1 Identifying Biomarkers of Retinal Pigment Epithelial Cell Stem Cell-

2 derived RPE Cell Heterogeneity and Transplantation Efficacy

3

4 Farhad Farjood[1], Justine D. Manos[2], Yue Wang[1], Anne L. Williams[1], Cuiping Zhao[3], Susan

5 Borden[1], Nazia Alam[4], Glen Prusky[4], Sally Temple[1], Jeffrey H. Stern[1,5], Nathan C. Boles[1,5,*]

6

7 [1] Neural Stem Cell Institute, Rensselaer NY 12144.

8 [2] Current address: Abbvie, Cambridge Research Center, 200 Sidney Street, Cambridge,

9 Massachusetts, USA 02139

10 [3] Current address: Regeneron Pharmaceuticals. 81 Columbia Turnpike. Rensselaer, NY 12144

11 [4] Burke Neurological Institute at Weill Cornell Medicine, White Plains, NY 10605

12 [5] Co-senior authors.

13

14 *Correspondence: nathanboles@neuralsci.org

Summary

16   Transplantation of retinal pigment epithelial (RPE) cells holds great promise for patients

17   with retinal degenerative diseases such as age-related macular degeneration. In-depth

18   characterization of RPE cell product identity and critical quality attributes are needed to enhance

19   efficacy and safety of replacement therapy strategies. Here we characterized an adult RPE

20   stem cell-derived (RPESC-RPE) cell product using bulk and single cell RNA sequencing (sc-

21   RNA-seq), assessing functional cell integration *in vitro* into a mature RPE monolayer and *in vivo*

22   efficacy by vision rescue in the Royal College of Surgeons rats. scRNA-seq revealed several

23   distinct subpopulations in the RPESC-RPE product, some with progenitor markers. We

24   identified RPE clusters expressing genes associated with *in vivo* efficacy and increased cell

25   integration capability. Gene expression analysis revealed a lncRNA (TREX) as a predictive

26   marker of *in vivo* efficacy. TREX knockdown decreased cell integration while overexpression

27   increased integration *in vitro* and improved vision rescue in the RCS rats.

Introduction

29   The retinal pigment epithelium (RPE) supports the overlying neural retinal photoreceptor

30   cells that initiate vision. RPE cells provide nourishment, phagocytose photoreceptor cell outer

31   segments, recycle components of the visual cycle, absorb scattered light, regulate ionic

32   homeostasis and contribute to the blood-retinal barrier (Sparrow et al., 2010; Strauss, 2005). In

33   age-related macular degeneration (AMD), RPE cells undergo changes in morphology,

34   proteome, and phagocytic capacity, resulting in dysfunction, cell death and vision loss (Gu et al.,

35   2012; Kopitz et al., 2004; Lin et al., 2011; Vives-Bauza et al., 2008). RPE replacement therapies

36   are under development to treat AMD using embryonic stem cell (ESC)- and induced pluripotent

37   stem cell (iPSC)-derived RPE cells, which have shown promise to rescue vision in animal

38   models and in patients in early clinical trials (da Cruz et al., 2018; Mandai et al., 2017; Sharma

39   et al., 2019). Adult RPE stem cell (RPESC)-derived RPE progeny (RPESC-RPE) are another

40   stem cell-derived source of RPE cells for transplantation to treat AMD (Salero et al., 2012). In

41   our previous work, we showed that subretinal transplantation of a suspension of RPESC-RPE

42   cells in the Royal College of Surgeons (RCS) rat can preserve vision in this model of retinal

43   degeneration (Davis et al., 2017). A clinical trial of RPESC-RPE transplantation for non-

44   exudative (dry) AMD is underway (NCT04627428).

45

46   An important step to characterize the identity of cell products for transplantation is to

47   describe the extent of cellular heterogeneity. RPE cells have been traditionally considered a

48   homogenous cell population composed of one cell type, but accumulating evidence points

49   towards RPE cellular diversity (Voigt et al., 2019; Xu et al., 2021). Differences in RPE cell

50   behavior, appearance, and gene expression suggest cellular diversity in the native RPE layer

51   and in stem cell-derived RPE preparations(Cuomo et al., 2020; Ortolan et al., 2022; Whitmore et

52   al., 2014). Previously, we showed that transplantation of RPESC-RPE cells cultured for

53   approximately 4 weeks after thaw from the master cell bank (MCB) is more effective at vision

54   rescue compared to earlier and later time points of *in vitro* differentiation. The molecular

55   mechanisms underlying improved effectiveness of transplants at the 4-week stage of

56   differentiation, however, were unknown.

57   Here we use single cell and bulk RNA sequencing to investigate the heterogeneity of

58   RPESC-RPE cells and the underlying molecular signatures that confer transplantation efficacy.

59   Our bioinformatic analyses revealed a transcriptomic signature that correlated with successful

60   transplantation. Furthermore, the data revealed a remarkable heterogeneity of RPESC-RPE cell

61   identities, each with a distinct gene expression signature, indicating that multiple, previously

62   unrecognized cell subpopulations are present in the RPESC-RPE product, including a subset of

63   cells with potentially improved chance of integrating into the host RPE after transplantation.

64   Finally, we identified a novel long noncoding RNA (lncRNA) TREX as a biomarker of transplant

65   efficacy in the RCS rat. These findings broaden our understanding of RPE cell product identity

66   and critical quality attributes (CQAs) needed to enhance regenerative approaches to treat RPE

67   dysfunction and vision loss.

68   **Results**

69   **Bulk RNA-seq uncovers a transplant efficacy gene signature**

70   We sought to determine genes associated with efficient RPE cell transplantation

71   by comparing the transcriptome of RPESC-RPE cultures over an eight-week time course,

72   following our earlier work indicating that the 4-week developmental stage of RPESC is more

73   suitable for transplantation than 2-week or 7-8 week cells (Davis et al., 2017). We utilized the

74   same RPE lines from the earlier study with previously established transplantation effects or a

75   transplant status that could be predicted based on the transplant status of neighboring samples

76   in the timeline (lines 228, 229 and 230; Fig. 1A-B). In addition, we isolated RPE from a fourth

77   donor (line 233) for the experiments. During culture of each RPE line, we collected RNA at the

78   2, 3, 4, 5, 7, and 8-week time points for library preparation and sequencing. The data were then

79   mapped using STAR aligner (Dobin et al., 2013) (Table S1). We accounted for cell line to cell

80   line variance by batch correction using combat-seq (Zhang et al., 2020); (note that the RPE line

81   233 at the 4-week time point had much lower counts and that timepoint was discarded from

82   subsequent analysis).

83   We selected the 4000 most variable genes in the whole bulk-RNA-seq dataset and

84   utilized the Singular Value Decomposition (SVD) approach to examine the relationship of

85   samples to each other. Even though previously we utilized weeks in culture as a surrogate

86   measure to separate groups by transplantation efficacy (Davis et al., 2017), the SVD results did

87   not support time in culture as a good variable for grouping samples (Fig. 1C).However, utilizing

88   the known and predicted transplant status of the samples clearly separated them into three

89   groupings based on transplantation status: 2-week non-efficacious cultured RPE cells (W2-NE),

90   Efficacious RPE (EFF-RPE), or Non-Efficacious RPE (NE-RPE) (Fig. 1D). Using this grouping,

91   we proceeded to look for differential gene expression to identify potential biomarkers of

92   transplant efficacy. We used the general linear model approaches in the edgeR and DESeq2

93   packages to identify differentially expressed genes (DEGs), and included genes identified by

94   both approaches for additional confidence. There were 1465 DEGs with approximately 14%

95   being long-noncoding RNAs (lncRNAs) (Fig. 1E, Table S2). We next determined the extent to

96   which these genes were associated with a transplant group based on their maximum expression

97   and found that the majority of coding DEGs were associated with the W2-NE group (Fig. 1F)

98   while the lncRNA DEGs were associated with each group in roughly equal numbers (Fig. 1G).

99   Proportionately, the lncRNAs made up a much larger fraction of the DEGs in both the EFF-RPE

100   and the NE-RPE. This result is consistent with previous studies on heart cells that highlighted

101   the tissue specificity of lncRNAs; suggesting that these genes may distinguish between groups

102   based on transplant efficacy (Lee et al., 2011).

103   To gain insight into the biological processes differing between RPE cultures from the

104   W2-NE, EFF-RPE and NE-RPE groups we performed enrichment analysis using the goseq

105   package (Young et al., 2010) (Table S3). Following the GO enrichment testing, semantic

106   similarity analysis (Sayols 2020) was used to group terms (Fig. 1H). Similarly, we performed

107   enrichment for REACTOME pathways (Fig. 1I; Table S4). Some of the most over-represented

108   GO parent terms and REACTOME pathways were associated with proliferation and

109   developmental processes, supporting the concept that the RPESCs are undergoing

110   developmental processes during cell culture and that the cells reach a specific point of

4

111   intermediate maturation that is ideal for transplantation. However, as demonstrated by our SVD

112   analysis (Fig. 1C,D), time in culture is an imperfect measure of transplant efficacy. One

113   possibility is that as the RPESC cultures develop, multiple subpopulations arise, with one or

114   more subpopulations more effectively conferring transplantation efficacy, and that exactly when

115   those subpopulations arise, and their duration shows variability over time.

116   **Single cell sequencing reveals changes in RPE heterogeneity over time**

117   We next sought to identify if changes in RPE subpopulations influence transplant

118   efficacy. For these experiments, P2 RPESC cells were cultured for 2, 4, and 8 weeks using the

119   same methodology as for the bulk RNA-seq experiments described above. The ICELL8 platform

120   (Takara) was used to isolate single RPE cells and generate libraries for sequencing with

121   Illumina NovaSeq 6000. Data were processed and mapped with the Cogent NGS Analysis

122   Pipeline (Takara) utilizing the STAR aligner (Dobin et al., 2013). The mapped data was then

123   analyzed using the Seurat (v3) package in R and normalization was performed using the

124   SCTransform pipeline (Hafemeister and Satija, 2019). The dataset was analyzed, and 13

125   clusters (decreasing in size from 0 to 12) representing subpopulations of RPE cells (Fig. 2A)

126   were discovered. All clusters contained cells from all time points (Fig. 2B), but the subpopulation

127   composition of the whole RPE population changed with time in culture (Fig. 2C). Subpopulations

128   were identified as RPE cells based on expression of previously identified human RPE cell

129   signature genes (Bennis et al., 2015; Strunnikova et al., 2010); all RPE cell clusters

130   demonstrated expression of the majority (136-163/171, 79-95%) of RPE signature genes,

131   underscoring the RPE identity of the subpopulations (Fig. S1).

