

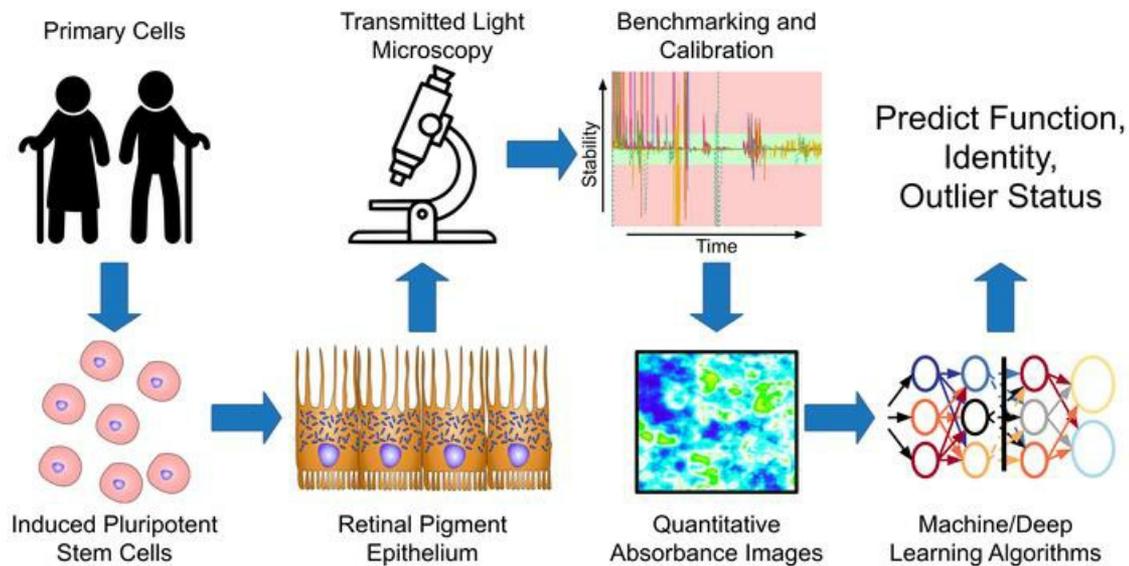
Deep learning predicts function of live retinal pigment epithelium from quantitative microscopy

Nicholas J. Schaub, ... , Peter Bajcsy, Kapil Bharti

J Clin Invest. 2019. <https://doi.org/10.1172/JCI131187>.

Research In-Press Preview Ophthalmology

Graphical abstract



Find the latest version:

<https://jci.me/131187/pdf>



Title: Deep learning predicts function of live retinal pigment epithelium from quantitative microscopy

Authors: Nicholas J. Schaub^{+,1,4,5}, Nathan A. Hotaling^{+,2,5}, Petre Manescu³, Sarala Padi³, Qin Wan², Ruchi Sharma², Aman George², Joe Chalfoun³, Mylene Simon³, Mohamed Ouladi³, Carl G. Simon, Jr.^{1,6*}, Peter Bajcsy^{3,6*}, Kapil Bharti^{2,6,7*}

Affiliations:

¹Materials Measurement Laboratory, Biosystems & Biomaterials Division, National Institute of Standards & Technology, Gaithersburg, MD, 20899

²Unit on Ocular and Stem Cell Translational Research, National Eye Institute, National Institutes of Health, Bethesda, MD 20892

³Information Technology Laboratory, Information Systems Group, National Institute of Standards & Technology, Gaithersburg, MD, 20899

⁴Department of Neurology, University of Michigan, Ann Arbor, MI, 48109

⁵These authors contributed equally

⁶Senior author

⁷Lead contact

+ Co-First Author

* Correspondence:

Kapil.Bharti@nih.gov - 301-451-9372, 10 Center Drive, NIH Building 10, Room 10B10, Bethesda, MD 20892

Carl.Simon@nist.gov - (301) 975-8574, 100 Bureau Drive, Building 224, Gaithersburg, MD 20899

Peter.Bajcsy@nist.gov - (301) 975-2958 - 100 Bureau Drive, Building 222, Gaithersburg, MD 20899

Abstract:

