTLs and ascertained target genes at six loci. Furthermore, using transcriptome wide association analysis (TWAS), we identified 23 additional AMD-associated genes, including RLBP1, HIC1 and PARP12. Our studies expand the genetic landscape of AMD leading to direct targets for biological evaluation and establish the Genotype-Retina Expression (GREx) database as a resource for post-GWAS interpretation of retina-associated traits including glaucoma and diabetic retinopathy. Overall design: Retinal samples from 523 aged post-mortem human subjects from a spectrum of age-related macular degeneration (AMD) were RNA-seq profiled.	SRP151763	50.100

sample accession	Tissue	Sub Tissue	Origin	Age Days	Kept	study title	sample attribute	study abstract	study accession	mapping rate
SRS1597869	RPE	RPE - Adult Tissue	Adult Tissue	.	Kept	Region-specific Transcriptome Analysis of the Human Retina and RPE/Choroid	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: S16 || submitted sample id: S16 || submitted subject id: 3 || gap_sample_id: 1975628 || 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	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.)	SRP080886	84.000
SRS1747723	ESC	ESC - Stem Cell Line	Stem Cell	.	low salmon mapping rate	Transcriptomes during lens differentiation of human embryonic stem cells	isolate: human embryonic stem cells || age: not applicable || biomaterial_provider: Zhongshan Ophthalmic Centre, 54th 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	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 dayo 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.	SRP091605	17.900
SRS3492989	Retina	Retina - Adult Tissue AMD MGS 4	Adult Tissue	.	Kept	Integrated analysis of genetic variants regulating retinal transcriptome (GREx) identifies genes underlying age-related macular degeneration	id: GSM3191108: R42015-211pf_232-IR_L3; Homo sapiens; RNA-Seq | sample: SRS3492989 | run: SRR7460971 | patient_number: 232 | r_id: 232_4 | donor: 15-0040 | A69S_rs10490924: G/G | Y402H_rs1061170: C/T | os_od: OS | age: 103 | sex: M | mgs_level: 4 | cause_of_death: pneumonia | death_category: Pulmonary | cataracts: NA | cataract_timing: NA | ocular_history: confirmed phakic (OU), AMD | compiled_medical_history: NA | smoking_history: NA | cholesterol: NA | heart_disease: NA | hypertension: NA | postmortem_interval_hrs: 17.68 | rna_isolation_date: 6/24/15 | rna_isolation_batch: isobatch2 | rin: 7.6 | library_sequenced_date: 11/20/15	Age-related macular degeneration (AMD) is a complex multifactorial disease with at least 34 loci contributing to genetic susceptibility. To gain functional understanding of AMD genetics, we generated transcriptional profiles of retina from 453 individuals including both controls and cases at distinct stages of AMD. We integrated retinal transcriptomes, covering 13,662 protein-coding and 1,462 noncoding genes, with genotypes at over 9 million common single nucleotide polymorphisms (SNPs) for expression quantitative trait loci (eQTL) analysis of a tissue not included in Genotype-Tissue Expression (GTEx) and other large datasets. Cis-eQTL analysis revealed 10,474 genes under genetic regulation, including 4,541 eQTLs detected only in the retina. We then integrated the AMD-genome-wide association studies (GWAS) data with eQTLs and ascertained target genes at six loci. Furthermore, using transcriptome wide association analysis (TWAS), we identified 23 additional AMD-associated genes, including RLBP1, HIC1 and PARP12. Our studies expand the genetic landscape of AMD leading to direct targets for biological evaluation and establish the Genotype-Retina Expression (GREx) database as a resource for post-GWAS interpretation of retina-associated traits including glaucoma and diabetic retinopathy. Overall design: Retinal samples from 523 aged post-mortem human subjects from a spectrum of age-related macular degeneration (AMD) were RNA-seq profiled.	SRP151763	48.400
SRS3493218	Retina	Retina - Adult Tissue AMD MGS 2	Adult Tissue	.	Kept	Integrated analysis of genetic variants regulating retinal transcriptome (GREx) identifies genes underlying age-related macular degeneration	id: GSM3191338: R42016-403pf_442-IR_L7; Homo sapiens; RNA-Seq | sample: SRS3493218 | run: SRR7461201 | patient_number: 442 | r_id: 442_2 | donor: 15-1220 | A69S_rs10490924: G/G | Y402H_rs1061170: T/T | os_od: OS | age: 92 | sex: F | mgs_level: 2 | cause_of_death: complete heart block | death_category: Cardiovascular | cataracts: yes | cataract_timing: 2005 & 2011 (unsure which eye first) | ocular_history: confirmed pseudophakic (OU), cataracts (OU) | compiled_medical_history: CABG - 1979, hypertension, hyperlipidemia | smoking_history: NA | cholesterol: yes | heart_disease: NA | hypertension: yes | postmortem_interval_hrs: 18.82 | rna_isolation_date: 8/15/16 | rna_isolation_batch: isobatch2 | rin: 8.4 | library_sequenced_date: 9/23/16	Age-related macular degeneration (AMD) is a complex multifactorial disease with at least 34 loci contributing to genetic susceptibility. To gain functional understanding of AMD genetics, we generated transcriptional profiles of retina from 453 individuals including both controls and cases at distinct stages of AMD. We integrated retinal transcriptomes, covering 13,662 protein-coding and 1,462 noncoding genes, with genotypes at over 9 million common single nucleotide polymorphisms (SNPs) for expression quantitative trait loci (eQTL) analysis of a tissue not included in Genotype-Tissue Expression (GTEx) and other large datasets. Cis-eQTL analysis revealed 10,474 genes under genetic regulation, including 4,541 eQTLs detected only in the retina. We then integrated the AMD-genome-wide association studies (GWAS) data with eQTLs and ascertained target genes at six loci. Furthermore, using transcriptome wide association analysis (TWAS), we identified 23 additional AMD-associated genes, including RLBP1, HIC1 and PARP12. Our studies expand the genetic landscape of AMD leading to direct targets for biological evaluation and establish the Genotype-Retina Expression (GREx) database as a resource for post-GWAS interpretation of retina-associated traits including glaucoma and diabetic retinopathy. Overall design: Retinal samples from 523 aged post-mortem human subjects from a spectrum of age-related macular degeneration (AMD) were RNA-seq profiled.	SRP151763	43.800

