

Single-cell transcriptomic profiling provides insights into retinal endothelial barrier properties

Mark I. Watson,¹ Peter Barabas,¹ Mary McGahon,¹ Megan McMahon,¹ Marc A. Fuchs,² Tim M. Curtis,¹ David A. Simpson¹

¹Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast, Belfast, Northern Ireland; ²Genomics Core Technology Unit, Faculty of Medicine, Health & Life Sciences, Queen's University Belfast, Belfast, Northern Ireland

Purpose: To better characterize retinal endothelial barrier properties through analysis of individual transcriptomes of primary bovine retinal microvascular endothelial cells (RMECs).

Methods: Individual RMECs were captured on the Fluidigm C1 system, cDNA libraries were prepared using a Nextera XT kit, and sequencing was performed on a NextSeq system (Illumina). Data analysis was performed using R packages Scater, SC3, and Seurat, and the browser application Automated Single-cell Analysis Pipeline (ASAP). Alternative splicing events in single cells were quantified with Outrigger. Cytoscape was used for network analyses.

Results: Application of a single-cell RNA sequencing (scRNA-seq) analysis workflow showed that RMECs form a relatively homogeneous population in culture, with the main differences related to proliferation status. Expression of markers from along the arteriovenous tree suggested that most cells originated from capillaries. Average gene expression levels across all cells were used to develop an *in silico* model of the inner blood–retina barrier incorporating junctional proteins not previously reported within the retinal vasculature. Correlation of barrier gene expression among individual cells revealed a subgroup of genes highly correlated with PECAM-1 at the center of the correlation network. Numerous alternative splicing events involving exons within microvascular barrier genes were observed, and in many cases, individual cells expressed one isoform exclusively.

Conclusions: We optimized a workflow for single-cell transcriptomics in primary RMECs. The results provide fundamental insights into the genes involved in formation of the retinal–microvascular barrier.

The inner blood–retina barrier (iBRB) controls fluid exchange across retinal capillary beds. Tight junctions (TJs) and adherens junctions (AJs) between adjacent vascular endothelial cells result in high transendothelial electrical resistance (TEER) and greatly restrict the paracellular flow of water, proteins, lipids, and immune cells into the retina [1]. Breakdown of the iBRB can have serious pathophysiological consequences in the retina and causes the edema that contributes to vision loss and blindness in diabetic retinopathy, retinal vein occlusions, retinopathy of prematurity, and uveitis [2].

The TJ proteins that make up the iBRB include members of the tetraspanin (claudins, occludin, tricellulin, and marvelD3 proteins) and junctional adhesion molecule (JAM) families [3]. These proteins interact with cytosolic scaffolding proteins, such as zonula occludens (ZO-1), afadin (AF6), and cingulin, which, in turn, anchor the TJ complex to the actin cytoskeleton [4]. The main AJ protein of the iBRB is vascular endothelial (VE)-cadherin, which is a member of the

classical cadherin superfamily [3]. VE-cadherin complexes with β -catenin, which links through other proteins to the actin cytoskeleton [5]. However, the specific barrier genes employed by retinal endothelial cells, their interactions with one another, and the heterogeneity in their expression between cells are not well understood. The recent development of single-cell transcriptomic approaches now provides an opportunity to investigate these factors, all of which have consequences for the development of strategies for regulating permeability.

Traditional bulk RNA sequencing or microarray approaches compare average gene expression between populations of cells and are unable to discriminate heterogeneity among individual cells. In contrast, single-cell RNA sequencing (scRNA-seq) technology [6–10] measures the gene expression profiles of individual cells and has enabled characterization of existing and new retinal cell types at the transcriptome level in adult [10] and developing retinas [11]. The cellular resolution provided by scRNA-seq enables co-expression analysis between cells to suggest potential functional interactions between genes.

In many cases, the function of a gene depends not only on its level of expression but also on the specific alternatively spliced variants present, which can give rise to different

Correspondence to: David Simpson, The Wellcome – Wolfson Institute for Experimental Medicine, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, 97 Lisburn Road, Belfast BT9 7BL. Phone: +44-28- 90976470; email: David.Simpson@qub.ac.uk

protein isoforms. Alternative splicing has been recognized as a regulatory force in TJ assembly for almost two decades [12]. Several ZO-1 splice isoforms have been localized to TJs [13], and the ZO-1 alpha(+) variant has been shown to enhance tightness, while the ZO-1 alpha(-) variant is present in dynamic junctions that are readily opened by physiologic signals [14,15]. Occludin 1B's wide epithelial distribution and conservation across species suggest a potentially important role in the structure and function of the TJ [16]. An alternatively spliced occludin isoform generated by skipping exon 4 has been identified [17], and four differentially spliced occludin mRNA transcripts have been reported in human epithelial tissues (the placenta and the colon) and colonic epithelial cells [18]. MarvelD3 colocalizes with occludin at TJs [19] and is expressed as two alternatively spliced isoforms by different types of epithelial as well as endothelial cells. TJ protein claudin-10 exists in two isoforms with alternative exons, 1a and 1b (Cldn10a, Cldn10b) with Cldn10a restricted to the kidney [20]. Four additional claudin-10 splice variants have since been found in mice and humans [21]. Some isoforms insert into the TJ, whereas those lacking exon 4 are retained in the endoplasmic reticulum.

Bulk RNA sequencing has demonstrated the rich diversity of alternatively spliced transcript variants [22,23], but it cannot determine whether cell populations are homogeneous or heterogeneous [24]. This knowledge is important to identify previously unrecognized subpopulations expressing distinct splice isoforms and to determine the extent of differential isoform expression that might change in response to environmental conditions. For example, are there subpopulations of endothelial cells expressing different isoforms of junction proteins, and does differential isoform expression support alternative splicing as a possible mechanism for regulating cell junction properties? Single-cell RNA sequencing has revealed heterogeneity of alternative splicing between the individual cells of many populations [25-30], but it is unclear to what extent all cells in a population express only one transcript variant, one of several variants (bimodality), or simultaneous expression of multiple transcript variants.

Although TJ and AJ transcripts can undergo alternative splicing giving rise to multiple protein isoforms, there have been no studies specifically investigating the splicing patterns of iBRB genes. Single-cell RNA-seq can be used to assess heterogeneity in overall gene expression and the choice of splice isoforms between individual cells [31].

We report the application of scRNA-seq to characterize the transcriptomes of primary bovine retinal microvascular endothelial cells (RMECs). It would be difficult to isolate a pure population of RMECS directly from retinal tissue, and

primary, early passage cultured RMECs provide a widely used model for examining the iBRB that provides the practical utility of working with cultured cells while maintaining as many *in vivo* features of the cells as possible. These data underpin a retina-specific model of endothelial barrier function, including widespread cell-specific patterns of alternative splicing.

METHODS

Isolation and culture of retinal microvascular endothelial cells: RMECs were isolated from bovine retinas using protocols well established in our laboratory that have been demonstrated previously to consistently generate highly pure endothelial cultures with minimal contamination [32]. Bovine eyes were transported from a local abattoir on ice. Ten eyes were used per RMEC isolation. The retinas were removed, washed free of RPE, and homogenized. Microvessel fragments were trapped on an 85 μ m filter. The microvessels were subsequently digested using an enzyme cocktail of Pronase E from *Streptomyces griseus* (Cat No 1074330001, Sigma-Aldrich, Gillingham, UK), DNaseI (Cat No LS002139, Worthington Biochemicals, Lakewood, NJ) and Collagenase from *Clostridium histolyticum* (Cat No C9891, Sigma-Aldrich, Gillingham, UK) and passed through a 53 μ m filter. They were then centrifuged (800 \times g for 10 min) and the pellet resuspended in Dulbecco's modified Eagle's medium (DMEM) containing 0.1 mg/ml Primocin (InvivoGen; San Diego, CA), 1.4 ng/ml insulin (Sigma-Aldrich), 5 μ g/ml heparin (Sigma-Aldrich), and 20% (v/v) porcine serum (Sigma-Aldrich). This mixture was seeded into a 25 cm², 1% gelatin-coated culture flask (TPP, Trasadingen, Switzerland) and maintained at 37 °C in 5% CO₂. The culture medium was changed every 48 h for approximately 5–6 days before the cells were passaged. The RMECs were passaged in a 1:3 ratio at approximately 80% confluency by washing the cells with phosphate buffered saline NaCl (PBSA; 1X; 171 mM NaCl, 3.4 mM KCl, 10.1 mM Na₂HPO₄, 1.8 mM KH₂PO₄, pH 7.4) before TrypLE Express (Gibco, ThermoFisher Scientific, Leicestershire, UK) was added, and the flask was incubated for approximately 2 min to allow the cells to detach. The reaction was stopped with the addition of culture medium. From passage 1 onward, the cells were maintained in DMEM supplemented with 0.1 mg/ml Primocin, 1.4 ng/ml insulin, 5 μ g/ml heparin, and 10% (v/v) porcine serum and used for experiments at passage 3.