Increases in the number of cell therapies in the preclinical and clinical phases have prompted the need for reliable and non-invasive assays to validate transplant function in clinical biomanufacturing. We developed a robust characterization methodology composed of quantitative bright-field absorbance microscopy (QBAM) and deep neural networks (DNNs) to non-invasively predict tissue function and cellular donor identity. The methodology was validated using clinical-grade induced pluripotent stem cell derived retinal pigment epithelial cells (iPSC-RPE). QBAM images of iPSC-RPE were used to train DNNs that predicted iPSC-RPE monolayer transepithelial resistance, predicted polarized vascular endothelial growth factor (VEGF) secretion, and matched iPSC-RPE monolayers to the stem cell donors. DNN predictions were supplemented with traditional machine learning algorithms that identified shape and texture features of single cells that were used to predict

tissue function and iPSC donor identity. These results demonstrate non-invasive cell therapy characterization can be achieved with QBAM and machine learning.

Introduction:

Advancements in developmental biology and regenerative medicine have helped produce cell-based therapies to treat retinal degeneration, neurodegeneration, cardiopathies and other diseases by replacing damaged or degenerative native tissue with a new functional implant developed in vitro. Induced pluripotent stem cells (iPSCs) have extended the potential of cell therapies to permit transplantation of autologous tissues (1). However, as these autologous cell therapies are translated from the laboratory to the clinic, barriers to large scale biomanufacturing have become evident (2). Often, cell function is determined in the laboratory using sophisticated assays including electrophysiology, ELISAs, immunolabeling, small molecule quantification, and DNA/RNA-based assays. These assays require a trained user, are low throughput, expensive, end-point, and/or are time consuming. Thus, they are challenging to implement under current Good Manufacturing Practices (cGMP) which are required to produce any cell therapy product (3). Therefore, clinical translation of any iPSC-based therapy is made more challenging due to the scarcity of non-invasive, automated, fast, and robust assays (2) .

Here a methodology was developed that uses quantitative bright-field microscopy and neural networks to non-invasively predict tissue function. Clinical-grade iPSC derived retinal pigment epithelium (iPSC-RPE) from age related macular degeneration (AMD) patients and healthy donors were used as a model system to determine if tissue function could be predicted from bright-field microscopy images.

The retinal pigment epithelium (RPE) is a cellular monolayer, and RPE are of clinical interest because at least 11 IND-enabling or Phase I clinical trial stage studies are using RPE to treat AMD (4). Additionally, the appearance of RPE cells within the monolayer is known to be critical to RPE function (5–7) and recently a Phase I clinical trial used visual inspection of RPE by an expert technician as a biomanufacturing release

criterion for implantation (8). RPE cell appearance is largely dictated by the maturity of junctional complexes between neighboring RPE cells, and the characteristic pigmented appearance from melanin production (9). The junctional complex is linked to tissue maturity and functionality including barrier function (transepithelial resistance and potential (TER and TEP) measurements) (10) and polarized secretion of growth factors (ELISA) (9). Thus, cell appearance and function are correlated and may be predictive of each other.

The variability of transmitted light microscopy images makes them challenging to use for automated cell analysis and segmentation (11, 12) . Thus, the methodology developed in this study consists of two components. The first is quantitative bright-field absorbance microscopy (QBAM), which is an automated method of capturing images that are reproducible across different microscopes. The second component is artificial intelligence, which uses QBAM images to predict multicellular function. The artificial intelligence techniques were split into the categories of deep neural networks (DNNs) and traditional machine learning (TML). These techniques were chosen because of the high accuracy achieved in recent biomedical classification (13) and regression applications (14), their computational speeds for industrial scalability (15, 16), and their robustness to noise/variability (17, 18). We anticipate that the speed, reproducibility, and accuracy of these non-invasive, automated methods will aid in scaling the biomanufacturing process as cell therapies translate from the laboratory to the clinic.

Results

Methodology Overview and Test Case Description

QBAM was developed to achieve reproducibility in bright-field imaging across different microscopes. QBAM converts pixels from relative intensity units to absorbance units, an absolute measure of light attenuation (Figure 1a). To improve reproducibility of imaging, QBAM calculates statistics on images in real time as they are captured to ensure the absorbance value measured at every pixel has a 95% confidence of 10 mill-absorbance units (mAU, Figure S1). In this study three different band-pass filters were used for imaging, but the method scales to any number of wavelengths. QBAM imaging was implemented as a plugin for

Micromanager (for microscopes with available hardware) or a modular python package (for microscopes not supported by Micromanager) so that the user needs to only push a few buttons to obtain QBAM images.