132   To identify genes associated with each cluster, we used the FindAllMarkers function

133   from the Seurat package in R (Fig. 2D, Table S6). Each cluster was associated with between 6

134   (cluster 11) and 1967 (cluster 10) genes with a median of 300 genes across clusters (Fig. 2D;

135   Table S5). RPE associated genes such as *RPE65* and *BEST1* demonstrated differential

136   expression across the subpopulations. *BEST1*, which has known variable expression levels

137   between central and peripheral retina (Mullins et al., 2007), has higher expression in cluster

138   2,3,8,9 and *RPE65* showed a similar expression pattern but with increased expression in cluster

139   7 rather than cluster 9. Transthyretin (*TTR*), which is highly expressed in the RPE of rats and

140   primates (Cavallaro et al., 1990; Pfeffer et al., 2004), also demonstrated variable expression,

141   with cluster 3 having the highest expression followed by cluster 8. *CXCL14*, an immune

142   regulator (Lu et al., 2016) with growth factor activity (Augsten et al., 2009), is highly expressed

143   by RPE cells in the macular region (Whitmore et al., 2014) and demonstrated higher expression

5

144   in clusters 7 and 5 suggesting a potential macular phenotype for these clusters. Overall, the

145   results support prior work indicating that the RPE layer contains a more heterogenous

146   population of cells than previously considered (Xu et al., 2021).

Enrichment analysis reveals RPE subpopulation functional specialization

148   GO and REACTOME pathway enrichment with the hypeR package were utilized to

149   uncover high-level functional differences among the RPE subpopulations (Table S6, S7). The

150   number of enriched GO categories per cluster ranged from 18 categories in cluster 4 to 465

151   categories in cluster 10, with a median of 76 enriched categories per cluster. After GO

152   enrichments were calculated, a semantic similarity analysis for the enriched terms was

153   performed to group the terms together and assist in visualizing and summarizing the analysis

154   (Fig. 2E; Table S6). For visualizing the signaling pathway enrichments we utilized the Reactome

155   pathways database (v72), which is arranged into multiple hierarchical trees composed of related

156   pathways. For our analysis, we took the top-level terms in the trees and counted the number of

157   enriched pathways under those terms (Fig. 2F; Table S7). The number of enriched pathways for

158   each cluster ranged from 0 in cluster 9 to 329 in cluster 10 with a median of 12 enriched

159   pathways per cluster.

160   Most clusters showed enrichment for 'homeostatic processes', 'intracellular transport',

161   and 'sensory organ development' GO categories as well as metabolism- and signaling-related

162   REACTOME pathways (Fig. 2E-F; Table S7-S8), as expected given the role of RPE to maintain

163   the retinal microenvironment. Cluster 0 (~20% of cells) demonstrated strong enrichments for

164   metabolic pathways and pathways involved in protein and ion transport. Cluster 1 (~17% of

165   cells) also demonstrated enrichment for pathways associated with metabolism and GO terms

166   associated with metabolism and molecule transport. Cluster 2 (~9% of cells) showed

167   enrichment in a broader number of GO categories ranging from 'homeostatic process', and 'ion

168   transport' to 'growth', 'neurogenesis' and a variety of development-related terms. Cluster 2 is

169   enriched for many REACTOME pathways associated with 'signal transduction', indicating these

170   RPE cells are likely reactive to environmental challenges and may have a particularly active role

171   in regulating the retinal microenvironment. Cluster 3 (~8% of cells) had a similar enrichment

172   profile to cluster 2, however cluster 3 had less enrichment in developmental pathways and an

173   increased enrichment for metabolic and response to stress functions. In contrast, Cluster 4

174   (~8% of cells) enriched for fewer GO terms and pathways and the enriched functions were

175   focused on metabolic activities. Together these 5 clusters make up more than 60% of the RPE

176   cells sequenced; they show a spectrum of functions, with cluster 4 being highly metabolic,

6

177 cluster 2 being highly reactive and clusters 0, 1, and 3 falling between these categories of
178 function.

179   The pathway enrichments in the remaining clusters also revealed specialization in
180 function. Clusters 5 and 7 showed very similar patterns of enrichment in the GO and pathway
181 analysis for functions associated with locomotion and morphogenesis. The enrichment profile of
182 these clusters could be indicative of subpopulations that could successfully integrate into a RPE
183 monolayer. Notably, cell cycle was substantially enriched in Clusters 6 and 10, indicating these
184 may be proliferative subpopulations. Clusters 6 and 10 were also enriched in signal
185 transduction, and immune and cytokine responses. These clusters are similar to cluster 2 in that
186 they have enrichment profiles with a significant developmental component that could be
187 indicative of populations that benefit cell manufacture and transplantation. Cluster 8 was
188 enriched for functions of molecule transport and stress response. Clusters 9 and 11 were not
189 enriched for any functions and both had few marker genes. Cluster 12, which also had a small
190 number of marker genes, did show enrichment for functions associated with metabolic activities
191 and cation transport. Overall, the RPE subpopulations have overlapping, but distinct functional
192 profiles. Based on the enrichment data, several clusters are candidate subpopulations for
193 efficacious transplantation.

194 **Intersection of scRNA-seq and bulk RNA-seq data implicates three clusters that are more**
195 **likely to confer transplantation efficacy.**
196   To gain a better understanding of which subpopulations may play a role in
197 transplantation efficacy, we intersected the DEGs correlating positively or negatively with
198 efficiency (Bulk-Eff) from the bulk RNA-seq data with the single cell RNA-seq data. Two sets of
199 DEGs were determined using the FindAllMarkers function in Seurat: the first set was DEGs
200 across the 2-, 4- and 8-week culture timepoints (Time-SC) (Table S8) and the second set were
201 DEGs across the clusters (Cluster-SC) (Table S5). Forty-five percent of the Bulk-Eff DEGs
202 intersected with at least one of the two gene sets from the scRNA-seq data (Fig. 3A). As
203 expected, the Bulk-Eff DEGs were most abundantly expressed in the single-cell data at the 4-
204 week timepoint (Fig. 3B). Importantly, when we compared the Bulk-Eff DEGs to the Cluster-SC
205 data, we found that the genes associated with efficacy were most abundantly expressed in
206 clusters 2, 6 and 10 (Fig. 3C). As noted in the previous section, clusters 2, 6, and 10 were
207 enriched for a variety of functions related to development, environmental reactivity, and, notable
208 in cluster 10, proliferation. This enrichment profile along with the large number of marker genes

7

209   associated with transplant efficacy make these three subpopulations the lead candidates for

210   playing a key role in successful engraftment and vision rescue after transplantation.

211   Following up on this discovery, we intersected the markers of clusters 2, 6 and 10 with

212   the Bulk-Eff DEGs and performed GO (Fig. 3D; Table S9) and REACTOME pathway (Fig. 3E;

213   Table S10) enrichment analyses on the intersected gene list for each cluster. The GO analysis

214   revealed some striking differences in the intersected genes from the selected subpopulations

215   (Fig. 3D,E) when compared to the overall cluster GO and REACTOME data (Fig. 2E,F). The

216   intersected gene set from Cluster 10 was highly enriched for terms dealing with proliferation and

217   cell organization. The cluster 2 intersected gene set was enriched for several terms dealing with

218   metabolism and homeostatic processes. On the other hand, the intersected genes from cluster

219   6 were more specifically enriched for terms related to development and cell differentiation than

220   the total cluster 6 marker genes. In the Reactome analysis, the cluster 2 intersected gene set

221   was enriched for either pathways related to vision or sensing external stimuli. The cluster 6

222   intersected gene set was enriched for developmental pathways and pathways dealing with the

223   extracellular matrix (ECM), and specifically for pathways interacting with ECM components and

224   for degrading the matrix (Table S9). This cluster 6 enrichment profile may suggest a

225   subpopulation capable of breaking down and integrating into the RPE monolayer. The cluster 10

226   intersected gene set was again highly enriched for proliferation pathways. Overall, this data

227   analysis reveals that three subpopulations of RPE cells, clusters 2, 6 and 10, have gene

228   expression correlating with properties that could contribute to efficacious transplantation.

229   **Cluster 10, but not Cluster 2 RPE subpopulations can integrate into an RPE monolayer**

230   The adult RPESC-RPE product is subretinally injected as a cell suspension to enable

231   the possibility of cells integrating into the existing RPE monolayer *in vivo*. As a surrogate of the

232   cell integration process, we developed an *in vitro* assay (Fig. S2). The assay workflow begins by

233   culturing RPE test cells and labeling these with GFP. Then approximately 12,000 GFP+ test

234   cells are plated onto a pre-existing 8-week-old RPE monolayer grown in 24-well Transwell

235   format; at the 8-week stage, the monolayers are highly polarized and exhibit typical mature RPE

236   cobblestone morphology that serves as an *in vitro* model of the native RPE layer. Seven days

237   after plating the GFP+ labeled test cells on the mature RPE monolayer, the cultures are fixed

238   and imaged by confocal microscopy over a pre-set 20 position grid covering approximately 4%

239   of the Transwell surface. GFP+ cells that integrate into the mature GFP-negative monolayer are

240   then counted and the percentage of integrated cells are calculated.

8

241   We performed this *in vitro* integration assay by plating 4-week-old GFP labelled RPESC-

242   RPE test cultures on mature monolayers. To determine if integrating cells came from a

243   particular subpopulation, we identified markers for our candidate cluster 2 (YEATS2) and cluster

244   10 (EZH2) subpopulations and used confocal immunofluorescence to determine if the integrated

245   cells expressed either marker along with GFP. Our results showed that ~90% of EZH2+ cells

246   present in the original suspension had integrated into the RPE monolayer. Cluster 10 only

247   makes up ~3% of the original 4-week RPESC-RPE population, yet EZH2+ cells comprised 22%

248   of all integrated cells (Fig. 3F). This demonstrates that cluster 10 cells will successfully integrate

249   and establish in a preformed RPE monolayer. None of the integrated RPE cells exhibited

250   YEATS2 staining (Fig. S3), indicating that the cluster 2 subpopulation, which represents ~9% of

251   the original isolate, did not successfully establish within the monolayer. These results support

252   our hypothesis that specific RPE subpopulations contribute to an efficacious transplant.

253   **A long non-coding RNA as a biomarker of efficacious transplantation**

254   After probing for a possible connection between transplantation and changes in the RPE

255   subpopulation composition, we sought to identify a biomarker of efficacious transplantation.

256   Potential efficacy marker candidates were identified by selecting the genes that had a maximum

257   expression of at least 100 counts in EFF-RPE and at least a 2-fold increase in expression over

258   the W2-NE- and NE-RPE groups in our bulk sequencing data. This yielded a total of 36

259   candidate markers (Table S12). The most consistent and differentially expressed candidate

260   gene is a lncRNA, TCONS_00005049, hereafter referred to as **TREX** (**T**ransplanted **R**PE

261   **Ex**pressed). We examined the level of TREX in the bulk-RNA-seq data of RPESC-RPE samples

262   that had available *in vivo* efficacy data in the RCS rat quantified by optokinetic tracking (OKT)

263   measures of visual acuity. There was a striking positive correlation between OKT data indicating

264   improved vision and increased levels of TREX (Fig. 4A). We next verified TREX levels in these

265   samples by performing qPCR, and again EFF-RPE demonstrated higher levels of TREX (Fig.

266   4B) than the non-efficacious groups. These data suggest that a threshold level of TREX

267   expression was associated with effective transplantation and improved vision.

268   We next undertook qPCR for nine RPE lines including clinical grade cultures obtained

269   under GMP conditions at different cell culture times (Fig. 4C). Most of the lines follow a similar

270   trend with a low level at 2 weeks peaking at 4 weeks. Based on the known transplant data

271   outcomes described previously, we were then able to determine a minimal level of TREX

272   expression (compared to control 18s levels) that was associated with a successful transplant.