Immunohistochemistry: Animal use conformed to the standards of the ARVO Statement for the Use of Animals in Ophthalmic and Vision Research and was authorized by Queen's University of Belfast Animal Welfare and Ethical

Review Body (AWERB). Adult C57BL/6J mice were euthanized with CO₂ asphyxiation and cervical dislocation. The eyes were immediately enucleated, punctured, and fixed in 4% paraformaldehyde (PFA) in PBS (1X; 136.9 mM NaCl, 2.7 mM KCl, 8.1 mM Na₂HPO₄, 1.5 mM KH₂PO₄) for 1–2 h at room temperature. The retinas were dissected out and placed in permeabilization buffer (0.5% Triton X-100 with 5% normal donkey serum in PBS) overnight at 4 °C. Primary antibodies listed in Appendix 1 were selected based on their specificity toward the associated protein and prepared in permeabilization buffer with isolectin B4-biotin (1:200, Sigma-Aldrich). The retinas were incubated with primary antibodies and isolectin B4-biotin at 4 °C for 4 days. Following extensive washing in PBS for 8–9 h, the retinas were incubated with donkey anti-rabbit Alexa 488 and streptavidin Alexa 568 (both 1:200, ThermoFisher, Gillingham, UK) at 4 °C overnight. Nuclei were labeled with the far-red nuclear stain TOPRO3 (1:1,000, ThermoFisher) in PBS. The retinas were rinsed once with PBS and flatmounted on glass slides in Vectashield antifade solution (Vector Laboratories, Peterborough, UK). Images were acquired using a Leica SP5 confocal laser scanning microscope (Leica Geosystems; Heerbrugg, Switzerland; HCX PL APO x63 oil immersion lens). Each fluorescent channel was excited and captured sequentially to minimize bleed-through. Following acquisition, images were digitally segmented using the isolectin-B4 channel to isolate retinal capillaries from the surrounding retinal tissue. All secondary-only controls were negative for staining.

RNA isolation and RT-PCR: RNA was isolated from RMECs at passage 3 using the RNeasy Mini Kit (Qiagen, Manchester, UK) according to the manufacturer's instructions. cDNA synthesis was performed using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, ThermoFisher Scientific). PCR was performed with REDExtract-N-Amp PCR Reaction mix (buffer, salts, dNTPs, Taq polymerase, REDTaq dye, and JumpStart Taq antibody; Merck, Darmstadt, Germany) in conjunction with gene-specific forward and reverse PCR primers (Appendix 2). Agarose gel electrophoresis was conducted at 200 V for 45 min on a 2.5% agarose gel made with 2.5 g agarose powder (Apollo Scientific, Manchester, UK) and 100 ml 1X lithium borate buffer (10 mM lithium acetate, 10 mM boric acid) [33,34]. The GeneRuler low-range DNA ladder (ThermoFisher Scientific) was used to identify band sizes.

Single-cell RNA library preparation: RMECs were grown to confluency at passage 3 and dissociated using TrypLE Express (Gibco, ThermoFisher Scientific). Cell concentration was determined using an Eve Automated Cell Counter (10 µl

cells/10 µl Trypan blue, Nanoentek USA Inc., Loughborough, UK), and the suspension was diluted to 2×10^5 cells/ml so that about 500 cells entered the cell inlet of a C1 Single-Cell Auto Prep integrated fluidic circuit (IFC) for mRNA Seq, 10–17 µm (Fluidigm, San Francisco, CA, Part No. 100–5760). Cell buoyancy was assessed according to the Fluidigm Single-Cell Preparation Guide (PN 100–7697). Cell capture sites were visualized using the Auto Find Fluidigm protocol on an EVOS FL Auto Imaging System (Thermo Fisher). cDNA synthesis and amplification were performed with a Clontech SMART-Seq v4 Ultra Low Input RNA Kit for Fluidigm C1 IFCs. cDNA harvested from the Fluidigm C1 was quantified using the AccuBlue High Sensitivity dsDNA Quantification Kit (Thermo Scientific) in combination with an Echo 525 Acoustic Dispenser (Labcyte, San Jose, CA). The Clontech modified Illumina® Nextera® XT DNA sample preparation protocol was also miniaturized using an Echo 525 Acoustic Dispenser to prepare single-cell mRNA sequencing libraries from cDNA acquired using the C1 system. cDNA was fragmented and tagged with adaptor sequences using a Nextera XT DNA Library Preparation Kit (Illumina). For each sample well (96 samples in a 384 Echo plate), 500 nl of Tagment DNA Buffer, 250 nl of Amplicon Tagment Mix, and 250 nl of sample (at 50 pg) were added to give a reaction volume of 1,000 nl. The plate was then sealed and centrifuged at 1,500 ×g for 30 s before incubation at 55 °C for 5 min and held at 10 °C. Once the sample reached 10 °C, 250 nl of Neutralize Tagment Buffer was added. Tagmented cDNA then underwent limited-cycle PCR amplification to add index adapters and sequences required for cluster formation. To each sample, 1.875 µl of Nextera PCR Master Mix, 1.875 µl of water, and 0.625 µl of i5 (S5XX) and i7 (N7XX) adapters were added to give a total reaction volume of 6 µl. The plate was sealed and centrifuged at 1,500 ×g for 1 min before incubation on a preheated thermocycler; 72 °C 3 min, 95 °C 30 s and then 12 cycles; 98 °C 10 s, 55 °C 30 s, 72 °C 30 min followed by 72 °C 5 min, and held at 10 °C. The libraries were purified using Agencourt AMPure XP Bead Clean-up (1.8X). The libraries were then pooled and normalized, and the molarity was quantified using the High Sensitivity NGS Fragment Analysis Kit on a Fragment Analyzer (Agilent Technologies).

Sequencing: Sequencing was performed by the Genomics Core Technology Unit (GCTU) unit at Queen's University Belfast (Genomics) using the NextSeq® 500 system (Illumina).

Bioinformatic analyses: Fastq files were generated using bcl2fastq2 in the BaseSpace cloud platform (Illumina) and sequencing quality assessed using fastqc. Samples with no cell capture, >1 cell per capture site, or <500 pg/µl cDNA

library were discarded before alignment. Reads were aligned to *Bos taurus* UMD3.1 genome using the Spliced Transcripts Alignment to a Reference (STAR) read mapper [35], and cells with <100,000 aligned reads were discarded. Alternative splicing within single cells was quantified using Outrigger [31], which uses the junction reads from the STAR output to quantify percent spliced in values (PSI / ψ) for each event within each single cell. The Integrated Genomics Viewer (IGV) [36] was used to visualize splicing and create sashimi plots of alternative splicing events. Gene counts were calculated using HTseq [37], and an expression matrix was created. Quality control (QC) was visualized at each stage using multiQC [38]. Seurat [39] is an R package that was used for QC, preprocessing, dimension reduction, clustering, and differential expression of scRNA-seq data. Seurat takes the raw expression matrix as input and was able to mitigate the effects of cell cycle heterogeneity in scRNA-seq data by calculating cell cycle phase scores and regressing them out of the data during preprocessing. The R-based Single-cell Analysis Toolkit for expression in R (Scater) [40] was used to generate violin plots of gene expression. The single-cell latent variable model (scLVM) [41] within the web-based single-cell transcriptomic analytical tool Automated Single Cell Analysis Pipeline (ASAP) [42] was used for normalization of gene expression, accounting for differences in sequencing depth and technical bias. Single-cell consensus clustering (SC3) [43] was used for unsupervised clustering of the filtered data set within ASAP [42]. The ideal number of clusters was determined by maximum silhouette value [43,44], which indicated the stability of each cluster. ASAP was also used for exploration and visualization of the data, specifically expression of genes related to the microvascular barrier.

The BAM files from STAR were analyzed using Monovar [45], a single nucleotide variant (SNV) detection and genotyping algorithm for single-cell sequencing data, to generate variant call files (VCFs). They were processed using VCFtools [46] in a matrix with individual cells in columns and SNVs in rows (labeled 0, 1, or 2 for variant absent, present on one allele, or present on both alleles, respectively). This matrix was used to cluster cells based on their specific SNV fingerprint into groups of cells originating from the same animal. The Cytoscape open source software platform [47] was used in conjunction with the GENEmania [48] and String [49] applications for network analyses and generation of the blood-retina barrier model. Protein functions and interactions were analyzed using the STRING database [50]. CorrelationCalculator was used to calculate and generate heatmaps of pairwise correlations between genes based on their expression pattern among individual cells [51]. Correlation results

were imported into Cytoscape and MetScape used to view correlation networks [51,52].