sample accession	Tissue	Sub Tissue	Origin	Age Days	Kept	study title	sample attribute	study abstract	study accession	mapping rate
SRS1305277	Cornea	Cornea - Adult Tissue	Adult Tissue	.	Kept	Transcriptome profiling of human keratoconus corneas through RNA sequencing identifies collagen synthesis disruption and downregulation of core elements of TGF-ÂƒÂ., Hippo, and Wnt pathways	source_name: Cornea || disease state: non-KTCN || study: discovery || tissue: cornea	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-ÂƒÂ., 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	SRP070148	46.500
E-MTAB-4377.RNA49	Retina	Retina - Adult Tissue	Adult Tissue	.	Kept	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.	E-MTAB-4377	81.600
SRS3493229	Retina	Retina - Adult Tissue AMD MGS 3	Adult Tissue	.	Kept	Integrated analysis of genetic variants regulating retinal transcriptome (GREx) identifies genes underlying age-related macular degeneration	id: GSM3191346: R42016-110pf_45-IR_L3; Homo sapiens; RNA-Seq | sample: SRS3493229 | run: SRR7461209 | patient_number: 45 | r_id: 45_3 | donor: 14-1029 | A69S_rs10490924: G/G | Y402H_rs1061170: T/T | os_od: OD | age: 90 | sex: M | mgs_level: 3 | cause_of_death: respiratory failure | death_category: Pulmonary | cataracts: NA | cataract_timing: NA | ocular_history: cataracts (OU), confirmed pseudophakic | compiled_medical_history: HTN, Non-Hodgkins lymphoma, prostate cancer, bladder cancer, UTI, light smoker (20yrs) | smoking_history: yes | cholesterol: NA | heart_disease: NA | hypertension: yes | postmortem_interval_hrs: 16.02 | rna_isolation_date: 12/19/14 | rna_isolation_batch: isobatch1 | rin: 8.5 | library_sequenced_date: 3/8/16	Age-related macular degeneration (AMD) is a complex multifactorial disease with at least 34 loci contributing to genetic susceptibility. To gain functional understanding of AMD genetics, we generated transcriptional profiles of retina from 453 individuals including both controls and cases at distinct stages of AMD. We integrated retinal transcriptomes, covering 13,662 protein-coding and 1,462 noncoding genes, with genotypes at over 9 million common single nucleotide polymorphisms (SNPs) for expression quantitative trait loci (eQTL) analysis of a tissue not included in Genotype-Tissue Expression (GTEx) and other large datasets. Cis-eQTL analysis revealed 10,474 genes under genetic regulation, including 4,541 eQTLs detected only in the retina. We then integrated the AMD-genome-wide association studies (GWAS) data with eQTLs and ascertained target genes at six loci. Furthermore, using transcriptome wide association analysis (TWAS), we identified 23 additional AMD-associated genes, including RLBP1, HIC1 and PARP12. Our studies expand the genetic landscape of AMD leading to direct targets for biological evaluation and establish the Genotype-Retina Expression (GREx) database as a resource for post-GWAS interpretation of retina-associated traits including glaucoma and diabetic retinopathy. Overall design: Retinal samples from 523 aged post-mortem human subjects from a spectrum of age-related macular degeneration (AMD) were RNA-seq profiled.	SRP151763	50.500
SRS2582161	Retina	Retina - Fetal Tissue	Fetal Tissue	73	Kept	Molecular anatomy of the developing human retina	D73P		SRP119766	74.500