273   Note that due to the introduction of proprietary changes made to the 18s taqman probe by its

9

274   manufacturer, we also established a threshold using HPRT as an internal control for future use,

275   using RPE samples that had demonstrated *in vivo* efficacy (229 at 7 weeks) versus not

276   efficacious (230 at 8 weeks) (Fig. 4D). We next sought to manipulate the level of TREX. As a

277   starting point, we determined if TREX, was primarily in the cytoplasm or the nucleus. Using two

278   different RPE cell lines cultured for 4 weeks, we performed qPCR for TREX on the cytoplasmic

279   fraction or on isolated nuclei and found a 40%:60% distribution respectively (Fig. 4E). Based on

280   this distribution, we decided to use a gapmer-based approach to knock down TREX levels.

281   Gapmer knockdown relies on RNase H to cleave the targets and can be effective in both the

282   nucleus and cytoplasm (Liang et al., 2017). We designed and tested four gapmers against

283   TREX and achieved over 50% knockdown with one of the gapmers compared to scrambled and

284   untransfected controls (Fig. 4F).

285   To assess the functional effect of TREX levels on RPE transplantation we utilized our *in*

286   *vitro* integration assay. Using the gapmer based approach, we knocked down TREX in RPESC-

287   RPE cells that had been cultured for 4 weeks and then plated the TREX-knockdown cells on the

288   established mature RPE monolayers used for the integration assay. The scrambled control cells

289   had an integration rate of approximately 2% while significantly fewer TREX-knockdown cells

290   were integrated (~0.2%) (Fig. 4G). We next used a lentiviral overexpression system to

291   overexpress TREX in RPESC-RPE cells cultured for 4 weeks and then performed the

292   integration assay. The empty vector control cells had an integration rate of ~3% whereas the

293   TREX overexpressing cells demonstrated a marked increase in integration with ~18% or 6 times

294   the number of cells integrating into the mature monolayer (Fig. 4H). These results indicate that

295   TREX is not only associated with but is also necessary (Fig. 4G) and sufficient (Fig. 4H) to

296   increase engraftment and transplant efficacy, strongly supporting an important role of TREX to

297   mediate RPE cell integration into a mature RPE monolayer.

298   Based on our *in vitro* results, we proceeded to assess the role of TREX in RPE

299   transplant efficacy at vision rescue. As previously observed, RPE line 230 cells cultured for 7

300   weeks were not efficacious after subretinal transplantation in the RCS rat model of retinal

301   degeneration (Fig. 1B). We transfected RPE 230 cells with either an empty control vector or a

302   dox inducible TREX OE virus and cultured them for 7 weeks prior to transplantation. Vision was

303   measured by OKT at 60 days after transplantation (Table S13). Eight rats were transplanted for

304   each condition and five sham (vehicle only) subretinal injections were used as control. The OKT

305   results generally have a bimodal distribution indicative of either a positive effect on vision rescue

306   or a lack of vision rescue. An animal's vision was considered 'rescued' if the OKT measure was

10

307   above 0.4 cycles/degree and considered 'not rescued' if below this threshold. We took a

308   Bayesian approach to the analysis and identified a '**R**egion **O**f **P**ractical **E**quivalence' (ROPE, a

309   range of values representing no effect) based on the highest density interval (HDI, range of

310   values that 95% of the distribution lays under) from the OKT results at p90 of the cumulative

311   sham (vehicle injected) (n=92, Table S13). Using the sham results as a prior probability, we

312   applied a Bernoulli likelihood function to calculate the posterior distribution of the probability of

313   an efficacious transplant of our empty vector control (Fig. 4I) and of TREX OE cells (Fig. 4J).

314   We then sampled from the posterior probabilities to identify the HDI of each condition. In the

315   case of the empty vector control cells, the HDI completely encompassed the ROPE indicating

316   that the control RPE cells performed similarly to the sham as previously observed with the

317   inefficacious 7-week-old RPE 230 cultures. However, in the case of the TREX OE cells, the HDI

318   laid completely to the right of the ROPE, indicating an increase in efficacy associated with RPE

319   cells overexpressing TREX. In addition to our Bayesian approach, we also performed binomial

320   tests for each comparison and found that while the control cells were not significantly different

321   from sham (p=0.09157), the TREX OE cells were different from both sham (p=1.92e-06) and the

322   control cells (p= 0.004227). Hence, there was a positive effect of creating high TREX levels on

323   cell transplant efficacy; we essentially converted a non-efficacious RPE cell preparation to one

324   that effectively rescued vision in this animal model. We next sought to determine if having

325   endogenous high levels of TREX was sufficient to demonstrate transplant efficacy.

326   RPE line 255 has exceptionally high levels of TREX after only 2 weeks of culture (Fig.

327   4C), a time point in RPE culture that is usually ineffective at vision rescue compared with 4

328   weeks of culture prior to transplantation (Davis et al., 2017). We cultured RPE 255 cells for 2

329   weeks without modification and transplanted them into the RCS rat model at p30, followed by

330   OKT 60 days later. The results demonstrated that the RPE 225 cells with high endogenous

331   TREX were indeed efficacious for transplantation (5 out of 5 animals; Fig. 4K). This result

332   indicates that TREX level was more accurate than time in culture as a biomarker of RPE255

333   transplantation efficacy. The HDI of the posterior probability of a successful transplant was

334   outside our ROPE and the binomial test also support a significant change (p=1.18e-06). Overall,

335   these results provide strong evidence that TREX is not only a biomarker of efficacious RPE, but

336   that TREX also functions to mediate RPE transplantation efficacy.

337   **Discussion**

338   The recent expansion in the number and diversity of RPE cell transplant products

339   developed for treating RPE loss in AMD has increased the need to characterize the identity and

340   efficacy attributes of RPE cell products. To advance this knowledge, we performed bulk and

341   single cell sequencing on adult RPESC-derived RPE cell products during *in vitro* differentiation.

342   These studies characterized the individual RPE subpopulations present and identified

343   transplantation efficacy gene markers. Our analysis revealed a clear distinction between the

344   transcriptomes of efficacious and non-efficacious RPE cell products in animal studies that were

345   consistent across different donors, allowing us to identify molecular pathways associated with

346   transplant efficacy and identify potential biomarkers for efficacious RPE cultures.

347   Our study identified several cell product attributes related to successful transplants by

348   exploring the transcriptional changes in RPESC-RPE cultures between cell populations with

349   varying transplant success in RCS rats. Our earlier work (Davis et al., 2017) showed that time in

350   culture correlated with vision rescue, indicating that time served as a surrogate measure of RPE

351   cell differentiation stage and vision rescue efficacy. Our present transcriptomic analysis

352   demonstrated an incomplete ability to separate the RPE cell transcriptomic profiles based solely

353   on time, pointing out the need to dissect RPE cellular heterogeneity at each timepoint. Recent

354   studies using single cell approaches have reported multiple RPE subpopulations within the stem

355   cell-derived, native human and mouse RPE with unique transcriptional and morphological

356   characteristics, indicating that the RPE cell layer is composed of heterogenous cell populations

357   that may have differing functional and clinical capabilities (Lee et al., 2022; Petrus-Reurer et al.,

358   2022; Voigt et al., 2019; Xu et al., 2021). Identification of the subpopulations responsible for

359   vision rescue can improve regulatory evaluation and improve RPE cell products to result in

360   better transplantation outcomes.

361   This transcriptomic analysis identified 13 subpopulations within the RPESC-RPE

362   cultures with distinct patterns of gene expression, consistent with recent findings from other

363   laboratories (Petrus-Reurer et al., 2022; Xu et al., 2021). Three clusters, 10, 6 and 2,

364   demonstrated enrichment for pathways associated with cell differentiation and proliferation

365   implying that these subpopulations had progenitor characteristics and were candidates to

366   improve both cell manufacture and transplant success. To enhance our ability to identify

367   subpopulations with more effective transplantation outcomes and uncover potential biomarkers,

368   we developed an *in vitro* integration assay to mimic engraftment after *in vivo* transplantation

369   when transplanted cells insert into the host RPE monolayer. This assay provides a more cost

370   effective and higher throughput method than animal models to examine how an RPE cell

12

371   population performs in the early engraftment step of a successful transplantation. Additionally, it

372   allows insight into how the genetic manipulation of RPE populations may affect the transplant

373   competency of specific cellular subpopulations. Using the *in vitro* integration assay, we found

374   that a subset of RPE cells were better able to integrate into a mature RPE monolayer. The

375   integrating RPE cells were more likely to express EZH2+, a marker of cluster 10. EZH2 is a

376   polycomb transcription factor involved in histone methylation and stem cell self-renewal and

377   differentiation (Cao et al., 2002; Kamminga et al., 2006; Karantanos et al., 2016). Notably,

378   cluster 10 showed the highest level of cell cycle markers, underscoring its stem- or progenitor-

379   like state. Further study of the prospectively enriched cluster 10 subtype will be worthwhile to

380   understand the roles of this important subpopulation in RPE layer cell biology, RPE cell product

381   manufacture, and RPE transplantation safety and efficacy outcomes. It is possible that

382   interactions between different clusters within the product are beneficial, and we anticipate future

383   studies to alter RPE subtype composition will be needed to determine the optimal product

384   composition.

385   Our analysis identified potential biomarkers of transplant efficacy, including our top

386   candidate, the lncRNA TREX. Recently, lncRNAs have emerged as potential markers for

387   predicting the quality of other types of tissues and cell products used for transplantation (Wong

388   et al., 2019; Zou et al., 2019). Manipulating TREX expression in our *in vitro* RPE integration

389   assay suggests that TREX may regulate cellular processes involved in engraftment into the host

390   RPE. However, TREX could be utilizing a wide variety of processes ranging from upregulating

391   cell migration to sustaining cell survival to increase RPE potency and further research is needed

392   to understand the mechanism. *In vivo* transplantation experiments support a role for TREX in

393   successful transplantation. An RPE cell line that expressed unusually high levels of TREX was

394   effective at the 2-week stage of differentiation that is usually ineffective. Furthermore,

395   knockdown of TREX caused efficacious REP lines to no longer support cell integration *in vitro.*

396   Finally, over-expression of TREX in non-efficacious RPE lines improved the probability of vision

397   rescue. These results provide strong evidence that TREX is directly involved in mediating

398   transplant efficacy and that TREX can predict the efficacy of RPE cell therapy products. Thus,

399   TREX is a critical attribute and candidate potency biomarker for the RPESC-RPE cell product.

400   While the emphasis of RPE cell product development has been placed on overall RPE purity

401   assessed by canonical RPE markers, it is likely that subpopulation heterogeneity found in adult

402   RPESC-RPE cells is reflected in other RPE cell products such as those derived from pluripotent

403   stem cells (35705015). Our work provides a path to establish a level of TREX expression to

404   identify RPESC-RPE cell products, which may apply to other types of RPE cell products, with

405   the highest likelihood of successful transplantation.