RESULTS

In this study, we implemented an optimized workflow for scRNA-seq analysis of RMECs (Appendix 3). Two replicate batches of cells were processed, with 49 and 51 single cells, respectively, passing QC and taken through the downstream analysis pipeline. The raw FastQ files and expression matrices are available from Gene Expression Omnibus (GEO) accession number [GSE136800](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE136800).

Population structure of cultured RMECs: To assess the existence of distinct subgroups within RMECs, unsupervised clustering was performed to group the cells based on the similarity of their transcriptomes (i.e., all genes expressed within individual cells). SC3 [23] demonstrated that heterogeneity between RMECs was due principally to cell cycle variation, with proliferating cells forming a small distinct cluster (Appendix 4). All genes differentially expressed between the non-dividing and cycling cells are detailed in Appendix 5.

Use of the Seurat package [39] to regress out cell cycle effects confirmed that the RMECs otherwise formed a relatively homogeneous population (Figure 1). The RMEC cultures were prepared by pooling retinas from five cows; therefore, the genetic origin of each cell could be traced using its characteristic SNV profile. Cells from one individual animal did not form subclusters with similar gene expression (Figure 1, Appendix 4). This result suggests that the genetic background or individual differences in the donor cows' phenotypes did not have a marked effect on an individual cell's gene expression, but more cells would be required to draw any firm conclusions.

Arteriovenous marker genes have been well defined at the single-cell level for brain vasculature [53]. We observed predominant expression of genes highly expressed in brain capillary endothelial cells and limited expression of most arterial or venous markers, although wider expression of *Vegfc* and *Nr2f2* (Figure 2).

Development of a BRB model: Structural junctional proteins: Maintenance of the iBRB is a key function of the retinal microvascular endothelium, and we used the genes expressed in RMECs as a predictor of the main proteins involved in this function. First, a subset of barrier genes listed by the Gene Ontology consortium (GO:0061028 - Establishment of endothelial barrier, GO:0005911 - cell-cell junction and GO:0071603 - endothelial cell-cell junction) that were expressed within the RMECs was created. Known interactions between the proteins encoded by these genes were added

to create the model depicted in Figure 3 (the list of genes in the network is provided in Appendix 6, and the Cytoscape session file is available upon request). Several of the proteins in this model have not been previously shown in RMECs or reported to contribute to the formation of the iBRB. These proteins include desmocollin-2 (*DSC2*), protocadherin 12 (*PCDH12*), fibronectin leucine rich transmembrane protein 3 (*FLRT3*), latrophillin-3 (*ADGRL3/LPHN3*), and cell adhesion molecule-1 (*CADMI*). To validate the model, the presence of previously reported and novel junctional proteins within the retinal microvasculature was confirmed using immunohistochemistry (Figure 3B). It is possible that some of the proteins detected in RMECs are involved in junctions with other cell types, such as pericytes or glial cells *in vivo*.

The availability of expression data from each cell enables the assessment of not only the average expression of a gene, akin to bulk RNA-seq, but also of the distribution of expression values across the population. Analysis of the variation in expression of genes involved in formation of the microvascular barrier between cells showed that some genes are high in all cells and some low in all cells. Others exhibit a pattern suggestive of bimodal expression, low or undetected in one population of cells and highly expressed in another. Examples of potentially bimodally expressed genes include *junctional adhesion molecule 3* (*JAM3*; Gene ID 83700, OMIM 606871) and *Desmocollin 2* (*DSC2*, Gene ID 1824, OMIM 125645; Figure 4). The modality of each gene can be visualized using violin plots of the expression of the gene across all cells with each dot representing a single cell, as shown in Figure 4A. Dots are sized according to the total reads per cell, and their

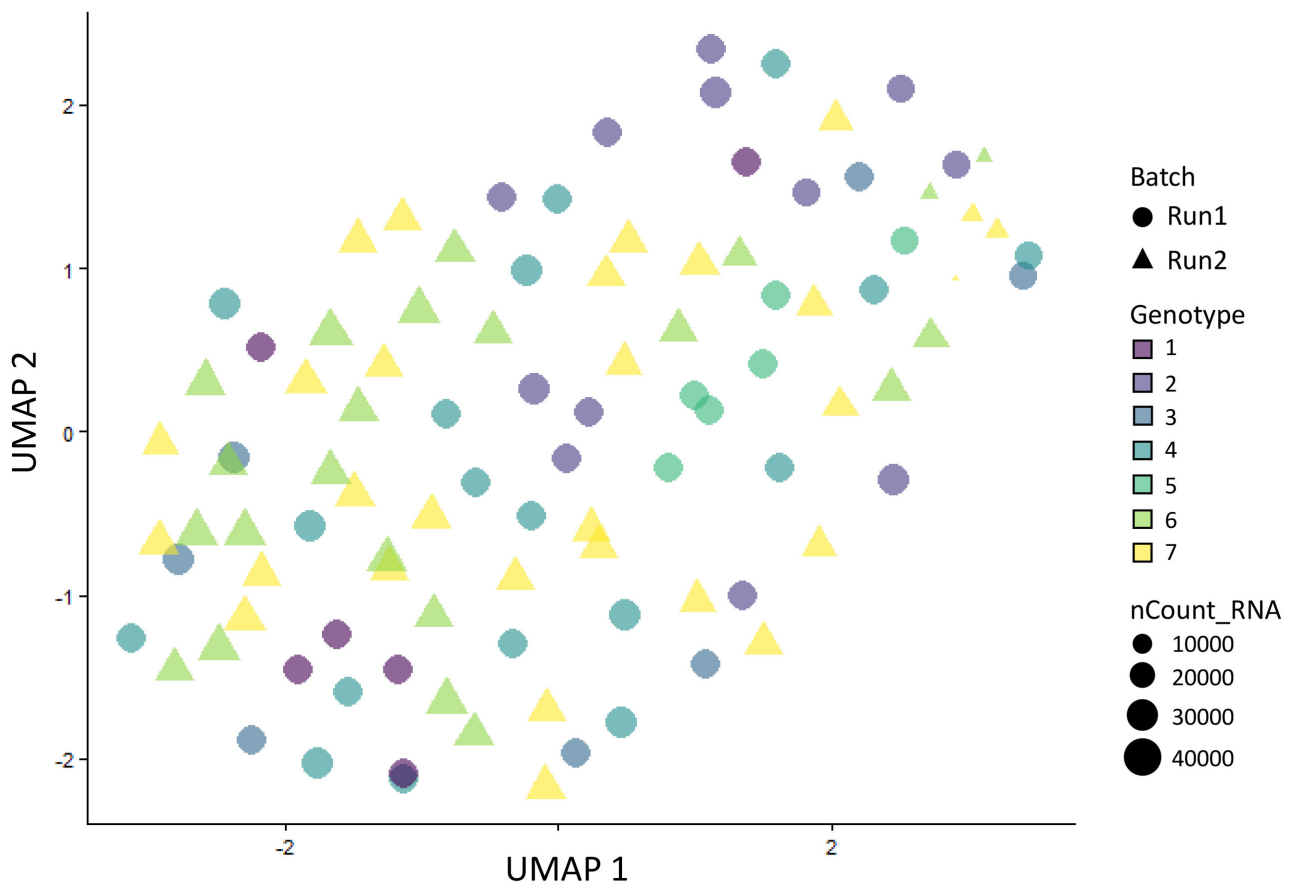


Figure 1. UMAP visualization of the RMEC population according to the individual cell transcriptomes. There are no distinct clusters that would indicate the presence of subpopulations of retinal microvascular endothelial cells (RMECs; following batch correction and regression of cell cycle effects). The scattered distribution of cells with different genotypes suggests that the variability between cells from the same individual is similar to that between cells from different individuals. The random distribution of cells from each batch (shape) and with differing total counts (size) demonstrates that these potential technical effects do not have a significant influence.

even distribution demonstrates that normalization has minimized potential technical artifacts due to the variable read number.

In addition to determining expression modality, which provides insights into how genes are regulated, single-cell resolution enables correlation of the expression patterns of different genes between cells. A subset of barrier genes has highly correlated gene expression across individual RMECs (Figure 4B). These genes can be divided into two subgroups within which expression is highly positively correlated but between which there is negative correlation. The observation that a group of genes are expressed together suggests that they may be involved in the same function, although not necessarily physically interacting. A small group of genes centered on *PECAM-1* (Gene ID 5175, OMIM 173445) and

including *CDH5* (Gene ID 1003, OMIM 601120), *MCAM* (Gene ID 4162, OMIM 155735), *ESAM* (Gene ID 90952, OMIM 614281), and *ADGRL4* (Gene ID 64123, OMIM 616419) stand out for having highly correlated expression. As expected for any subgroup of barrier-related genes, these two sets of coexpressed genes are connected in a network of known protein–protein interactions. Although correlated gene expression is extremely weak evidence for functional interactions, the possibility that these two groups may reflect functional relationships is supported by the clustering of the two sets of proteins in distinct regions of the protein–protein interaction network (Appendix 7).