Supplemental Table 3: Full metadata for 10 random eye samples. Full metadata available as supplementary file "metadata.csv."

Sub Tissue	Cluster	Count
Retina - Adult Tissue	1	107
RPE - Cell Line	2	50
Cells - EBV-transformed lymphocytes	3	30
Pancreas	4	30
Lens - Stem Cell Line	5	2
RPE - Fetal Tissue	5	7
RPE - Stem Cell Line	5	16
Cornea - Adult Tissue	6	25
Cornea - Endothelium	7	16
Cornea - Fetal Endothelium	7	2
Cornea - Stem Cell Endothelium	8	4

Sub Tissue	Cluster	Count
ESC - Stem Cell Line	9	6
Lens - Stem Cell Line	9	2
Retina - 3D Organoid Stem Cell	9	5
Retina - Fetal Eye	9	3
Retina - 3D Organoid Stem Cell	10	47
Retina - RGC Stem Cell	10	3
RPE - Adult Tissue	11	48
Retina - Fetal Tissue	12	35
Brain - Amygdala	13	30
Brain - Anterior cingulate cortex (BA24)	13	27
Brain - Caudate (basal ganglia)	13	28
Brain - Cerebellar Hemisphere	13	1
Brain - Cortex	13	28
Brain - Frontal Cortex (BA9)	13	29
Brain - Hippocampus	13	29
Brain - Hypothalamus	13	28
Brain - Nucleus accumbens (basal ganglia)	13	30
Brain - Putamen (basal ganglia)	13	28
Brain - Spinal cord (cervical c-1)	13	28
Brain - Substantia nigra	13	28
Pituitary	13	1
Lung	14	29
Whole Blood	15	19
Brain - Cerebellar Hemisphere	16	28
Brain - Cerebellum	16	29
Brain - Cortex	16	1
Thyroid	17	27
Heart - Atrial Appendage	18	30
Heart - Left Ventricle	18	28
Skin - Not Sun Exposed (Suprapubic)	19	29
Skin - Sun Exposed (Lower leg)	19	24
Muscle - Skeletal	20	29
Prostate	20	1
Artery - Aorta	21	30