406   In conclusion, our findings highlight the importance of characterizing at the single cell

407   level stem-cell derived RPE cell products designed for cell transplantation. Knowledge about the

408   identity and quality attributes of RPE cell products is needed to guide their successful

409   development. Our findings highlight the importance knowing the subpopulations present within

410   an overall bulk RPE population. Further in-depth characterization of the different subpopulations

411   of RPE cells present in cell replacement products will be valuable to improve manufacture,

412   regulatory evaluation and transplant efficacy of RPE cell products to benefit retinal patients with

413   degenerative diseases such as dry AMD.

414   **Methods**

415   *RPE cell culture and subject details*

416   All RPE cell lines were generated from donor eyes obtained from certified eye banks

417   with consent for research use. The donor details are listed in Table S14. RPE cells were

418   cultured in a 24-well plate at 100K cells per well. Different RPE lines were used for different

419   experiments in this study due to the limited availability of adult RPESC-derived RPE cells. RPE

420   cells used in scRNA-Seq experiments were cultured on Transwell inserts (Corning, Corning,

421   NY). Cultures were maintained in RPE medium: DMEM F12 50/50 medium (Corning), MEM

422   alpha modification medium (Sigma-Aldrich), 1.25 mL Glutamax (Gibco), 2.5 mL Sodium

423   Pyruvate (Gibco), 2.5 mL Niacinamide (1M; Spectrum Chemical Inc., CA), 2.5 mL MEM NEAA

424   (Gibco), 2% or 10% heat inactivated fetal bovine serum, supplemented with THT (Taurine,

425   Hydrocortisone, Triiodo-thyronin), and 1.25 mL N1 medium supplement (Sigma-Aldrich). Cells

426   were incubated in a humidified incubator at 37 °C and 5% CO_2 and the medium was replaced

427   every 3 days.

428   *Bulk and single cell RNA sequencing preparation*

429   Bulk and single cell RNA sequencing were performed on cultured RPE cells collected at

430   2-, 4-, and 8-weeks post plating. Single cell suspensions were prepared using 0.25% Trypsin

431   (Thermo Fisher Scientific, Waltham, MA) and processed for either bulk or scRNA sequencing.

432   For bulk sequencing RNA was isolated with Direct-zol RNA kit (Zymo Research, Irvine, CA).

433   Library preparation was then carried out with the TruSeq Stranded Total RNA kit (Illumina, San

14

434   Diego, CA) ribo-depleted by the University of Rochester Genomics Research Center and

435   sequenced using a NextSeq550 high-output flow cell generating 2 × 151-bp read lengths.

436   For scRNA-seq, the RPE single cell suspension was stained with 1 µl SYTO64 dye

437   (Invitrogen, Carlsbad, CA) in 1mL PBS for 20 minutes at room temperature, then washed twice

438   in fresh PBS. Next, cells were diluted to 25,000 cells/mL and dispensed into ICELL8 3' DE chips

439   (Takara Bio, CA) using an MSND device (Takara Bio). Cell dispensing, and in-chip reverse

440   transcription PCR were performed using a 3' DE Chip and Reagent kit (Takara Bio) according to

441   the manufacturer's instructions. Following the extraction of PCR products, cDNA samples were

442   concentrated and purified using a DNA Clean & Concentrator-5 kit (Zymo Research) and

443   purified using a 0.6X proportion of AMPure XP magnetic beads (Beckman Coulter, Brea, CA)

444   according to manufacturer protocols. Library preparation was performed using a Nextera XT

445   DNA Library Preparation Kit (Illumina) according to the Takara's 3' DE chip and reagent kit

446   instructions. The quantification, and quality checks of the cDNA products were carried out at the

447   University at Albay's NextGen Sequencing core facility. The concentration of cDNA products

448   was quantified using a Qubit Fluorometer and the Qubit dsDNA HS Assay Kit (Thermo Fisher

449   Scientific). The quality of the cDNA product was checked using an Agilent High Sensitivity DNA

450   Kit and Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA) to ensure the complete

451   removal of contaminants. Libraries were sequenced using a NovaSeq 6000 high-output flow cell

452   generating 2 × 150-bp read lengths (GeneWiz).

453   *Data processing and analysis*

454   For bulk sequencing, the University at Rochester core facility processed the raw Illumina

455   BCL files and provided fastq files. Files were then mapped to hg19 and converted to bam files

456   using STAR (v2.4). Bam files were then read into R using the GenomicAlignments package

457   (code available at https://github.com/neural-stem-cell-institute/RPESC_TREX) to generate a

458   counts matrix for further analysis. For single cell sequencing, raw Illumina read BCL Files were

459   converted to fastq files using the bcl2fastq2 software (bcl2Fastq v2.19.1, Illumina, Inc). For

460   scRNA-seq the, fastq files were then merged into read1 (ICELL8 barcode sequence) and read2

461   (transcript sequence) fastq files. Using the metadata file from the ICELL8 system containing

462   single cell well information their specific nanowell barcodes, read1 and read2 files were

463   demultiplexed based on nanowell barcodes. The sequence reads were then converted to bam

464   files and mapped to hg19 using STAR (v2.5). The code utilized for processing the ICELL8 data

465   can be found at github (https://github.com/neural-stem-cell-institute/sc-pipeline). The final

466   transcript read counts RData file was used as the input reads matrix for further analysis in R.

467   All code as well as package versions for the analysis can be found at

468   https://github.com/neural-stem-cell-institute/RPESC_TREX. Briefly, bulk data was analyzed with

469   the EdgeR and DESeq2 packages to identify differentially expressed genes. Single cell data

470   was analyzed with the Seurat package (V3.1). Strict criteria were used for the QC of the single

471   cell dataset, including a minimum feature cutoff of 200 and a minimum cell cutoff of 3. Different

472   clusters were identified according to Seurat's clustering workflow. scRNA-seq enrichment

473   analysis was performed with the screp package (https://github.com/neural-stem-cell-

474   institute/screp/)

475   *Lentiviral Vectors Production and Infection*

476   The exonic TREX sequence (TCONS_00005049) was downloaded from the UCSC

477   genome browser (hg19) and a TREX insert with 15bp overhangs was synthesized with GeneArt

478   Gene Synthesis (Thermo Fisher Scientific). The insert was then cloned into a EcoRI cut TetO-

479   FUW vector (Addgene) using an In-Fusion HD Cloning kit (Takara Bio) and sequenced

480   (GeneWiz) to verify plasmid. Lentivirus was generated in 75% confluent 293FT cells by co-

481   transfecting the packaging plasmids pCMV-pLNV and pCMV-pVSVG, as well as either the

482   TetO-FUW-TREX or TetO-FUW plasmids into the cells with the XtremeGene HP DNA

483   transfection reagent at a ratio of 1:2.5 according to manufacturer's protocol. Supernatant was

484   collected 24 and 72 hours after transfection followed by centrifugation at 21,700g for 2.5 hours

485   to concentrate the viral particles. Viral particles were tittered via qPCR (ABM qPCR Lentivirus

486   Titration Kit) before storage at -80°C.

487   *Animal maintenance, transplantation, and analysis*

488   Royal College of Surgeons (RCS) rats were obtained from Dr. Shaomei Wang and Long

489   Evans rats from Taconic Biosciences, Inc. were maintained under a 12-hour light/dark cycle

490   according to IACUC-approved procedures. Transplantations were carried out as previously

491   described (Zhao et al., 2017). Briefly, RCS rats at P28–P32 days were treated with cyclosporine

492   (210 mg/L), then a 33-gauge needle was used to inject 1.5 μL of RPESC-RPE cell suspension

493   or BSS vehicle control under the retina under isoflurane anesthesia. Surgical success was

494   confirmed by visualization of a subretinal bleb using optical coherence tomography.

495   The spatial frequency threshold for opto-kinetic tracking (OKT) was measured (Table

496   S1s) by observers masked to treatment group using a device (CerebralMechanics) and

497   methods previously described (Douglas et al., 2005; Prusky et al., 2004) (Table S13). The

498   results were characterized by a non-normal bimodal distribution leading us to assign transplants

499   as either efficacious or not efficacious. In addition to binomial testing, a Bayesian approach was

500   utilized to compare groups using R. The code for the analysis is available at

501   https://github.com/neural-stem-cell-institute/RPESC_TREX. Briefly, a region of practical

502   equivalence (ROPE) was established using the results from 92 sham experiments. Each group

503   was then compared to this ROPE to determine if a treatment demonstrated an unambiguous

504   difference to the sham experiments.

505   *In vitro Integration assay*

506   A 'receiver' monolayer of RPE (RPE line 270) was grown on Transwells inserts (6.5 mm

507   diameter) for 8 weeks. 'Donor' RPE cells at P2 (transfected with a GFP lentiviral maker) were

508   grown for 4 weeks, made into a single cell suspension, and 12,000 RPE cells were transplanted

509   on to the 'receiver' RPE monolayer. After one week, Transwells were washed, fixed with 4%

510   paraformaldehyde, and imaged using an LSM780 Zeiss confocal microscope (Zeiss, Germany).

511   Twenty images, whose locations were distributed across the Transwell and in the same relative

512   position to each other, were taken for each Transwell. The images covered ~4% of the wells

513   surface. Confocal images were analyzed using the Zeiss ZEN software (V3.1). Integrated cells

514   were identified by determining if a nucleus from a 'donor' cell was in the same z-plane as

515   'receiver' nuclei using the 2.5D view. The percent of integrated cells for each sample was

516   calculated using the following equation:

517     $$\frac{Number\ of\ Integrated\ cells}{(plated\ donor\ cells \times covered\ well\ surface)} \times 100 = \%\ \text{Integrated RPE}$$

518   *Immunofluorescence staining*

519   Fixed Transwell membranes were immunostained for EZH2 and YEATS2 to determine

520   the identity of integrated cells. Cells were permeabilized using 1% Triton X-100 in DPBS for one

521   hour at room temperature. After removing Triton X-100 and rinsing 2 times with DPBS, cells

522   were incubated with EZH2 and YEATS2 primary antibodies (Thermo Fisher Scientific) according

523   to the manufacturer's instructions. Next, primary antibodies were removed and cells were

524   washed with DPBS twice and incubated with secondary donkey anti rabbit or goat anti mouse

525   antibody conjugated with Alexaflour 647 or Alexaflour 555 dyes and DAPI (1:1000) diluted in

526   DPBS with 0.5% bovine serum albumin (BSA) for 1 hours at room temperature. Immunostained

527   cells were then washed 2 times with DPBS for 5 minutes at room temperature and mounted on

528   glass slides using Fluoromount G (Thermo Fisher Scientific) and a coverslip. Samples were

529   imaged using an LSM780 Zeiss confocal microscope (Zeiss, Germany).

530   *Quantitative RT-PCR*

531   RNA was isolated with the Direct-zol RNA kit (Zymo Research) and cDNA was

532   generated using the Superscript VILO kit (Thermo Fisher Scientific) according to manufacturer's

533   instructions. Quantitative RT-PCR was performed using TaqMan gene expression assays and

534   TaqMan Universal Master Mix (Thermo Fisher Scientific). Either 18s-VIC or HPRT-VIC was

535   used as an internal control for all reactions.