Cell-specific and novel alternative splicing: In addition to analysis of the gene expression levels between individual cells, scRNA-seq can be used to assess variations

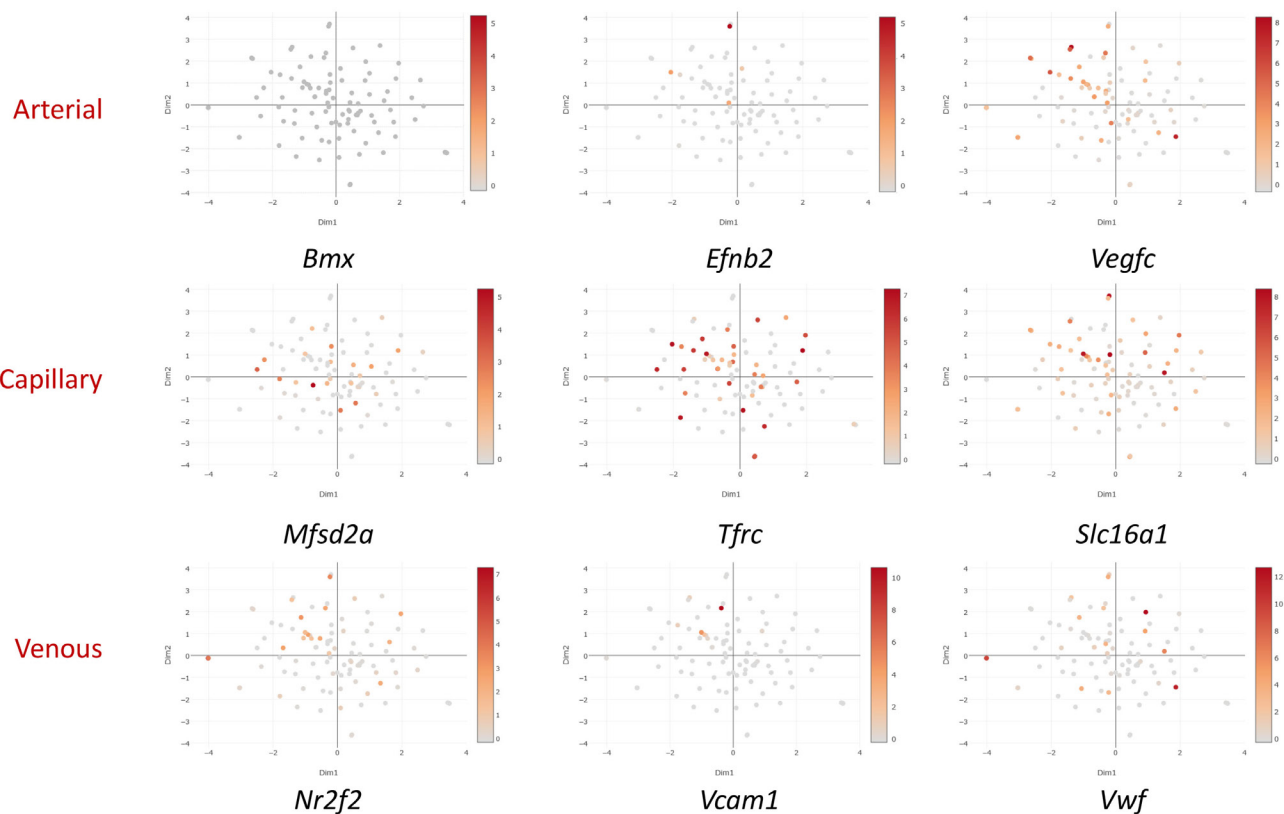


Figure 2. Expression of arteriovenous markers. The expression of arterial, capillary, and venous markers, as defined by single-cell analysis of the brain vasculature [32], is shown on tSNE plots that display the retinal microvascular endothelial cells (RMECs) arranged by overall transcriptome similarity. The aim is to establish whether differences in gene expression between the cells might reflect their origin from different regions of the arteriovenous tree. The majority of cells express genes characteristic of capillaries with very few expressing two of three arterial or venous markers. The discrepant expression of *Vegfc* and *Nr2f2* may be due to species differences, the retinal rather than brain origin of the vessels, or a change in phenotype in culture. The broad expression of most capillary markers, and lack of consistent expression of arterial or venous markers in a subset of cells, suggests that most of the cells originate from capillaries.

in alternative splicing. To achieve this, alternative splicing events within single cells were detected using the Python program Outrigger [31]. The detection of >4,000 skipped exon events and >400 mutually exclusive events emphasizes the diversity of splice isoforms present within RMECs. Although for most events almost all cells have the exon either spliced in (PSI = 1) or spliced out (PSI = 0), there are some for which both isoforms are common, as illustrated by a plot of the global average PSI values for all genes (Figure 5). Differential splicing does not separate the RMECs into obvious clusters, as illustrated with a t-distributed stochastic neighbor embedding (t-SNE) plot and heatmap based on PSI values (Figure 5).

Multiple instances of previously annotated and novel alternative splicing events were observed in the microvascular barrier genes (Appendix 8). For example, the penultimate exon of *adducin 1* (*ADD1*; Gene ID 118, OMIM

102680), which encodes the alpha subunit of a family of cytoskeletal proteins, is skipped in some transcripts that encode a truncated protein with 11 different amino acids at the C-terminus (Figure 6A). This almost certainly has functional consequences due to the loss of phosphorylation sites and the domain involved in interaction with calmodulin [54]. Remarkably, although both isoforms were present in some cells, many expressed one variant exclusively. Similarly, an exon in the *adducin 3* (*ADD3*; Gene ID 120, OMIM 601568) gene was skipped in all transcripts in some cells (i.e., percent spliced in or PSI = 0) and present in all transcripts in other cells (PSI = 1). The functional consequences of the insertion of 32 amino acids toward the C-terminus of the encoded protein due to inclusion of the exon are unknown. Inclusion of an additional exon in transcripts of *ECT2* (Gene ID 1894, OMIM 600586), a guanine nucleotide exchange factor (GEF) that participates in the formation of epithelial TJs, also maintains the reading frame and results in the addition of 31 amino