Analysis of QBAM images was performed at the field of view (FOV) scale or the single cell scale (Figure 1b). At both scales DNNs were used for analysis, but for different purposes. The DNNs at the FOV scale were designed to directly predict two things: the outcome of functional/maturity assays (DNN-F, Figure S2a-b) or whether two sets of QBAM images came from the same donor (DNN-I). No image processing was performed prior to feeding images into DNN-F or DNN-I.

Single cell analysis began with a DNN that identified cell borders in QBAM images (DNN-S, Figure S2c). Next, visual features of individual cells were extracted from the QBAM images using the web image processing pipeline (WIPP, extracted features in Table S1) (19). The extracted visual features were then used to train TML algorithms to predict a variety of tissue characteristics including function, identity of the donor the cells came from, and developmental outliers (abnormal cell appearance). TML algorithms were then used to identify critical cell features that contributed to the prediction of tissue characteristics. To demonstrate the effectiveness of the imaging and analysis method, a proof of principle study was carried out on iPSC-RPE from the following donor types: healthy, oculocutaneous albinism disorder (OCA), and age-related macular degeneration (AMD). The iPSC-RPE from healthy donors were imaged as they matured, while AMD and OCA donors were imaged at a terminal timepoint once they had reached maturity. These two experimental designs were used to address two different use cases for the platform: (1) Non-invasively predicting iPSC-RPE function throughout the long culture to reduce the total number of replicates needed and provide more comprehensive/continuous assessment of samples and (2) Non-invasively providing assurance of tissue function, tissue identity, and if that tissue is an outlier from all other replicates on the implant that is being placed into patients on the day that it will be implanted. The assessment of cells was done using TML algorithms and DNNs (rather than traditional statistical approaches) due to the large heterogeneity of cells across scales (within a well, across replicates, between cultures, across clones, and across donors). To both

assess and incorporate this heterogeneity high-dimensional models were needed for accurate prediction of cell function, identity, or if cells were outliers.

iPSC-RPE from five different OCA patients and two healthy donors were imaged using QBAM to determine the sensitivity of the imaging method to biological variation and naturally low levels of melanin in OCA iPSC-RPE. In addition to weekly imaging, iPSC-RPE from healthy donors were assessed for the trans-epithelial resistance (TER) and polarized vascular endothelial growth factor (VEGF) secretion. TER is a measure of RPE maturity that increases as tight junctions form between neighboring cells. Polarized VEGF secretion is a measure of RPE function, where more VEGF is secreted on the basal side relative to the apical side of the cell monolayer (VEGF-Ratio). Finally, eight iPSC-RPE clones were derived from three AMD donors using clinical grade protocols. Here, clinical grade refers to production of cells using xenogeneic free reagents and cGMP compliant production processes. QBAM imaging was carried out on AMD iPSC-RPE once they had reached maturity.

Accuracy, Reproducibility, and Sensitivity of QBAM

QBAM imaging was validated with a combination of reference neutral density (ND) filters and biological samples. ND filters with known absorbance values were used as a reference to validate QBAM imaging by comparing absorbance measured on a UV-Vis spectrometer to absorbance measured using QBAM imaging. Absorbance measured with QBAM imaging strongly correlated with absorbance measured by the spectrometer across the visible spectrum (Figure 2a, $R^2 > 0.999$). To further validate the method, reproducibility of QBAM on ND filters was determined across three additional microscopes, each equipped with different filters, objectives, and light sources. The root mean squared error (RMSE) across all filters and microscopes was 66 milli-absorbance units (mAU), or $\approx 4.4\%$ at the highest absorbance value measured (Figure 2b).

QBAM imaging was then tested on live, progressively maturing iPSC-RPE derived from two different healthy donors. As expected from published literature (20), a general trend of increasing mean absorbance as

time progressed was found (Figure 2c). To determine how sensitive QBAM imaging was with respect to iPSC-RPE pigmentation, QBAM was used to image iPSC-RPE from five different patients with OCA (a disease known to reduce iPSC-RPE pigmentation). Mutant loci in OCA iPSC-RPE were sequenced to confirm the albinism type (OCA1A or OCA2) and the disease severity. OCA1A iPSC-RPE have severe albinism and produce no melanin (OCA8 and OCA26) and thus had the lowest image absorbance. OCA2 patients had a range of phenotypes from moderate (OCA103 and OCA9) to mild (OCA71), which corresponded with absorbance measures made by QBAM (Figure 2d). Despite iPSC-RPE from OCA1A patients producing low levels of pigment, the absorbance values were 2x higher than the lowest sensitivity of QBAM (10 mAU). Taken together, these data demonstrate the accuracy, reproducibility, and sensitivity of QBAM imaging.