536

537   **References**

538   Augsten, M., Hägglöf, C., Olsson, E., Stolz, C., Tsagozis, P., Levchenko, T., Frederick, M.J.,
539   Borg, A., Micke, P., Egevad, L.*, et al.* (2009). CXCL14 is an autocrine growth factor for
540   fibroblasts and acts as a multi-modal stimulator of prostate tumor growth. Proceedings of the
541   National Academy of Sciences of the United States of America *106*, 3414-3419.

542   Bennis, A., Gorgels, T.G.M.F., Ten Brink, J.B., van der Spek, P.J., Bossers, K., Heine, V.M.,
543   and Bergen, A.A. (2015). Comparison of Mouse and Human Retinal Pigment Epithelium Gene
544   Expression Profiles: Potential Implications for Age-Related Macular Degeneration. PloS One *10*,
545   e0141597.

546   Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and
547   Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing.
548   Science *298*, 1039-1043.

549   Cavallaro, T., Martone, R.L., Dwork, A.J., Schon, E.A., and Herbert, J. (1990). The retinal
550   pigment epithelium is the unique site of transthyretin synthesis in the rat eye. Investigative
551   Ophthalmology & Visual Science *31*, 497-501.

552   Cuomo, A.S.E., Seaton, D.D., McCarthy, D.J., Martinez, I., Bonder, M.J., Garcia-Bernardo, J.,
553   Amatya, S., Madrigal, P., Isaacson, A., Buettner, F.*, et al.* (2020). Single-cell RNA-sequencing
554   of differentiating iPS cells reveals dynamic genetic effects on gene expression. Nature
555   communications *11*, 810.

556   da Cruz, L., Fynes, K., Georgiadis, O., Kerby, J., Luo, Y.H., Ahmado, A., Vernon, A., Daniels,
557   J.T., Nommiste, B., Hasan, S.M.*, et al.* (2018). Phase 1 clinical study of an embryonic stem cell-
558   derived retinal pigment epithelium patch in age-related macular degeneration. Nature
559   Biotechnology *36*, 328-337.

560   Davis, R.J., Alam, N.M., Zhao, C., Muller, C., Saini, J.S., Blenkinsop, T.A., Mazzoni, F.,
561   Campbell, M., Borden, S.M., Charniga, C.J.*, et al.* (2017). The Developmental Stage of Adult
562   Human Stem Cell-Derived Retinal Pigment Epithelium Cells Influences Transplant Efficacy for
563   Vision Rescue. Stem Cell Reports *9*, 42-49.

564   Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson,
565   M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics
566   (Oxford, England) *29*, 15-21.

567   Douglas, R.M., Alam, N.M., Silver, B.D., McGill, T.J., Tschetter, W.W., and Prusky, G.T. (2005).
568   Independent visual threshold measurements in the two eyes of freely moving rats and mice
569   using a virtual-reality optokinetic system. Vis Neurosci *22*, 677-684.

570   Gu, X., Neric, N.J., Crabb, J.S., Crabb, J.W., Bhattacharya, S.K., Rayborn, M.E., Hollyfield,
571   J.G., and Bonilha, V.L. (2012). Age-related changes in the retinal pigment epithelium (RPE).
572   PloS One *7*, e38673.

573   Hafemeister, C., and Satija, R. (2019). Normalization and variance stabilization of single-cell
574   RNA-seq data using regularized negative binomial regression. Genome Biology *20*, 296.

575   Kamminga, L.M., Bystrykh, L.V., de Boer, A., Houwer, S., Douma, J., Weersing, E., Dontje, B.,
576   and de Haan, G. (2006). The Polycomb group gene Ezh2 prevents hematopoietic stem cell
577   exhaustion. Blood *107*, 2170-2179.

578   Karantanos, T., Chistofides, A., Barhdan, K., Li, L., and Boussiotis, V.A. (2016). Regulation of T
579   Cell Differentiation and Function by EZH2. Front Immunol *7*, 172.

580   Kopitz, J., Holz, F.G., Kaemmerer, E., and Schutt, F. (2004). Lipids and lipid peroxidation
581   products in the pathogenesis of age-related macular degeneration. Biochimie *86*, 825-831.

582   Lee, H., Lee, H.Y., Chae, J.B., Park, C.W., Kim, C., Ryu, J.H., Jang, J., Kim, N., and Chung, H.
583   (2022). Single-cell transcriptome of the mouse retinal pigment epithelium in response to a low-
584   dose of doxorubicin. Commun Biol *5*, 722.

585   Lee, J.H., Gao, C., Peng, G., Greer, C., Ren, S., Wang, Y., and Xiao, X. (2011). Analysis of
586   transcriptome complexity through RNA sequencing in normal and failing murine hearts. Circ Res
587   *109*, 1332-1341.

588   Liang, X.H., Sun, H., Nichols, J.G., and Crooke, S.T. (2017). RNase H1-Dependent Antisense
589   Oligonucleotides Are Robustly Active in Directing RNA Cleavage in Both the Cytoplasm and the
590   Nucleus. Mol Ther *25*, 2075-2092.

591   Lin, H., Xu, H., Liang, F.-Q., Liang, H., Gupta, P., Havey, A.N., Boulton, M.E., and Godley, B.F.
592   (2011). Mitochondrial DNA damage and repair in RPE associated with aging and age-related
593   macular degeneration. Investigative Ophthalmology & Visual Science *52*, 3521-3529.

594   Lu, J., Chatterjee, M., Schmid, H., Beck, S., and Gawaz, M. (2016). CXCL14 as an emerging
595   immune and inflammatory modulator. Journal of Inflammation (London, England) *13*, 1.

596   Mandai, M., Watanabe, A., Kurimoto, Y., Hirami, Y., Morinaga, C., Daimon, T., Fujihara, M.,
597   Akimaru, H., Sakai, N., Shibata, Y.*, et al.* (2017). Autologous Induced Stem-Cell-Derived Retinal
598   Cells for Macular Degeneration. The New England Journal of Medicine *376*, 1038-1046.

599 Mullins, R.F., Kuehn, M.H., Faidley, E.A., Syed, N.A., and Stone, E.M. (2007). Differential
600 macular and peripheral expression of bestrophin in human eyes and its implication for best
601 disease. Investigative Ophthalmology & Visual Science *48*, 3372-3380.

602 Ortolan, D., Sharma, R., Volkov, A., Maminishkis, A., Hotaling, N.A., Huryn, L.A., Cukras, C., Di
603 Marco, S., Bisti, S., and Bharti, K. (2022). Single-cell-resolution map of human retinal pigment
604 epithelium helps discover subpopulations with differential disease sensitivity. Proc Natl Acad Sci
605 U S A *119*, e2117553119.

606 Petrus-Reurer, S., Lederer, A.R., Baque-Vidal, L., Douagi, I., Pannagel, B., Khven, I., Aronsson,
607 M., Bartuma, H., Wagner, M., Wrona, A.*, et al.* (2022). Molecular profiling of stem cell-derived
608 retinal pigment epithelial cell differentiation established for clinical translation. Stem Cell Reports
609 *17*, 1458-1475.

610 Pfeffer, B.A., Becerra, S.P., Borst, D.E., and Wong, P. (2004). Expression of transthyretin and
611 retinol binding protein mRNAs and secretion of transthyretin by cultured monkey retinal pigment
612 epithelium. Molecular Vision *10*, 23-30.

613 Prusky, G.T., Alam, N.M., Beekman, S., and Douglas, R.M. (2004). Rapid quantification of adult
614 and developing mouse spatial vision using a virtual optomotor system. Invest Ophthalmol Vis
615 Sci *45*, 4611-4616.

616 Salero, E., Blenkinsop, T.A., Corneo, B., Harris, A., Rabin, D., Stern, J.H., and Temple, S.
617 (2012). Adult human RPE can be activated into a multipotent stem cell that produces
618 mesenchymal derivatives. Cell Stem Cell *10*, 88-95.

619 Sayols , S. (2020). rrvgo: a Bioconductor package to reduce and visualize Gene Ontology
620 terms.

621 Sharma, R., Khristov, V., Rising, A., Jha, B.S., Dejene, R., Hotaling, N., Li, Y., Stoddard, J.,
622 Stankewicz, C., Wan, Q.*, et al.* (2019). Clinical-grade stem cell-derived retinal pigment
623 epithelium patch rescues retinal degeneration in rodents and pigs. Science Translational
624 Medicine *11*.

625 Sparrow, J.R., Hicks, D., and Hamel, C.P. (2010). The retinal pigment epithelium in health and
626 disease. Current Molecular Medicine *10*, 802-823.

627 Strauss, O. (2005). The retinal pigment epithelium in visual function. Physiological Reviews *85*,
628 845-881.

629 Strunnikova, N.V., Maminishkis, A., Barb, J.J., Wang, F., Zhi, C., Sergeev, Y., Chen, W.,
630 Edwards, A.O., Stambolian, D., Abecasis, G.*, et al.* (2010). Transcriptome analysis and
631 molecular signature of human retinal pigment epithelium. Human molecular genetics *19*, 2468-
632 2486.

633 Vives-Bauza, C., Anand, M., Shiraz, A.K., Shirazi, A.K., Magrane, J., Gao, J., Vollmer-Snarr,
634 H.R., Manfredi, G., and Finnemann, S.C. (2008). The age lipid A2E and mitochondrial
635 dysfunction synergistically impair phagocytosis by retinal pigment epithelial cells. The Journal of
636 biological chemistry *283*, 24770-24780.

637 Voigt, A.P., Mulfaul, K., Mullin, N.K., Flamme-Wiese, M.J., Giacalone, J.C., Stone, E.M., Tucker,
638 B.A., Scheetz, T.E., and Mullins, R.F. (2019). Single-cell transcriptomics of the human retinal
639 pigment epithelium and choroid in health and macular degeneration. Proc Natl Acad Sci U S A
640 *116*, 24100-24107.

641 Whitmore, S.S., Wagner, A.H., DeLuca, A.P., Drack, A.V., Stone, E.M., Tucker, B.A., Zeng, S.,
642 Braun, T.A., Mullins, R.F., and Scheetz, T.E. (2014). Transcriptomic analysis across nasal,
643 temporal, and macular regions of human neural retina and RPE/choroid by RNA-Seq.
644 Experimental Eye Research *129*, 93-106.

645 Wong, W.K., Jiang, G., Sorensen, A.E., Chew, Y.V., Lee-Maynard, C., Liuwantara, D., Williams,
646 L., O'Connell, P.J., Dalgaard, L.T., Ma, R.C.*, et al.* (2019). The long noncoding RNA MALAT1
647 predicts human pancreatic islet isolation quality. JCI Insight *5*.

648 Xu, Z., Liao, X., Li, N., Zhou, H., Li, H., Zhang, Q., Hu, K., Yang, P., and Hou, S. (2021). A
649 Single-Cell Transcriptome Atlas of the Human Retinal Pigment Epithelium. Front Cell Dev Biol
650 *9*, 802457.

651 Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A. (2010). Gene ontology analysis for
652 RNA-seq: accounting for selection bias. Genome Biol *11*, R14.

653 Zhang, Y., Parmigiani, G., and Johnson, W.E. (2020). ComBat-Seq: batch effect
654 adjustment for RNA-Seq count data. bioRxiv.