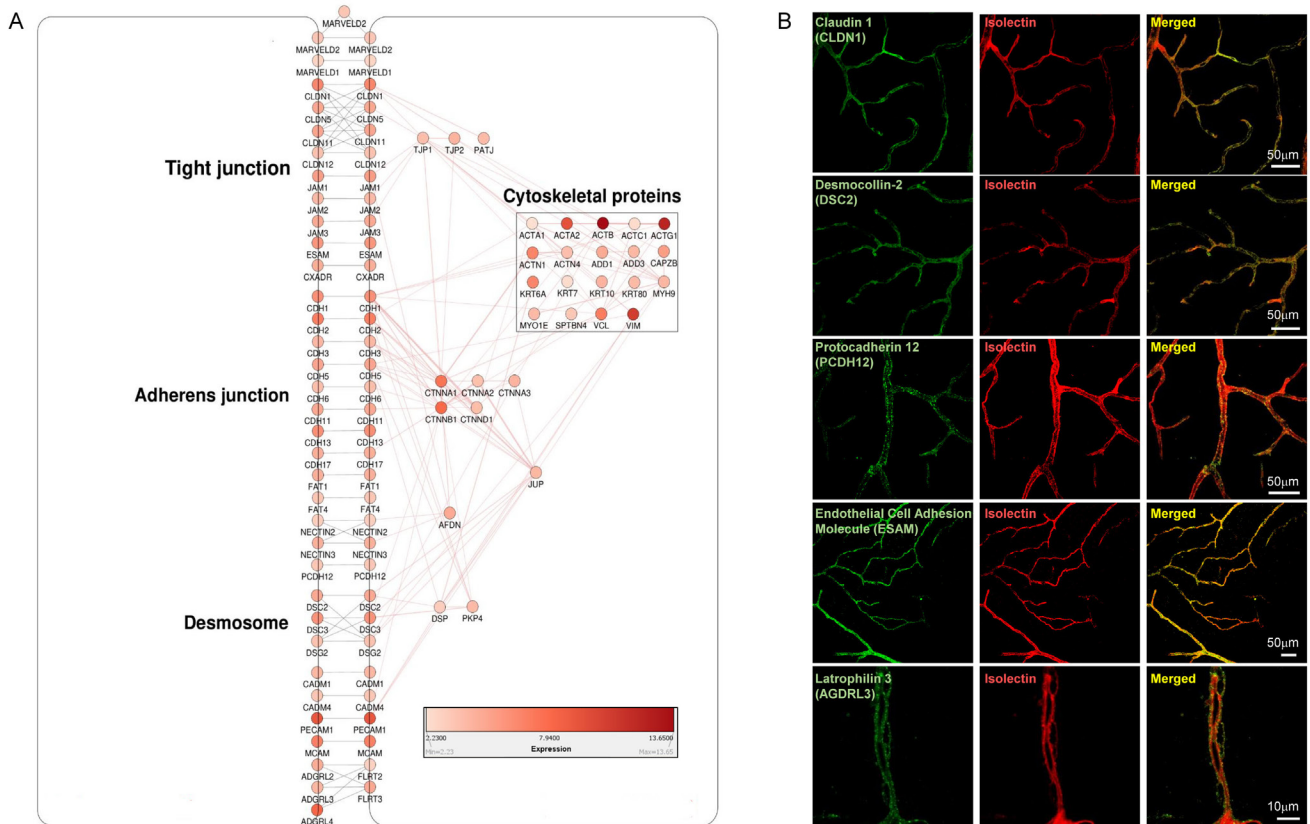


Figure 3. An in silico blood–retina barrier model. **A:** A model of the retinal microvascular barrier was created in Cytoscape using genes expressed in the retinal microvascular endothelial cell (RMEC) data set that were also present within barrier-related Gene Ontology lists. These included junctional proteins not previously reported within the retinal vasculature. GENEmania was used to identify protein–protein interactions between junctional proteins on the plasma membrane and structural proteins within the cytoplasm. A literature search then determined the types of interactions between junctional proteins, e.g., homotypic or heterotypic. **B:** Immunohistochemical staining confirmed the expression within the retinal vasculature of selected junctional proteins from the in silico barrier model.

acids of unknown function (Figure 6C). In addition to events previously annotated in cattle (or mice or humans), many novel events were observed (for many of which further investigation revealed supporting EST evidence). For example, the inclusion of an exon in the *SNAP23* (Gene ID 8773, OMIM 602534) transcript likely has a regulatory function mediated by its extension of the 5'-untranslated region (UTR). Again, the exclusive expression of one *SNAP23* isoform in many cells is supported by multiple reads. Reverse transcription (RT)-PCR with primers spanning the exons of interest validated the existence of the alternative splicing events described above (Figure 6E). The PSI values across all cells for the alternative splicing events detected in the BRB model genes can be visualized as a heatmap (Figure 6F). This shows that although most events are either present or absent in most cells, for a significant minority there are both forms in the same cell.

DISCUSSION

To the best of our knowledge, this is the first investigation of primary retinal endothelial gene expression at the level of individual cells. Although optimized for RMECs, the experimental and bioinformatics approach we developed would be equally applicable to cultures of other retinal cell types. Comparison of individual RMEC transcriptomes demonstrated that they form a relatively homogeneous population in culture, with the main variation related to the proliferative state rather than the genotype. The regressing out of these cell cycle effects enabled more cells to be retained within subsequent analyses. Expression of *in vivo* markers suggests that the RMECs most closely resemble capillary endothelial cells, which may reflect either their origin or convergence on a common phenotype in culture [55].

We exploited the single-cell transcriptome data to better understand the molecular properties of the retinal endothelial cell barrier. We have made the raw data publicly available

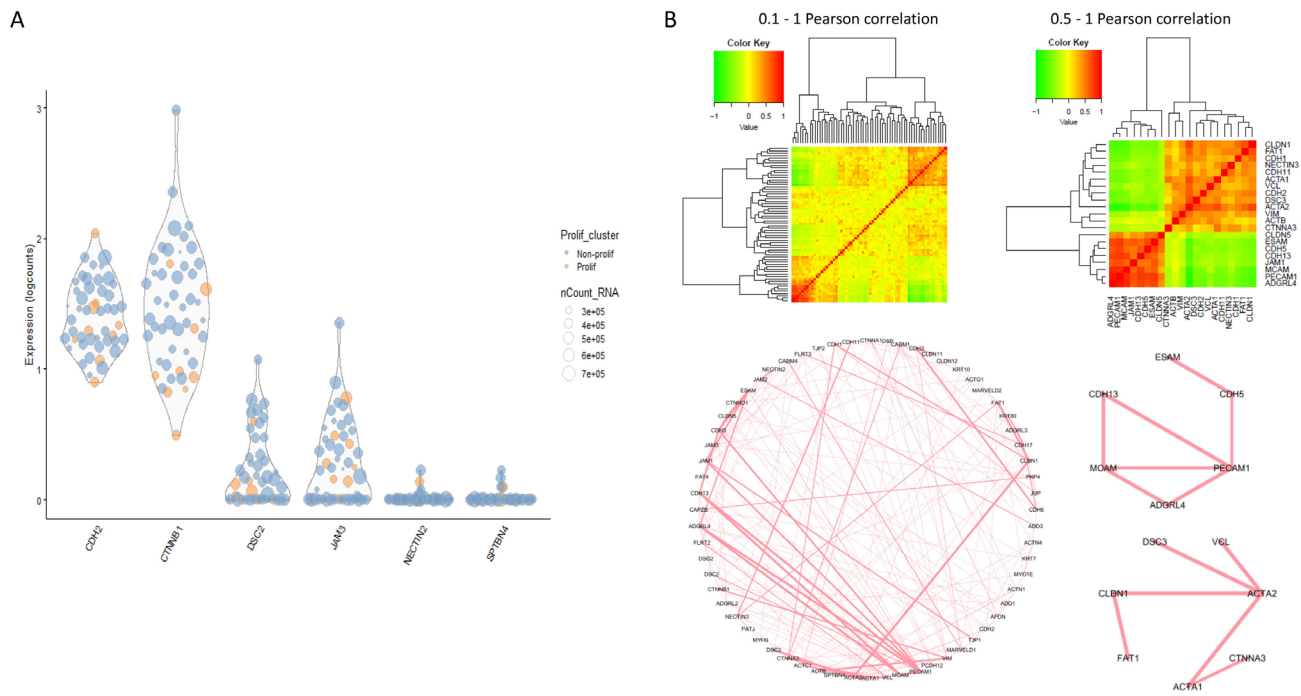


Figure 4. Differential expression of barrier genes between individual cells. **A:** Violin plots show the expression of selected barrier genes in single cells, with the size of the marker indicating total counts in that cell and the color indicating whether it is a proliferating or non-proliferating cell. Some barrier genes are highly expressed in all cells, others low or undetectable in some cells and higher in others (suggestive of bimodal expression), and others low in most cells. **B:** A heatmap of the correlation between the expression of all barrier genes across all individual cells reveals a subset of highly correlated genes that are shown in more detail in the right heatmap of genes with Pearson correlation >0.5. Correlation does not imply that the genes are involved in the same function, but known protein-protein interactions can be used to investigate this possibility. Visualization of interactions as networks, with the thickness of the edges reflecting the weight of the correlation, highlights the position of *PECAM-1* at the core of a small group of highly correlated genes.

(GEO ID [GSE136800](#)) so that they can be used by other researchers to focus on other features of retinal endothelial cell biology, such as cell matrix adhesion, receptor signaling pathways, proliferation, and angiogenesis. The combination of single-cell data with known protein–protein interactions enabled us to develop the most comprehensive *in silico* model of the iBRB published to date. Although the transcriptome data were obtained from a monolayer of RMECs in culture, the predicted expression of key proteins was validated in retinal vessels *in vivo*. In addition to confirming the involvement of known genes, we identified novel genes not previously associated with the iBRB. The model incorporates previously characterized interactions between all of these junctional and structural proteins including links to the cytoskeleton. The scRNA-seq data enabled us to investigate how these barrier genes were coexpressed between individual cells and identify networks of genes with correlated expression. These results suggest that individual cells differentially regulate specific subgroups of genes that influence their barrier properties. Such an insight would not have been possible without the use of a single-cell analysis approach.