Methodology to Predict iPSC-RPE Function from Absorbance Images

iPSC-RPE from healthy donors (Healthy-1, Healthy-2) were imaged to determine if QBAM imaging affected cell maturation and could measure a large range in variation of iPSC-RPE pigmentation. This was done using three culture conditions: (1) control iPSC-RPE (no treatment), (2) iPSC-RPE treated with a known inducer of RPE maturation (Aphidicolin), and (3) iPSC-RPE treated with a known inhibitor of RPE maturation (hedgehog pathway inhibitor-4, HPI4) (21) .

Control and aphidicolin-treated iPSC-RPE were found to mature as expected with increasing absorbance over the 8-week culture while HPI4-treated iPSC-RPE had a decreasing trend in absorbance over time (Healthy-2 is shown in Figure 3a-b and Healthy-1 in Figure S3a-b). Higher mRNA and protein expression of maturation markers were found in control and Aphidicolin treated iPSC-RPE than in HPI4 treated iPSC-RPE (Figure 3c and Figure S3d-f). The baseline electrical response (TEP and TER) and its change to physiological treatments of 5mM to 1mM potassium (K^+) or 100 μ M adenosine triphosphate (ATP) on the apical side was significantly greater in Aphidicolin treated iPSC-RPE and significantly lower in HPI4-treated iPSC-RPE relative to control (Figure 3d and Figure S3c). Further, iPSC-RPE maturation was evident from the presence of dense, native-like apical processes (Figure S3g-h)(21). From this set of experiments, it was concluded that: (1)

iPSC-RPE produced in clinical grade conditions had a mature-epithelial phenotype, (2) weekly QBAM imaging did not impact iPSC-RPE maturation, and (3) differences in pigmentation between mature (control and Aphidicolin) and immature (HPI4) iPSC-RPE could be quantified with QBAM imaging.

The capacity to predict iPSC-RPE monolayer function and phenotype from QBAM images of iPSC-RPE was assessed next. For healthy donor iPSC-RPE, mean QBAM pixel value had little correlation with TER (Figure 3f, $R^2=0.19$). However, the TER predictions by DNN-F highly correlated with actual TER values for the same samples (Figure 3g, $R^2=0.97$) and these predictions had a root mean squared error (RMSE) of $70.6 \Omega\cdot\text{cm}^2$. To incorporate this methodology into a biomanufacturing setting, a TER of $400 \Omega\cdot\text{cm}^2$ was used as a stringent threshold to classify iPSC-RPE monolayers as immature ($<400 \Omega\cdot\text{cm}^2$) or mature ($>400 \Omega\cdot\text{cm}^2$). Red-shaded regions in Figure 3f indicate TER values associated with immature iPSC-RPE, and red-shaded regions in Figure 3g indicate false positives and negatives. Based on this TER threshold, the DNN-F was 94% accurate, sensitivity was 100%, and specificity was 90% for classifying iPSC-RPE maturity. Similar trends were observed for VEGF-Ratios, where the polarized release of VEGF did not correlate well with QBAM mean pixel value (Figure S4a-b), but DNN-F predictions correlated highly with VEGF-Ratio measurements ($R^2=0.89$), and the RMSE of the VEGF-Ratio predictions was less than 1.0 (Figure S4c). The accuracy, sensitivity, and specificity of the VEGF-Ratio were all 100% (Figure S4c, samples with VEGF-Ratio < 3.0 were considered immature). From these experiments, we conclude (a) QBAM images of living cells can be used to predict TER and VEGF-Ratios with high fidelity, and (b) QBAM imaging could be used as a non-invasive means of functional validation of cells in lieu of measuring TER and/or VEGF-Ratio.