655 Zhao, C., Boles, N.C., Miller, J.D., Kawola, S., Temple, S., Davis, R.J., and Stern, J.H. (2017).
656 Development of a Refined Protocol for Trans-scleral Subretinal Transplantation of Human
657 Retinal Pigment Epithelial Cells into Rat Eyes. J Vis Exp.

658 Zou, Y., Zhang, W., Zhou, H.H., and Liu, R. (2019). Analysis of long noncoding RNAs for acute
659 rejection and graft outcome in kidney transplant biopsies. Biomark Med *13*, 185-195.

660

Acknowledgements

Author Contributions

669 Conceptualization, S.T., J.S. and N.C.B.; Methodology, F.F., J.M.,S.T., J.S. and N.C.B.;

670 Software, F.F. and N.C.B. Validation, F.F.,J.M. and N.C.B. Formal Analysis, F.F.,

671 Y.W.,A.L.W.,S.B. and N.C.B.; Transplantation and Analysis, J.M., C.Z, N.A., G.P., N.C.B.;

672   Investigation, F.F., J.M.,Y.W.,A.L.W.,C.Z.,S.B. and N.C.B.; Resources, S.T., J.S. and N.C.B.;

673   Writing – Original Draft, F.F. and N.C.B.; Writing – Review & Editing, F.F., S.T., J.S. and N.C.B.;

674   Visualization, F.F., Y.W., A.L.W. and N.C.B.; Supervision, S.T., J.S. and N.C.B.; Project

675   Administration, J.S. and N.C.B.; Funding Acquisition, J.S. and N.C.B.

676   **Competing Interests**

677   Glen Prusky is a Principal of Cerebral Mechanics. All other authors declare no competing

678   interests.

679   **Materials and Correspondence**

680   Further information and requests for resources and reagents should be directed to and will be

681   fulfilled by the Lead Contact, Nathan Boles (nathanboles@neuralsci.org).

682   **Data Availability**

683   The raw scRNA-Seq data generated during this study are available at GEO; GSE211189

684   (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE211189).

685   **Code Availability**

686   Code written for this study is available at github (https://github.com/neural-stem-cell-

687   institute/RPESC_TREX).

688   **Figure Legends**

689   **Figure 1. RPESC efficacious for transplant have a distinctive transcriptional profile.**

690   (A) Outline of experimental plan for RNA-sequencing. (B) Dot plot illustrating the timepoints

691   RNA was collected, the cell lines used, and the known and predicted transplant status. SVD

692   analysis of the 4000 most variable genes by time (C) and by transplant status (D) was carried

693   out. A discernable pattern was not seen in the time analyzed data, however the data analyzed

694   by transplant status showed a clear separation between groups. (E-G) Distribution of

695   significantly different features by RNA type (E), transplant group for coding gene (F) and long

696   noncoding genes (G). GO enrichment analysis followed by a semantic similarity analysis (Table

697   S3) visualized by treemap (H). Selected pathways from a REACTOME pathway enrichment

698   analysis (Table S4).

699 **Figure 2. Singe cell RNA- sequencing reveals heterogeneity in RPESC cultures.**

700 Single cell RNA-seq (scRNA-seq) was carried out using the well based ICELL8 system with

701 Passage 2 cells collected at 2, 4, and 8 weeks. Data was analyzed using the Seurat package.

702 Dimensionality reduction followed by clustering was used to identify differing clusters of RPESC

703 reveal considerable heterogeneity (A). The clusters appeared to be well mixed when looking at

704 time of collection (B). All clusters had cells from each time period with some clusters showing

705 more variation in their composition (C). Dot plot demonstrating the top five gene markers for

706 each cluster as identified by a Wilcoxon Rank Sum test (D). Using the marker genes for each

707 cluster GO enrichment followed by a semantic similarity analysis (E) and REACTOME pathway

708 analysis (F) was carried out. The numbers above the dots in figure 2F show the number of

709 pathways in each Reactome tree.

710 **Figure 3. Intersection of Bulk and scRNA-seq data provides insight into potential**

711 **subpopulations responsible for transplant efficacy.**

712 (A) Upset plot showing the overlap between significantly changing genes (DEGs) in the bulk

713 RNA-seq data (Bulk-eff) and the marker genes in the scRNA-seq data based on cluster

714 (Cluster-SC) or time (Time-SC). (B) Membership by time of Bulk-Eff genes and single cell

715 marker genes classified by time in the single cell data. (C) The cluster membership of single cell

716 cluster marker genes found in the Bulk-Eff data by counts and percentage of all marker genes in

717 cluster. We next performed enrichment analysis of the genes shared between the Bulk-Eff and

718 clusters 2,6, and 10. (D) GO enrichment followed by semantic similarity analysis and (E)

719 REACTOME pathway enrichment analysis. (F) RPE cells were immunostained for EZH2 after

720 transplantation of RPE cells cultured for 4 weeks on a RPE monolayer cultured for 8 weeks in

721 an *in vitro* integration assay. EZH2 expression in the monolayer and integrated cells was

722 examined by confocal microscopy. A fraction of integrated cells with GFP expression exhibited

723 EZH2 expression (white arrowheads) indicating that they were members of cluster 10. EZH2 is

724 expressed in the integrating GFP cells (top) and in a small subpopulation of RPE cells in the 8-

725 week-old monolayer (bottom). The boxes below each panel show a section across the z-stack

726 image along the horizontal line indicated by the black arrowhead.

727 **Figure 4. TREX as a biomarker of transplant efficacy.**

728 Potential candidates of efficacy were determined by taking the genes that were expressed at

729 least 2-fold higher in the efficacious RPESC cells over the W2-NE and NE RPESC with at least

730 a hundred counts after normalization in the bulk data. (A) A plot of the TREX expression data vs

731 the known OKT results from transplants. (B) Expression of TREX across transplantation groups.

23

732   (C) TREX expression across multiple lines over time using qPCR with 18s as an internal control.

733   The threshold of TREX expression for a successful transplant is calculated based on the

734   midpoint point the TREX expression of the lowest expressing efficacious cells and highest

735   expressing non-efficacious cells. (D) TREX expression using qPCR with HPRT as the internal

736   control. (E) The cytoplasmic or nuclear fractions of RPE cells were isolated and qPCR was used

737   to look at the distribution of TREX within the cell. (F) Four gapmers against TREX were tested

738   for efficacy in suppressing TREX levels. G2 was used for all knockdown experiments. Results of

739   integration assay (G) for TREX knockdown by gapmer or (H) TREX overexpression (n=3).

740   RPESC transplants were made into the RCS rats and OKT measurements were taken after 90

741   days. A 'Region of Practical Equivalence', or ROPE, was calculated based on the highest

742   density interval (HDI) of 92 sham transplants. The HDI of (I) Empty vector control RPESC, (J)

743   TREX overexpressing RPESC, and (K) RPESC from line 255 having unusually high TREX

744   expression after just 2 weeks were calculated and the densities for each comparison to sham

745   were plotted. The HDI of both the TREX-OE cells and the RPE_255 cells did not overlap with

746   the ROPE demonstrating higher efficacy in transplantation, whereas the empty vector control

747   cells encompassed the ROPE demonstrating no clear difference.

748   **Figure S1. Expression of RPE related genes.**

749   (A) Intersection of RPE signature genes with genes expressed in each cluster. (B) Select gene

750   expression using simplified violin plots. (C) Expression of RPE65 in 8-week-old RPE cultures.