Alternative splicing is a critical mechanism for generating diverse protein isoforms with differing functions. For example, the importance of VEGFA splice isoforms in regulating retinal angiogenesis is well-known [56], but analysis of alternative splicing of genes involved in the formation of the iBRB has been limited. To date, alternative splicing has been determined mainly from bulk RNA-seq data, but scRNA-seq provides several important advantages. First, it makes it possible to determine whether splice events occur with similar frequency in all cells or more or less frequently in subpopulations of cells. This makes it possible to address the question of whether two splice variants are present together in all cells or are mutually exclusive. Second, rare splice events that would be dismissed as background in bulk data may be revealed by scRNA-seq to be common in a small number of individual cells. Therefore, we exploited these RMEC scRNA-seq data to provide a comprehensive analysis of splicing within the genes in the iBRB model. In addition to known splice events, we identified numerous novel splicing events in barrier genes that have not been previously reported in the literature. Perhaps surprisingly, many events were detected as the dominant isoform in a subset of cells while being absent

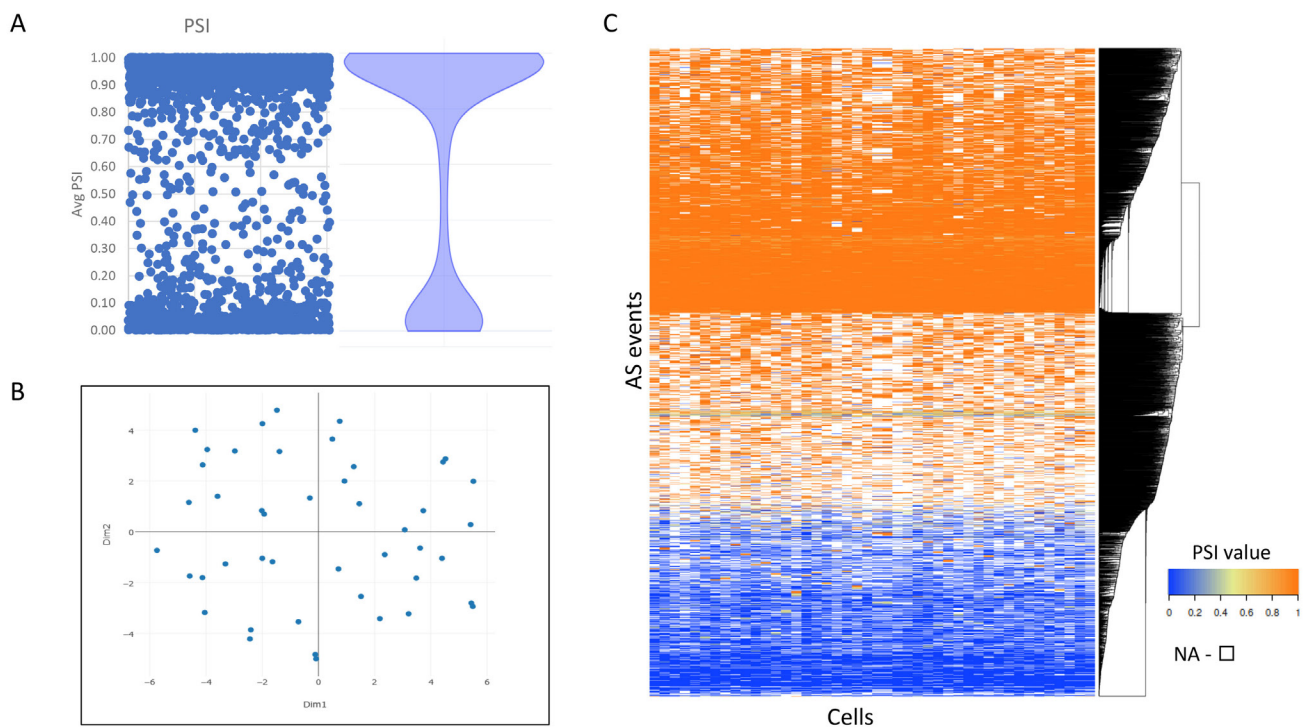


Figure 5. Global splicing events. **A:** The average percent spliced in (PSI) value across all cells is plotted for each alternative splicing event detected in more than ten cells and the distribution represented as a violin plot. This shows that the majority of events are either present or absent in most transcripts. **B:** A t-distributed stochastic neighbor embedding (t-SNE) plot based on PSI values does not reveal any distinct cell clusters. **C:** A heatmap depicting PSI values for each event (columns) across individual cells (rows).

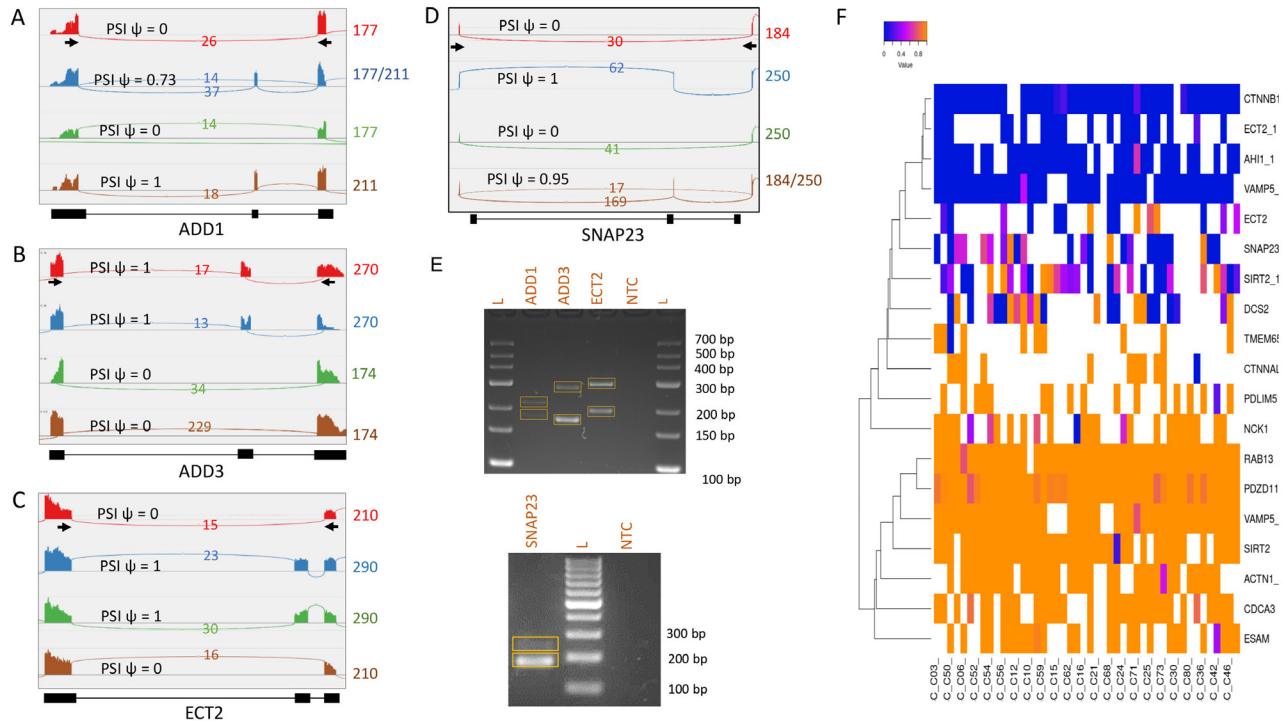


Figure 6. Quantification and validation of single-cell alternative splicing. **A–C:** Sashimi plots created in Integrated Genomics Viewer (IGV) display previously annotated alternative splicing events in the *ADD1* (**A**), *ADD3* (**B**), and *ECT2* (**C**) microvascular barrier genes within four exemplar individual cells. **D:** The alternative splicing detected in the *SNAP23* gene is a novel event not previously reported. Percent spliced in values (PSI ψ) for each event range from 0, indicating that an exon is always spliced out in this cell, to 1, indicating that all reads within this cell include this exon; values in between represent a mix of isoforms within this cell. The splicing events shown are supported by more than ten reads in each cell. The position of primers used for RT–PCR validation are indicated with arrows, and the length of the predicted amplification products in nucleotides is indicated to the right of each plot. **E:** The detected splicing events were validated using RT–PCR with primers flanking the exon of interest. For all genes, the two predicted products were amplified, differing by the length of the alternatively spliced exon. **F:** Heatmap of PSI values for each alternative splicing event identified in the barrier model genes and detected in more than ten cells.

in others, although this is in agreement with previous reports of the dominance of a single transcript variant in each cell [25]. However, we did not detect subpopulations of RMECs based on differential expression of alternative splice isoforms. Nonetheless, these data highlight the heterogeneity in splicing of barrier genes at the cell level that must be considered if we wish to fully understand the molecular properties of the iBRB.

In conclusion, we developed and optimized workflows for studying single-cell transcriptomics in primary RMECs. The gene expression and splicing data from these cells have been made publicly available for secondary analysis. This analysis of barrier genes enabled development of an in silico model of the iBRB, including heterogeneity of gene expression and splicing. This work has enabled a better understanding of the molecular basis of the iBRB and provides

an important platform to determine more precisely why this structure breaks down in multiple retinal disorders.

APPENDIX 1. PRIMARY ANTIBODIES USED FOR IMMUNOFLUORESCENCE.