Extraction of Single Cell Features from Live QBAM Images of iPSC-RPE Monolayers

DNNs are known to have superior predictive power relative to other machine learning algorithms, but it is difficult to determine what image features DNNs use to make predictions. To understand which cell image parameters of iPSC-RPE predict monolayer function, image features of individual iPSC-RPE cells in QBAM images were calculated and used to train TML algorithms to predict iPSC-RPE function. A DNN was created to

segment individual living iPSC-RPE cells in QBAM images (DNN-S, Figure 4a). The DNN-S segmentation was validated by comparing cell features calculated from 12,750 iPSC-RPE with the same cell features calculated from ground-truth hand segmentations (Figure 4b, red and green histograms are cell features from hand corrected and DNN-S segmentations respectively). A comparison of 44 different features for DNN-S versus hand-corrected segmentations (Figure S5 and Table S2) showed a difference of $7.94\% \pm 4.42\%$ (mean \pm standard deviation) between the feature histograms with good pixel-wise agreement ($F-2 = 0.71$) (Table 1).

QBAM imaging and live cell segmentation allows for hundreds of cell image features to be measured and tracked non-invasively on individual cells throughout iPSC-RPE maturation. Thus, the trained DNN-S was used to segment QBAM images of living iPSC-RPE (Healthy-1 and Healthy-2 donors) treated with Aphidicolin, HPI-4, or nothing (control). Previously published cell image features (22–24) and intensity metrics (20) known to correlate to RPE maturation and health were then assessed for significance. Figure 4c shows the average number of neighbors (22, 23) each iPSC-RPE cell had as a function of drug treatment, and shows that HPI4 had significantly lower ($p < 0.001$) mean number of neighbors across all time points. Importantly, this method allows for unprecedented hierarchical granularity to the data; enabling not just whole well “bulk” tissue measurements (Figure 4c) but also as a function of field of view, as shown in Figure 4d, or at an individual cell level, as shown in Figure 4e. Figure 4d shows clustering of treatments based off of two features known to be important to RPE maturation and health, cell area (22, 24) and mean cell intensity (20). Figure 4e shows cell minimum intensity, a new metric related to iPSC-RPE function identified using TML in section Figure 5 below. These results demonstrate (a) the accuracy of the DNN-S segmentation with respect to hand drawn segmentations of individual cells, and (b) the differences between iPSC-RPE treated with different molecules can be described with discrete cell features.

Single Cell Image Features Can Predict iPSC-RPE Maturation and Function

Five different TML methods (multilayer perceptron - MLP, linear support vector machine - L-SVM, random forest – RF, principle least squares regression - PLSR, and ridge regression – RR) were used to

predict TER and VEGF-Ratios from the Healthy-2 donor iPSC-RPE using cell features obtained from cell-border segmentations of QBAM images. Figure 5a shows TER predictions for the MLP, the most accurate TML approach (RMSE=84.7 $\Omega\cdot\text{cm}^2$ and $R^2=0.94$). The red shaded regions represent false positives/negatives when using 400 $\Omega\cdot\text{cm}^2$ as a Quality Assurance or Quality Control (QA/QC) threshold. The MLP had an accuracy of 94%, sensitivity of 100%, and of specificity 90% (Figure S6a). However, a comparison of all algorithms (Table 2 and Figure S6a) showed the DNN-F was the most accurate predictor of TER, with the MLP's RMSE being 14.1 $\Omega\cdot\text{cm}^2$ higher than DNN-F (Table 2). For VEGF-Ratio predictions, random forest (RF) was the best TML method (RF, Figure S6b), but DNN-F had a 1.4x lower RMSE (Figure S6c).

The benefit of predicting iPSC-RPE monolayer function from cell features is the ability to determine single cell characteristics that indicate tissue-level function. For each TML method, cell features were ranked by importance. When comparing all TML models there was similarity in the most important features for predicting TER (Figure 5b-c) or VEGF-Ratio regardless of which TML method was used. Interestingly, key cell image features for predicting TER were spread across intensity, texture, and shape of cells (Figure 5b). Of the 10 most important features, three related to the shape of cells (Shape), three to the intensity of pigment within cells (Intensity), and four describe the distribution of pigment within RPE (Texture). A table is included in the online data, with an excerpt in Table S3, that shows what metrics specifically these features represent and their 95% confidence interval for each time point and drug treatment.

Taken together, the above indicates that using live cell segmentation, feature extraction, and TML, tissue TER and VEGF-Ratio can be predicted with a level of accuracy approaching the prediction accuracy of DNNs to analyze QBAM images. The benefit to TML methods compared to DNNs is that TML models can identify discrete cell features that indicate iPSC-RPE monolayer function. This allows manufacturers and clinicians to better understand the models, determine why a culture has been predicted to have low function, and perhaps develop mitigation strategies based on the parameters that have shown to be predictive of the phenotype.