751   **Figure S2. Outline of Integration Assay.**

752   Description of novel integration assay's experimental design and analysis method.

753   **Figure S3. Expression of YEATS2 in RPE cells.**

754   YEATS2 was expressed in a subpopulation of 8-week-old RPE cells cultured on Transwells (top

755   row, white arrowheads), but no YEATS2 expression was found in integrated RPE cells labeled
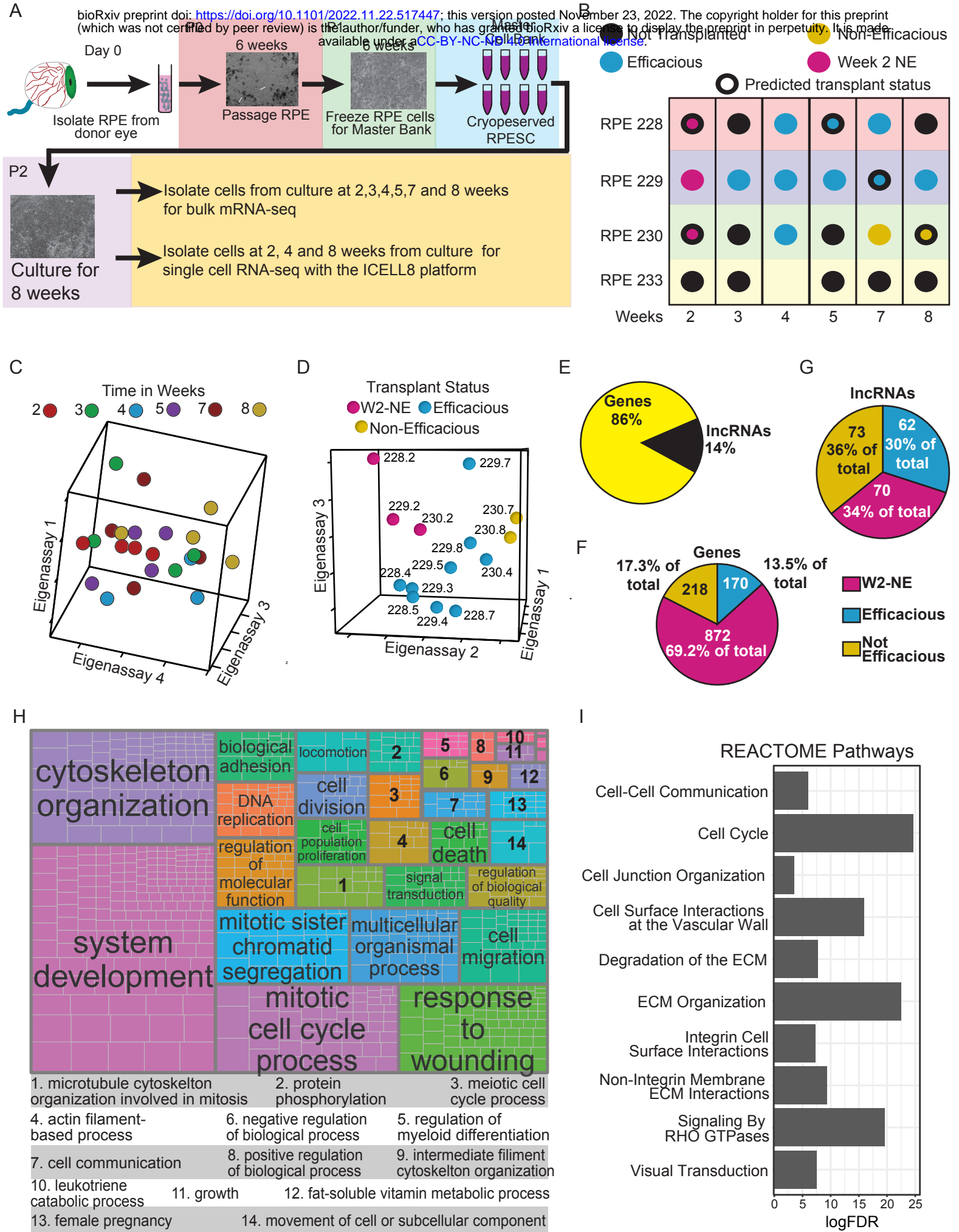
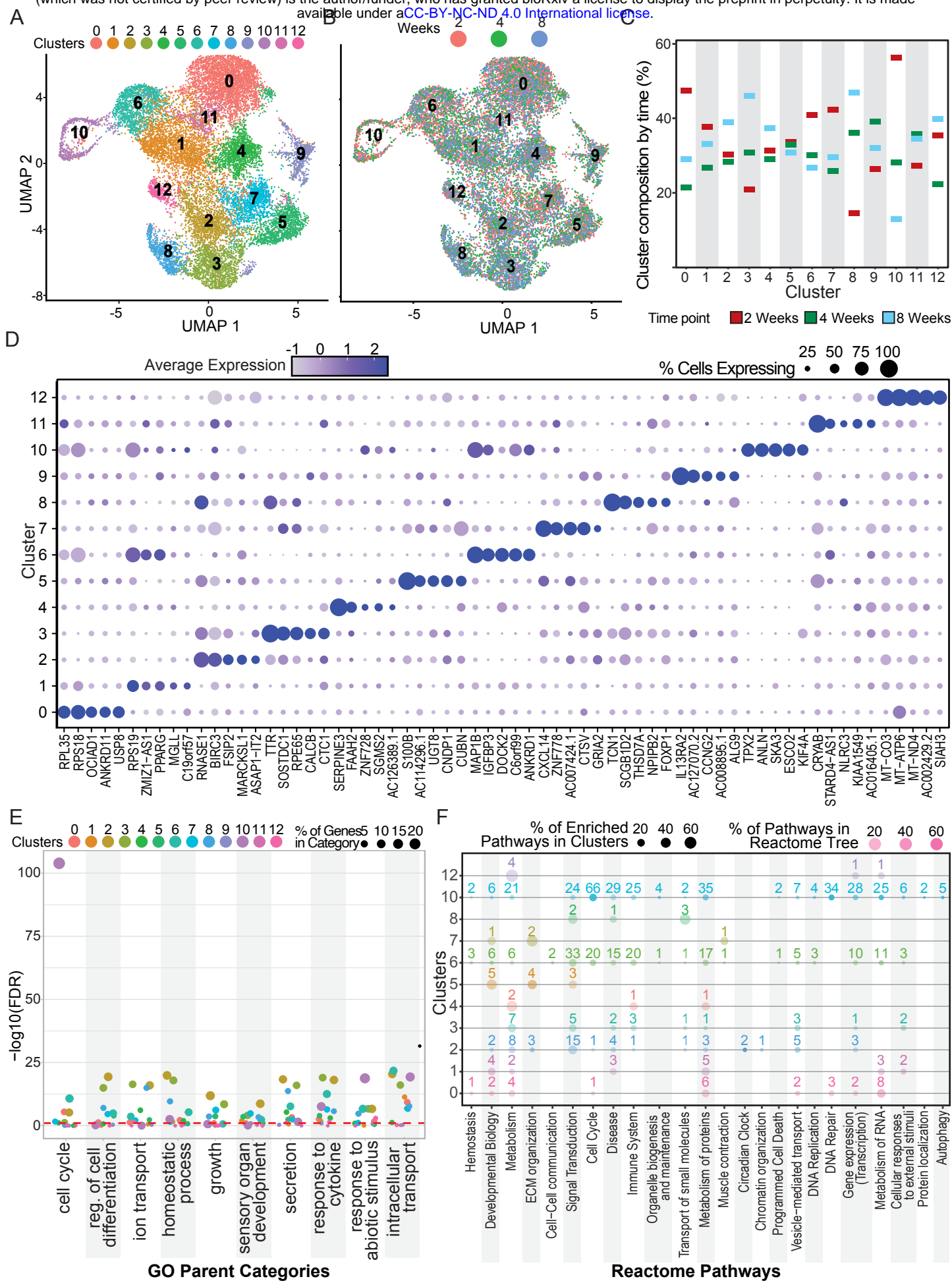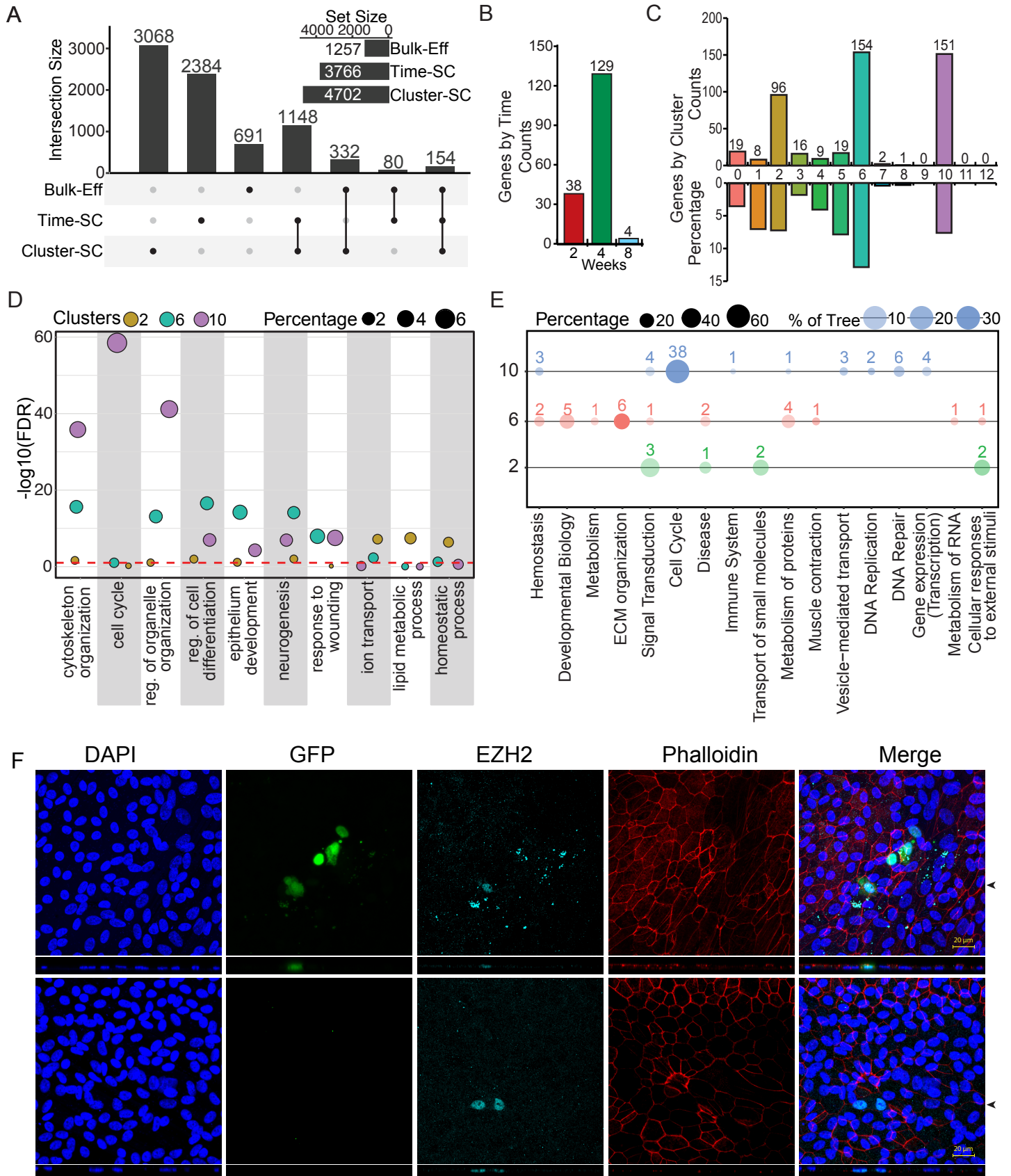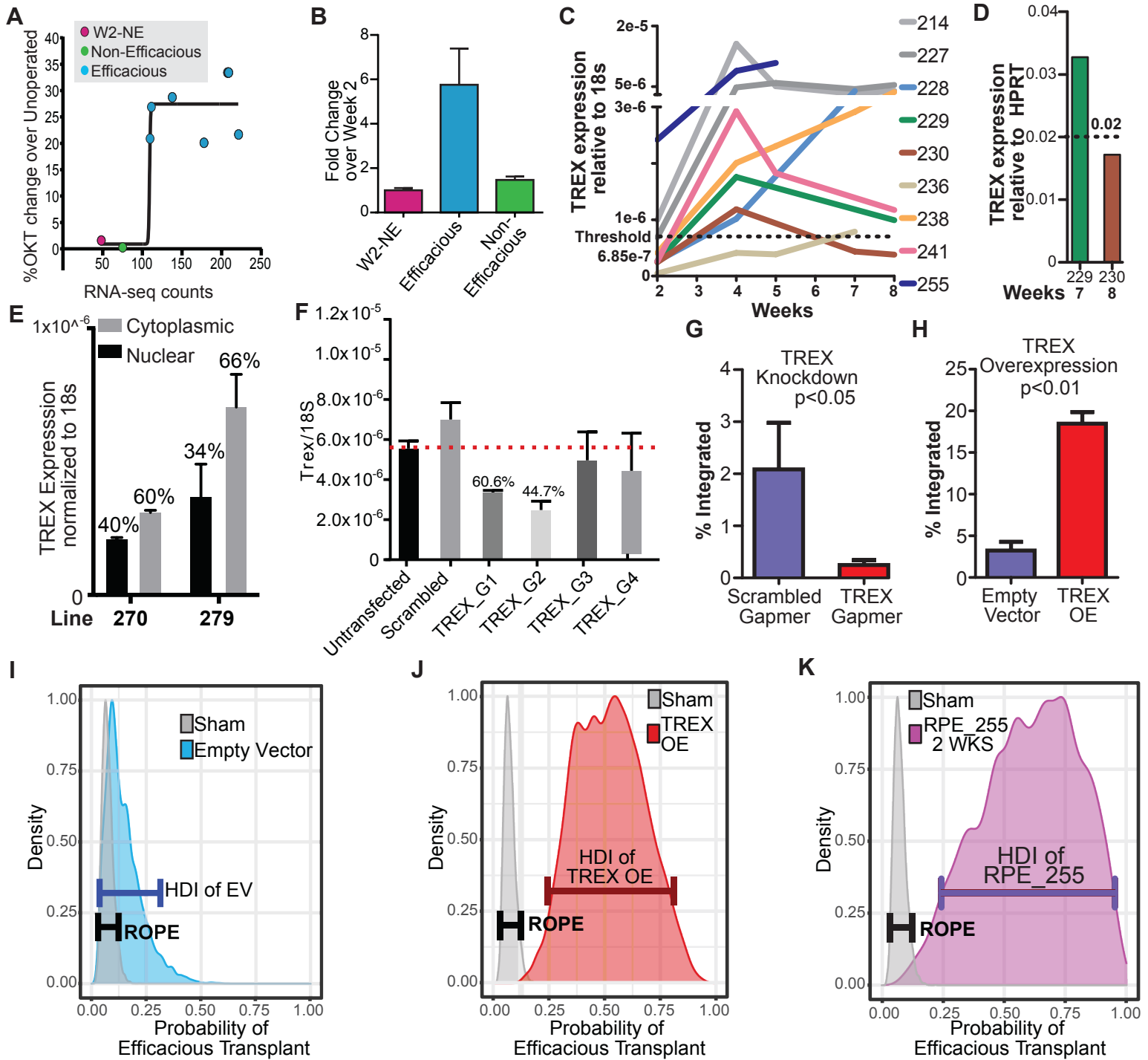756   with GFP (bottom row).