Sub Tissue	Cluster	Count
Artery - Coronary	21	27
Artery - Tibial	21	28
Adipose - Subcutaneous	22	27
Adipose - Visceral (Omentum)	22	27
Artery - Coronary	22	3
Breast - Mammary Tissue	22	8
Esophagus - Gastroesophageal Junction	22	1
Small Intestine - Terminal Ileum	22	1
Nerve - Tibial	23	30
Bladder	24	4
Breast - Mammary Tissue	24	22
Cervix - Endocervix	24	2
Fallopian Tube	24	2
Minor Salivary Gland	24	23
Prostate	24	4
Vagina	24	1
Pituitary	25	28
Kidney - Cortex	26	27
Stomach	27	21
Bladder	28	5
Colon - Sigmoid	28	29
Colon - Transverse	28	29
Esophagus - Gastroesophageal Junction	28	28
Esophagus - Muscularis	28	28
Small Intestine - Terminal Ileum	28	27
Stomach	28	9
Adrenal Gland	29	30
Small Intestine - Terminal Ileum	30	1
Spleen	30	30
Cells - Transformed fibroblasts	31	30
Liver	32	30
Cervix - Ectocervix	33	2
Esophagus - Mucosa	33	30
Minor Salivary Gland	33	7

Sub Tissue	Cluster	Count
Vagina	33	3
Cornea - Cell Line Endothelium	34	9
Cells - Leukemia cell line (CML)	35	5
Cervix - Ectocervix	36	4
Cervix - Endocervix	36	3
Fallopian Tube	36	5
Ovary	36	4
Testis	36	1
Uterus	36	4
Vagina	36	1
Testis	37	4

Supplemental Table 4: Counts of Sub Tissues in each tSNE - dbscan based cluster group

ID	Description	GeneRatio	pvalue	qvalue	geneID
GO:0030198	extracellular matrix organization	51/668	3.11e-17	0.000	VCAM1 FOXC1 P4HA1 LOX NID2 COL1A2 COL27A1 ICAM1 DDR2 WT1 COL14A1 POSTN COL3A1 PDPN HAS2 SMOC2 COL5A1 TGFBI COL6A3 DCN CDH1 TTR SULF1 ECM2 ADAMTS5 CCDC80 ITGB4 ITGA8 COL4A6 FOXF1 DPP4 ITGB6 FN1 COL1A1 ITGB3 SERPINE1 FMOD NPNT MMP19 CREB3L1 A2M BGN COL8A1 TNXB GFAP SPINT1 MMP7 ELF3 THBS1 COL17A1 MYH11
GO:0007156	homophilic cell adhesion via plasma membrane adhesion molecules	28/592	1.54e-12	0.000	PCDHGB5 PCDHGA6 PCDHGA7 PCDHGA3 PCDHGA5 PCDHGA2 PCDHGB3 PCDHGA4 PCDHGA11 PCDHGA9 PCDHGB2 PCDHA3 PCDHGA12 PCDHB7 PCDHGB4 PCDHGA1 PCDHGB1 PCDHGB7 FAT3 DSCAM PCDHA1 PCDHGC3 PCDHB8 VSTM2L KIRREL3 PCDHA2 PCDHA8 CDH13
GO:0050878	regulation of body fluid levels	50/668	6.53e-11	0.000	CYP26B1 FZD6 GUCA1B LYN COL1A2 XDH ADM DOCK6 TFPI2 COL3A1 PDPN HEG1 HAS2 SLC7A11 HK2 EDNRB AQP4 CLDN4 PLAU PLAT ENTPD2 F2RL1 HRG SCNN1A CD9 CLDN1 APOE FOXB1 EHD2 COL1A1 ITGB3 SERPINE1 TRPV4 PRTN3 PRKCG SCNN1B CAV1 KDF1 NFE2 A2M SCNN1G AQP1 GRHL3 ALOX12B TP63 THBS1 AGR2 ANXA8 FLG2 KRT1
GO:0061448	connective tissue development	34/668	1.85e-10	0.000	HOXB3 HOXC4 HOXA5 HOXA3 BMP5 BMPR1B COL27A1 MGP GDF7 WT1 WNT11 ALX1 COL14A1 PLA2G16 PRRX1 MSX2 SNAI2 COL5A1 FOXA1 TGFBI COL6A3 STC1 SULF1 EGR1 MATN2 FRZB OSR2 COL1A1 EFEMP1 PITX1 TRPV4 NKX3-2 ACTA2 CHI3L1