Accuracy of Function Predictions is Robust Across Multiple Clinical-Grade AMD Patient Derived iPSC-RPE

To determine if the methodology was robustness across multiple donors and multiple preparations, DNN-F and TMLs were used to predict TER and VEGF-Ratios of clinical grade iPSC-RPE from three AMD patients across 8 iPSC clones. Figure 6a shows representative absorbance images of one iPSC-RPE sample from each of the AMD donors and corresponding SEM images of iPSC-RPE apical processes, confirming iPSC-RPE polarized phenotype. Monolayer maturation was assessed by TER (Figure 6b) and VEGF-Ratio (Figure S7a) as well as the assays shown in Sharma et al (25). Mean QBAM pixel values of iPSC-RPE were measured for fully-mature AMD-iPSC-RPE. As with the healthy donors, mean absorbance did not correlate well with TER (Figure 7b, $R^2=0.015$) or VEGF-Ratio (Figure S7a, $R^2=0.50$). However, the random forest (RF) model was able to predict TER to a similar degree of accuracy (RMSE = $70.9 \Omega \cdot \text{cm}^2$, $R^2 = 0.92$) as with Healthy donors (Table 2). DNN-F was also used to model TER and the predicted vs. actual values correlated well ($R^2 = 0.91$, Figure S7b).

To further assess the robustness of the methodology, TML models were trained on different combinations of AMD-iPSC-RPE monolayers. A total of 18 unique training image subsets were formed, where each image subset contained test data which had images of one iPSC-RPE sample from each donor (Table S4 shows all combinations in detail). The mean TER RMSE was $86.9 \Omega \cdot \text{cm}^2 \pm 14.3 \Omega \cdot \text{cm}^2$ (Table S4) across all clone subsets showing that the prediction error is similar when scaled to a larger donor subset (eight AMD-iPSC-RPE samples, Figure 6) as compared to a single sample (Healthy-2, Figure 5). From the measured and predicted values, a 95% confidence interval was constructed and is shown as the blue region in Figure 6c and Figure S7c. iPSC-RPE falling outside of this region could be considered as “out of specification” in a biomanufacturing environment and would be recommended for further testing.

Finally, the most important cell features for predicting AMD-iPSC-RPE monolayer TER (Figure 6d) and VEGF-Ratio across all donor/sample combinations was assessed to determine if features used to predict AMD-iPSC-RPE function were similar to those of Healthy-2 (Figure 5). Interestingly, only four features

overlapped between these two groups - Zernike n4- l0 polynomial (Shape 1), mass displacement (Intensity 2), and the third inverse difference moment at 135° (Texture 2) and at 45° (Texture 1). Overall, the models derived from the clinical grade iPSC-RPE images were able to predict iPSC-RPE phenotype across multiple donors/samples and to determine the common cell image features of living cells across multiple donors that predict iPSC-RPE monolayer function. Further, the differences in the feature importance from Healthy-2 and the AMD-iPSC-RPE suggest there might be both donor specific features for predicting function, as well as features that are common to multiple donors.

Classification of Developmental Outliers and Identity of iPSC-RPE Donors using QBAM

QBAM images were used to determine if there were any developmental outliers based on cell image features in the eight clinical grade iPSC-RPE samples from three AMD patients. Developmental outliers are defined as an intra-donor comparison in which iPSC-RPE monolayers that are different from other iPSC-RPE monolayers based on cell image features may warrant additional analysis to determine if the monolayer developed properly. Figure 7a shows the principle component analysis of cell image features from the QBAM images of the AMD donor/sample preparation. iPSC-RPE derived from a given donor clustered well together based on cell image features, except AMD1 clone A (1A) and AMD3 clone A (3A) as identified in the hierarchical dendrogram of the PCA shown in Figure 7b. Analysis of clone 1A iPSC-RPE showed 894 changes in the onco-exome as compared to the starting donor material as shown by Sharma et al. (25). Clone 3A iPSC-RPE was found to have a lower pigment level than its “sibling” iPSC-RPE (Figure 6b). When the cell image features were analyzed the cell pigmentation and shape were found to be the two most dominant feature classes in identifying these iPSC-RPE as outliers; a full description of features can be seen in the online dataset.