To access the data, click or select the words “[Appendix 1.](#)”

APPENDIX 2. PRIMERS USED FOR RT–PCR.

To access the data, click or select the words “[Appendix 2.](#)”

APPENDIX 3. SINGLE CELL RNA-SEQ WORKFLOW AND ANALYSIS PIPELINE.

To access the data, click or select the words “[Appendix 3.](#)” (A) Cultured Primary Bovine Retinal Endothelial cells (RMECs) were dissociated. Cells were captured on the Fluidigm C1 system and lysed, mRNA reverse transcribed to cDNA then

PCR amplified using the Clontech SmartSeq-V4 protocol. (B) cDNA libraries were prepared using a miniaturised Nextera XT library prep protocol and sequenced on an Illumina NextSeq 500, using a 150 cycle (2×75 bp) NextSeq kit. (C) Bioinformatic analysis was performed using Python and R programs and the web-based automated single cell analysis pipeline ‘ASAP’.

APPENDIX 4. UNSUPERVISED CLUSTERING OF INDIVIDUAL RMECS ACCORDING TO GENE EXPRESSION WITHOUT CELL-CYCLE CORRECTION. (A)

To access the data, click or select the words “[Appendix 4.](#)” Unsupervised clustering indicated two stable clusters of 41 and 8 cells, as designated by a maximum silhouette value of 0.82 (the silhouette value for each cell is a measure of how similar that cell is to cells in its own cluster when compared to cells in other clusters and ranges from –1 to 1; a high silhouette value suggests that the clustering solution is appropriate). These are visualized in a heatmap with each column representing a single cell and rows genes (>10 fold differentially expressed ($p < 0.001$)). The color of origin, determined by genotype, is indicated above each cell. (B) Ball and stick visualization of the network of protein–protein interactions between the genes more highly expressed in the small cell cluster (False discovery rate (FDR) <0.01, Fold change >2). The network is significantly enriched for protein–protein interactions (p value: <1.0e-16) and the enrichment in genes in the Reactome Cell cycle pathway (BTA-1640170; nodes colored red) suggests that these are dividing cells. (C) There are fewer interactions, although significantly more than expected by chance (p value: 6.46e-14), among the genes more highly expressed in the larger cluster. Many are extracellular proteins, as indicated by the annotation of the most highly enriched terms from each gene ontology domain (in addition Circulatory System Development, the second most significantly enriched Biologic process is shown). Examples include collagens (COL1A1 and COL1A2), the matrix modifying enzyme Lysyl oxidase homolog 2 (LOXL2) and Connective tissue growth factor (CTGF), which mediates heparin- and divalent cation-dependent cell adhesion.

APPENDIX 5. GENES DIFFERENTIALLY EXPRESSED BETWEEN PROLIFERATING AND NON-PROLIFERATING CELLS.

To access the data, click or select the words “[Appendix 5.](#)”

APPENDIX 6. LIST OF BARRIER GENES WITH ALTERNATIVE IDS AND EXPRESSION VALUES.

To access the data, click or select the words “[Appendix 6.](#)”

APPENDIX 7.

To access the data, click or select the words “[Appendix 7.](#)” Interaction network between proteins encoded by the barrier genes that exhibit highly correlated intercellular gene expression. The two groups of proteins defined by their negatively correlated gene expression are located in distinct regions of the network.

APPENDIX 8. ALTERNATIVE SPLICING EVENTS WITHIN BARRIER GENES.

To access the data, click or select the words “[Appendix 8.](#)”

ACKNOWLEDGMENTS

Single cell libraries were prepared and sequenced with assistance from the Queens University Belfast Genomics CTU. Thanks to Michael O'Hare for assistance with imaging. This work was supported by a Northern Ireland Department for the Economy (DfE) studentship, the Health & Social Care R&D Division, Northern Ireland (STL/4748/13), the Medical Research Council (MC_PC_15026), the Biotechnology and Biological Sciences Research Council (BB/I026359/1), Fight for Sight project grant 5091/5092 and US-Ireland Partnership 14/US/B3116. Sequence data has been deposited in the Gene Expression Omnibus (GEO) repository, accession number GSE136800.

REFERENCES

- Hosoya K, Tachikawa M. The inner blood-retinal barrier: molecular structure and transport biology. *Adv Exp Med Biol* 2012; 763:85-104. [PMID: 23397620].
- Klaassen I, Van Noorden CJ, Schlingemann RO. Molecular basis of the inner blood-retinal barrier and its breakdown in diabetic macular edema and other pathological conditions. *Prog Retin Eye Res* 2013; 34:19-48. [PMID: 23416119].
- Díaz-Coránguez M, Ramos C, Antonetti DA. The inner blood-retinal barrier: Cellular basis and development. *Vision Res* 2017; 139:123-37. [PMID: 28619516].
- Guillemot L, Paschoud S, Pulimeno P, Foglia A, Citi S. The cytoplasmic plaque of tight junctions: A scaffolding and signalling center. *Biochim Biophys Acta* 2008; xxx:601-13. [PMID: 18339298].
- Giannotta M, Trani M, Dejana E. VE-Cadherin and Endothelial Adherens Junctions: Active Guardians of Vascular Integrity. *Dev Cell* 2013; 26:441-54. [PMID: 24044891].

6. Kulkarni A, Anderson AG, Merullo DP, Konopka G. Beyond bulk: a review of single cell transcriptomics methodologies and applications. *Curr Opin Biotechnol* 2019; 58:129-36. [PMID: 30978643].
7. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, Leonhardt H, Heyn H, Hellmann I, Enard W. Comparative Analysis of Single-Cell RNA Sequencing Methods. *Mol Cell* 2017; 65:631-643.e4. [PMID: 28212749].
8. Zheng GXY, Terry JM, Belgrader P, Ryvkin P, Bent ZW, Wilson R, Ziraldo SB, Wheeler TD, McDermott GP, Zhu J, Gregory MT, Shuga J, Montesclaros L, Underwood JG, Masquelier DA, Nishimura SY, Schnall-Levin M, Wyatt PW, Hindson CM, Bharadwaj R, Wong A, Ness KD, Beppu LW, Deeg HJ, McFarland C, Loeb KR, Valente WJ, Ericson NG, Stevens EA, Radich JP, Mikkelsen TS, Hindson BJ, Bielas JH. Massively parallel digital transcriptional profiling of single cells. *Nat Commun* 2017; 8:14049-[PMID: 28091601].
9. Picelli S. Single-cell RNA-sequencing: The future of genome biology is now. *RNA Biol* 2017; 14:637-50. [PMID: 27442339].
10. Macosko EZ, Basu A, Satija R, Nemesh J, Shekhar K, Goldman M, Tirosh I, Bialas AR, Kamitaki N, Martersteck EM, Trombetta JJ, Weitz DA, Sanes JR, Shalek AK, Regev A, McCarroll SA. Highly parallel genome-wide expression profiling of individual cells using nanoliter droplets. *Cell* 2015; 161:1202-14. [PMID: 26000488].
11. Clark BS, Stein-O'Brien GL, Shiao F, Cannon GH, Davis-Marcisak E, Sherman T, Santiago CP, Hoang TV, Rajaii F, James-Esposito RE, Gronostajski RM, Fertig EJ, Goff LA, Blackshaw S. Single-Cell RNA-Seq Analysis of Retinal Development Identifies NFI Factors as Regulating Mitotic Exit and Late-Born Cell Specification. *Neuron* 2019; 102:1111-1126.e5. [PMID: 31128945].
12. Cummins PM. Occludin: One Protein, Many Forms. *Mol Cell Biol* 2011; 32:242-50. [PMID: 22083955].
13. Willott E, Balda MS, Heintzelman M, Jameson B, Anderson JM. Localization and differential expression of two isoforms of the tight junction protein ZO-1. *Am J Physiol* 1992; 263:262-[PMID: 1590354].
14. Sheth B, Fesenko I, Collin JE, Moran B, Wild AE, Anderson JM, Fleming TP. Tight junction assembly during mouse blastocyst formation is regulated by late expression of ZO-1 α isoform. *Development* 1997; 124:2027-37. [PMID: 9169849].
15. Balda MS, Anderson JM. Two classes of tight junctions are revealed by ZO-1 isoforms. *Am J Physiol* 1993; 264:C918-24. [PMID: 7682777].
16. Muresan Z, Paul DL, Goodenough DA. Occludin 1B, a variant of the tight junction protein occludin. *Mol Biol Cell* 2000; 11:627-34. [PMID: 10679019].
17. Ghassemifar MR, Sheth B, Papenbrock T, Leese HJ, Houghton FD, Fleming TP. Occludin TM4-: An isoform of the tight junction protein present in primates lacking the fourth transmembrane domain. *J Cell Sci* 2002; 115:3171-80. [PMID: 12118072].
18. Mankertz J, Stefan Waller J, Hillenbrand B, Tavalali S, Florian P, Schöneberg T, Fromm M, Dieter Schulzke J. Gene expression of the tight junction protein occludin includes differential splicing and alternative promoter usage. *Biochem Biophys Res Commun* 2002; 298:657-66. [PMID: 12419305].
19. Steed E, Rodrigues NT, Balda MS, Matter K. Identification of MarvelD3 as a tight junction-associated transmembrane protein of the occludin family. *BMC Cell Biol* 2009; 10:95-[PMID: 20028514].
20. Van Itallie CM, Rogan S, Yu A, Vidal LS, Holmes J, Anderson JM. Two splice variants of claudin-10 in the kidney create paracellular pores with different ion selectivities. *Am J Physiol - Ren Physiol* 2006; 291:F1288-99. [PMID: 16804102].
21. Günzel D, Stuver M, Kausalya PJ, Haisch L, Krug SM, Rosenthal R, Meij IC, Hunziker W, Fromm M, Müller D. Claudin-10 exists in six alternatively spliced isoforms that exhibit distinct localization and function. *J Cell Sci* 2009; 122:1507-17. [PMID: 19383724].
22. Wang ET, Sandberg R, Luo S, Khrebukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. Alternative isoform regulation in human tissue transcriptomes. *Nature* 2008; 456:470-6. [PMID: 18978772].
23. Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* 2008; 40:1413-5. [PMID: 18978789].
24. Arzalluz-Luque Á, Conesa A. Single-cell RNAseq for the study of isoforms-how is that possible? *Genome Biol* 2018; 19:110-[PMID: 30097058].
25. Liu W, Zhang X. Single-cell alternative splicing analysis reveals dominance of single transcript variant. *Genomics* 2020; 112:2418-25. [PMID: 31981701].
26. Karlsson K, Linnarsson S. Single-cell mRNA isoform diversity in the mouse brain. *BMC Genomics* 2017; 18:126-[PMID: 28158971].
27. Linker SM, Urban L, Clark SJ, Chhatrivala M, Amatya S, McCarthy DJ, Ebersberger I, Vallier L, Reik W, Stegle O, Bonder MJ. Combined single-cell profiling of expression and DNA methylation reveals splicing regulation and heterogeneity. *Genome Biol* 2019; 20:30-[PMID: 30744673].
28. Shalek AK, Satija R, Adiconis X, Gertner RS, Gaublomme JT, Raychowdhury R, Schwartz S, Yosef N, Malboeuf C, Lu D, Trombetta JJ, Gennert D, Gnirke A, Goren A, Hacohen N, Levin JZ, Park H, Regev A. Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells. *Nature* 2013; 498:236-40. [PMID: 23685454].
29. Marinov GK, Williams BA, McCue K, Schroth GP, Gertz J, Myers RM, Wold BJ. From single-cell to cell-pool transcriptomes: Stochasticity in gene expression and RNA splicing. *Genome Res* 2014; 24:496-510. [PMID: 24299736].
30. Yap K, Makeyev EV. Functional impact of splice isoform diversity in individual cells. *Biochem Soc Trans* 2016; 44:1079-85. [PMID: 27528755].