Figure 1

Figure 2

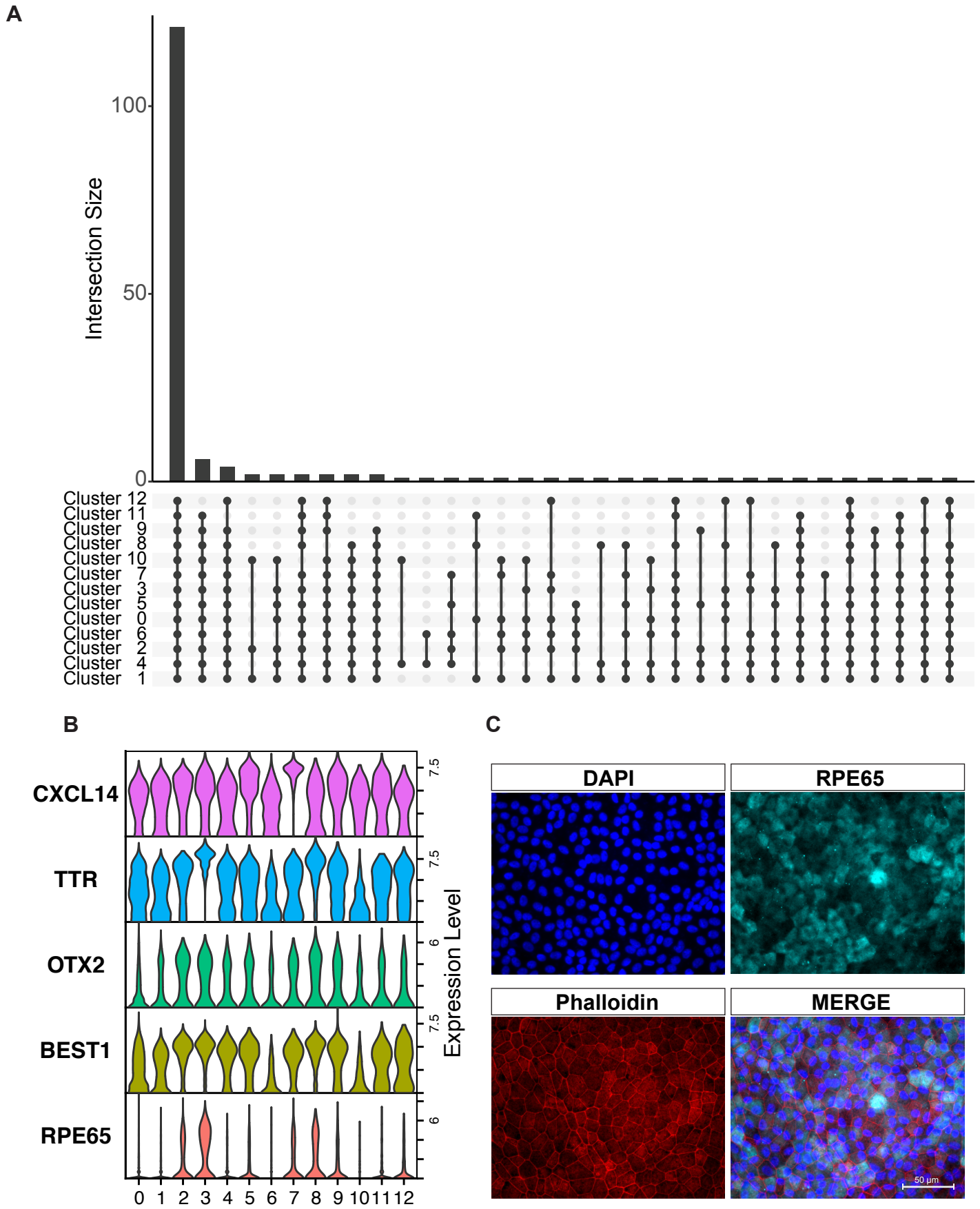**Figure 3**

Figure 4

A

B

C

Figure S1. Expression of RPE related genes. (A) Intersection of RPE signature genes with genes expressed in each cluster. (B) Select gene expression using simplified violin plots. (C) Expression of RPE65 in 8-week-old RPE cultures.
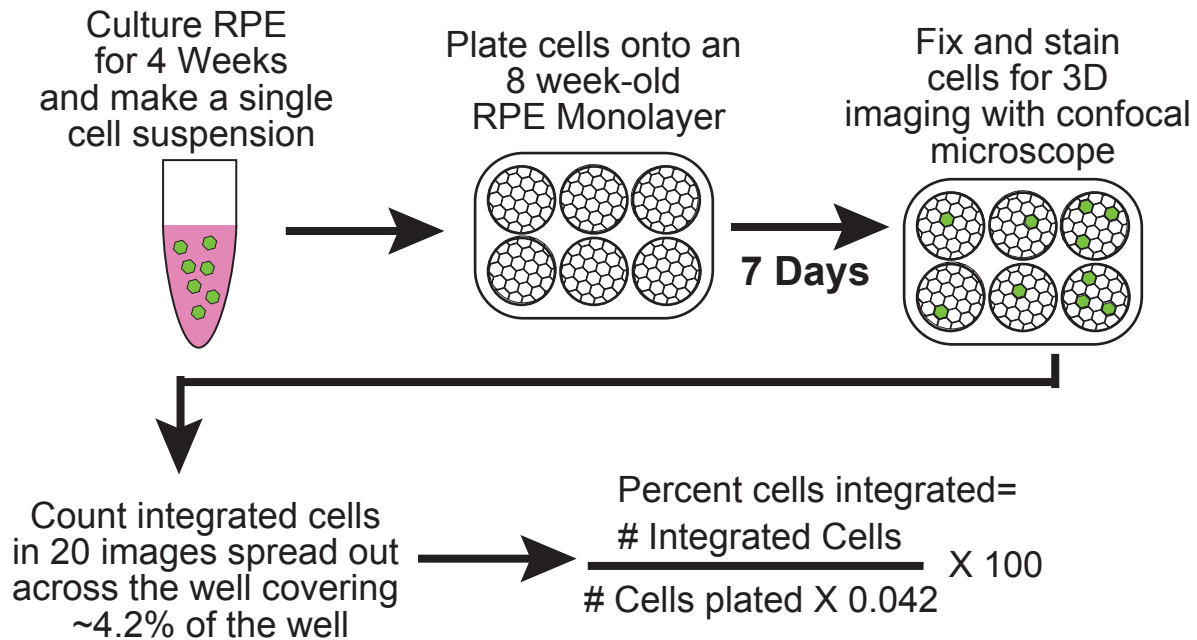
Figure S2. Outline of Integration Assay. Description of novel integration assay's experimental design and analysis method.
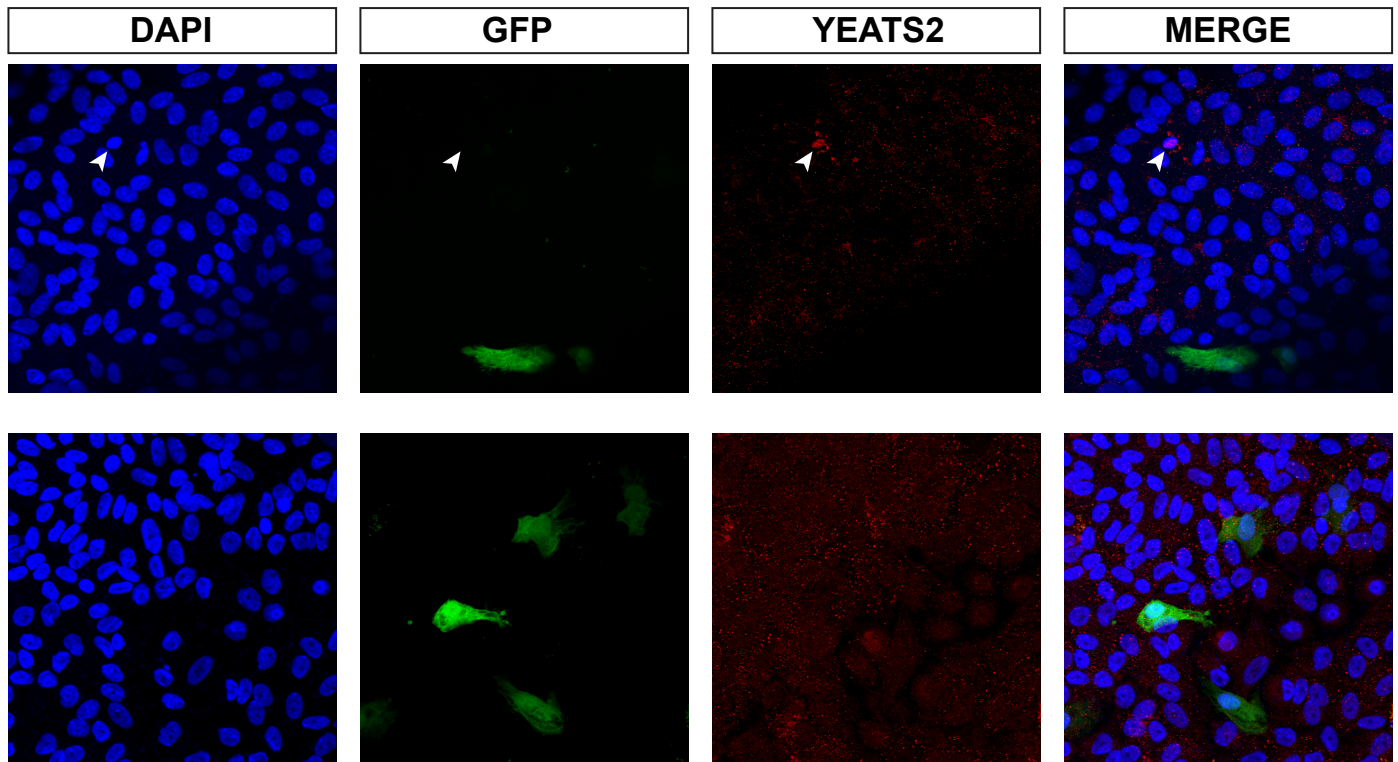
Figure S3. Expression of YEATS2 in RPE cells.
YEATS2 was expressed in a subpopulation of 8-week-old RPE cells cultured on Transwells (top row, white arrowhead), but no YEATS2 expression was found in integrated RPE cells labeled with GFP (bottom row).