For each iPSC-RPE monolayer, the cell donor identity was predicted from QBAM images using multiple TML models. Additionally, a new DNN (DNN-I) was developed to determine if two different iPSC-RPE images had the same identity (Table 3 and Table 5). Donor identity is defined as an inter-donor comparison to

determine which images came from which donor (TML models), or to identify if two images are from the same donor (DNN-I). The TML algorithms took the features derived from QBAM images as an input and gave a donor identification as an output. In comparison, DNN-I took two QBAM images as an input, and classified the images as coming from the same or different donors. The TML approach was able to classify the donor identity of RPE cells derived from new clones of a donor when it had been trained on images of RPE cells derived from other clones of that same donor, it was not able to classify “new” donors that were not present in the training data. While the DNN-I strategy for binary classification of two images as “same” or “not same” gives the DNN-I the potential to classify “new” donors that were not used during training. Linear support vector machines (L-SVM) were found to have the highest accuracy of all TML algorithms tested (Table S5) with an accuracy of 76.4% (2.3x random chance), a sensitivity of 64.6%, and a specificity of 82.3%. Across all donor/sample combinations, DNN-I had better performance with an accuracy of 85.4% (2.6x random chance), a sensitivity of 80.9%, and a specificity of 86.8% (Table 3-4). Interestingly, cell image features that were key to distinguishing AMD iPSC-RPE from each other were similar across different iPSC-RPE combinations (Figure 7c) and consist of features that were different from the features used to identify tissue function (compare Figs 5b,c with Figure 7c) and developmental outliers. A general difference between the top 50 features used in each application were: shape features were important to identify clones as outliers (23 of the top 50 features), texture features were important for donor classification of clones (25 of the top 50 features), and shape and texture features were important to classify iPSC-RPE function (40 of the top 50 features).

Discussion

Absorbance Imaging

Data inputs are critical to successful analysis. Thus, the methodology developed here starts with a rigorous, reproducible absorbance imaging method using a bright field microscope. Absorption imaging has been reported in the literature, (26–28) but the QBAM technology developed here can be implemented on any standard-bright field microscope and it uses real-time, automated, statistically robust methods to provide high

confidence in image quality and reproducibility. The advantage to using absorbance rather than raw pixel intensities is that absorbance is an absolute measure of light attenuation (29). Raw pixel intensities can vary with microscope configuration and settings (e.g. uneven lighting, bulb intensity and spectrum, camera, etc) that make comparison of images difficult even when the images are captured on the same microscope. Converting to absorbance values overcomes many issues related to image reproducibility (explained in more detail in the methods). The combination of automation, converting pixel intensities to absorbance values, calculating absorbance confidence, and establishing microscope equilibrium through benchmarking ensures the quality of image data captured with QBAM.

The robustness of QBAM was validated in three systems to ensure the measurement could be used in multiple circumstances: (1) synthetic standards (Figure 2a-b), (2) healthy biological samples (Figure 2c and Figure S3), and (3) drug induced models of iPSC-RPE maturity (Figure 3 and Figure S3). Analysis of the QBAM images showed absorbance values agreed with “known” synthetic standards and could assess the development of pigment in both healthy and diseased RPE. The results in Figure 2 and Figure S3 also highlight the robustness of the measurement across multiple microscopes or imaging configurations. Errors between different microscope measurements of the same sample were within 4.4% of the signal as compared to an average error of 31% on reference standards for VEGF ELISAs (30) and 100% for TER measurements in epithelium (31). This represents one to two orders of magnitude reduction in variability for a potential release assay for a cell therapy product or when used in a drug screening methodology.

QBAM is optimized for determining absorbance for cell types that absorb light. However, results with non-pigmented (severe) albinism cells indicate possible applications on other non-pigmented cell types. Due to the signal to noise ratio in these cells being over twice the lower limit of detection the Authors hypothesize that either Absorbance or Transmittance data could be broadly applicable to any cell type in culture. Therefore, it may be suitable for assays where light absorbance is used, such as viable cell counting using trypan blue staining, histological staining, or analysis of light-absorbing biological specimens such as pigmented skin cells

or dopaminergic neurons that express neuromelanin. In cases where transmittance values may be preferred such as histology, the statistics can be modified to generate reproducible images of tissue sections. Also, this methodology is generalizable to any multi-spectral modality since none of the calculations are wavelength specific. Finally, we expect that the methods could be suitable for hyperspectral autofluorescent imaging which can identify cell borders and sub-cellular organization in non-pigmented cells (32) .