31. Song Y, Botvinnik OB, Lovci MT, Kakaradov B, Liu P, Xu JL, Yeo GW. Single-Cell Alternative Splicing Analysis with Expedition Reveals Splicing Dynamics during Neuron Differentiation. *Mol Cell* 2017; 67:148-161.e5. [PMID: 28673540].
32. O'Leary C, McGahon MK, Ashraf S, McNaughten J, Friedel T, Cincolà P, Barabas P, Fernandez JA, Stitt AW, McGeown JG, Curtis TM. Involvement of TRPV1 and TRPV4 Channels in Retinal Angiogenesis. *Investig Ophthalmology Vis Sci*. 2019; 60:3297-[PMID: 31369032].
33. Brody JR, Kern SE. History and principles of conductive media for standard DNA electrophoresis. *Anal Biochem* 2004; 333:1-13. [PMID: 15351274].
34. Singhal H, Ren YR, Kern SE. Improved DNA electrophoresis in conditions favoring polyborates and lewis acid complexation. Cordaux R, editor. *PLoS One*. 2010;5:e11318.
35. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 2013; 29:15-21. [PMID: 23104886].
36. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. Integrative genomics viewer. *Nat Biotechnol* 2011; 29:24-6. [PMID: 21221095].
37. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* 2015; 31:166-9. [PMID: 25260700].
38. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 2016; 32:3047-8. [PMID: 27312411].
39. Butler A, Hoffman P, Smibert P, Papalexi E, Satija R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* 2018; 36:411-20. [PMID: 29608179].
40. McCarthy DJ, Campbell KR, Lun ATL, Wills QF. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics* 2017; 33:btw777-[PMID: 28088763].
41. Buettner F, Natarajan KN, Casale FP, Proserpio V, Scialdone A, Theis FJ, Teichmann SA, Marioni JC, Stegle O. Computational analysis of cell-to-cell heterogeneity in single-cell RNA-sequencing data reveals hidden subpopulations of cells. *Nat Biotechnol* 2015; 33:155-60. [PMID: 25599176].
42. Gardeux V, David FPA, Shajkofci A, Schwalie PC, Deplancke B. ASAP: a web-based platform for the analysis and interactive visualization of single-cell RNA-seq data. *Bioinformatics* 2017; [PMID: 28541377].
43. Kiselev VY, Kirschner K, Schaub MT, Andrews T, Yiu A, Chandra T, Natarajan KN, Reik W, Barahona M, Green AR, Hemberg M. SC3: consensus clustering of single-cell RNA-seq data. *Nat Methods* 2017; 14:483-6. [PMID: 28346451].
44. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J Comput Appl Math* 1987; 20:53-65. .
45. Zafar H, Wang Y, Nakhleh L, Navin N, Chen K. Monovar: Single-nucleotide variant detection in single cells. *Nat Methods* 2016; 13:505-7. [PMID: 27088313].
46. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, McVean G, Durbin R. The variant call format and VCFtools. *Bioinformatics* 2011; 27:2156-8. [PMID: 21653522].
47. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: A software Environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; 13:2498-504. [PMID: 14597658].
48. Montojo J, Zuberi K, Rodriguez H, Kazi F, Wright G, Donaldson SL, Morris Q, Bader GD. GeneMANIA cytoscape plugin: Fast gene function predictions on the desktop. *Bioinformatics* 2010; 26:2927-8. [PMID: 20926419].
49. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape StringApp: Network Analysis and Visualization of Proteomics Data. *J Proteome Res* 2019; 18:623-32. [PMID: 30450911].
50. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Von Mering C. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* 2019; 47:D1D607-13. [PMID: 30476243].
51. Basu S, Duren W, Evans CR, Burant CF, Michailidis G, Karnovsky A. Sparse network modeling and metscape-based visualization methods for the analysis of large-scale metabolomics data. *Bioinformatics* 2017; 33:1545-53. [PMID: 28137712].
52. Karnovsky A, Weymouth T, Hull T, Glenn Tarcea V, Scardoni G, Laudanna C, Sartor MA, Stringer KA, Jagadish HV, Burant C, Athey B, Omenn GS. Metscape 2 bioinformatics tool for the analysis and visualization of metabolomics and gene expression data. *Bioinformatics* 2012; 28:373-80. [PMID: 22135418].
53. Vanlandewijck M, He L, Mäe MA, Andrae J, Ando K, Del Gaudio F, Nahar K, Lebouvier T, Laviña B, Gouveia L, Sun Y, Raschperger E, Räsänen M, Zarb Y, Mochizuki N, Keller A, Lendahl U, Betsholtz C. A molecular atlas of cell types and zonation in the brain vasculature. *Nature* 2018; 554:475-80. [PMID: 29443965].
54. Matsuoka Y, Hughes CA, Bennett V. Adducin regulation. Definition of the calmodulin-binding domain and sites of phosphorylation by protein kinases A and C. *J Biol Chem* 1996; 271:25157-66. [PMID: 8810272].
55. Aird WC. Endothelial Cell Heterogeneity. *Cold Spring Harb Perspect Med* 2012; 2:a006429-006429. [PMID: 22315715].

56. Apte RS, Chen DS, Ferrara N. VEGF in Signaling and Disease: Beyond Discovery and Development. *Cell* 2019; 176:1248-64. [PMID: 30849371].

Articles are provided courtesy of Emory University and the Zhongshan Ophthalmic Center, Sun Yat-sen University, P.R. China. The print version of this article was created on 27 November 2020. This reflects all typographical corrections and errata to the article through that date. Details of any changes may be found in the online version of the article.