ID	Description	GeneRatio	pvalue	qvalue	geneID
GO:0031589	cell-substrate adhesion	37/668	3.08e-09	0.000	VCAM1 NID2 GBP1 POSTN COL3A1 PDPN HAS2 SMOC2 PTN ALOX15 ACTN3 HOXA7 ECM2 PLAU EPHA1 TNFRSF12A CCDC80 ITGB4 HRG ITGA8 FOXF1 RHOD CCL28 ANGPT1 ITGB6 FN1 COL1A1 ITGB3 SERPINE1 NPNT WNT4 COL8A1 CEACAM6 THBS1 AGR2 APOD COL17A1
GO:0030335	positive regulation of cell migration	45/668	1.04e-08	0.000	HGF LGALS3 IGFBP5 EGF LYN VEGFD ICAM1 DDR2 WNT11 ANXA1 POSTN PDPN HAS2 SEMA3C GPNMB SNAI2 SMOC2 PTP4A1 PLAU EPHA1 FPR2 F2RL1 IL1R1 PTGS2 FOXF1 RHOD FAM110C ANGPT1 EDN3 SRPX2 FOXC2 FN1 COL1A1 ITGB3 SERPINE1 TRPV4 ENPP2 GRB7 ANXA3 CAV1 CEACAM6 S100A14 RAB25 THBS1 NR4A3
GO:0001503	ossification	38/668	2.00e-08	0.000	IFITM1 HGF GPC3 EGR2 IGFBP5 FOXC1 HOXA3 DHRS3 BMP5 BMPR1B MGP DDR2 WNT11 IGF2 TOB1 MSX2 FHL2 GPNMB SNAI2 IGFBP3 FGFR2 PTN ASPN STC1 ALOX15 CEBPB PTGS2 PHOSPHO1 KREMEN2 OSR2 ISG15 COL1A1 NPNT DLX5 ALPL CREB3L1 WNT4 TP63
GO:0050678	regulation of epithelial cell proliferation	32/668	3.05e-07	0.000	GPC3 HOXA5 BMP5 KLF9 SIX1 PAX2 VEGFD XDH HAS2 EGR3 SNAI2 EDNRB FGFR2 PTN SULF1 DLX6 ALDH1A2 SERPINF1 APOE OSR2 ESRP2 ZFP36 ITGB3 DLX5 HTR2B CAV1 KDF1 CCL2 TP63 THBS1 NR4A3 NR4A1
GO:0030099	myeloid cell differentiation	29/592	8.11e-05	0.006	HIST1H4L HIST1H4F POU4F2 HIST1H3F POU4F1 HIST1H3I HIST1H3C HIST1H3B HIST1H3J PRKCB HIST1H4C SIGLEC15 HIST1H4A TREM2 HIST1H4I HIST1H3A LILRB4 FES C1QC HIST1H3G HIST1H4D CSF1R HIST1H4B TYROBP HIST1H4J GPR68 EVI2B CCL3 SPI1
GO:0030099	myeloid cell differentiation	29/668	6.08e-04	0.008	HIST1H4K HOXB8 HOXB7 PIR GPC3 LGALS3 HOXA5 LYN UBD INHBA EPO HOXA7 CEBPB B2M F2RL1 LIF MT1G IL31RA MB ISG15 ZFP36 HOXA9 PRTN3 NFE2 CA2 IRF4 FOS THBS1 NR4A3
GO:0050769	positive regulation of neurogenesis	28/592	1.20e-03	0.043	POU4F2 MYB P2RY12 HEYL E2F1 VWC2L ASCL1 TRIM67 NEFL DSCAM ISLR2 ALKAL2 PAX6 GDF6 CX3CR1 FES OLIG2 STMN2 SHH RIT2 NEUROG3 CUX2 CPNE6 ADRA2B IRX3 DKK1 ITPKA DLX1

Supplemental Table 5: Top GO terms enriched between fetal retina and organoid retina, a er paring of redundant terms with REVIGO[1]

Reference

1. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One*. 2011;6:e21800.