Prediction of iPSC-RPE Function

Neural networks and machine learning algorithms were trained with a full range of cell phenotypes by using healthy iPSC-RPE and drugs that were known to inhibit (HPI4) and promote (Aphidicolin) iPSC-RPE maturation (21). Having diverse phenotypes in the training set enhanced the robustness of the algorithms as was expected from the literature (18) . Additionally, the method worked on two different donors not only as an end-point assay of tissue health, but also as a non-invasive tool for tracking tissue development during the long maturation period (≈ 35 days) (33) . Importantly, the accuracy of the algorithms in predicting both TER and VEGF-Ratio was close to the measurement uncertainty for both TER (31) and VEGF (30, 34) . The generalizability of such metrics across cell lines, regardless of donor, adds to the broad utility of this work, as here we show that even with a limited data-set strong predictive ability, regardless of donor and clone, was shown (Figure 5).

The cut-off ratios for biomanufacturing, $400 \Omega \cdot \text{cm}^2$ for TER and 3.0 for VEGF-Ratio, were chosen according to the literature (33, 35, 36) . However, it is important to note that similar accuracies, sensitivities, and specificities were found when assessing a range of TER values from $200 \Omega \cdot \text{cm}^2$ to $1000 \Omega \cdot \text{cm}^2$ or VEGF-Ratio from 1 to 5 and thresholds should be set according to the manufacturer's specifications. In this study we observed a higher prediction error for the iPSC-RPE from Healthy-2 compared to the prediction error of iPSC-RPE from multiple AMD donors. We hypothesize two reasons for this: (1) Healthy-2 iPSC-RPE had a wider range of both TER and VEGF-Ratio values due to the inclusion of positive and negative controls than did AMD-iPSC-RPE which were manufactured using a cGMP based process with the goal of reproducibly

manufacturing healthy mature iPSC-RPE monolayers. (2) The AMD-iPSC-RPE had a larger dataset for training as it included all Healthy-2 data as well as the training data from the AMD-iPSC-RPE. In machine learning more data generally leads to more accurate models; however, implementation of our methodology in any application should give special consideration to ensuring a wide range of conditions (i.e. more donors and positive/negative controls) for training than presented in this proof-of-principle study to ensure the robustness of the model.

As expected, deep learning had a lower RMSE of prediction for both TER and VEGF-Ratio as compared to the TML approaches. However, using TML approaches allowed discovery of important cell image features (Figure 5b and 6d). These two approaches were chosen because we perceive two motivations for cellular product manufacturers, regulators, clinicians, and/or researchers. Motivation 1: often in manufacturing, clinical settings, or high throughput screens, time is a critical factor and a clear “go/no-go” or simple read-out is desired. In these cases, algorithms that provide the highest accuracy and are most robust to noise should be used, and deep learning is an excellent tool for this application (13, 18) . Motivation 2: frequently in research, insight into underlying mechanisms of function are important. In these cases, more scrutable methods in which the importance of cell image features can be determined are necessary. For this motivation TML approaches are desirable because their underlying architecture is simple enough to be understood and the importance of factors (here cell image features) to predicting tissue function can be obtained.

Feature extraction from QBAM images leads to hundreds of features based on the shape, intensity and texture of cells at both the single cell level and across larger cell populations. Many of these features are mathematical abstractions that lack meaningful connection to cell function. Therefore, even though TML models may be more interpretable than DNNs, the features that compose these models may not be relatable to the underlying biology (37–39) . Nonetheless, for cell manufacturing purposes what these features are and how they relate to the underlying biology is less important than being able to identify their 95% confidence

interval and ensuring that future batches/clones from donors fall within this range; making their use here in TML models relevant regardless of their relationship to the underlying biology.

Identification and Clustering of iPSC-RPE from Clinically Relevant Donors

Currently, there is a critical need to develop a non-invasive, clinically-compatible, assay to confirm the quality of cell therapy products just prior to implantation. The PCA and cluster analysis shown in Figure 7a-b could serve this unmet need. Using this approach, or similar clustering techniques, the similarity of the to-be implanted article to other technical replicates (or previous successful manufactured batches) can be non-invasively assessed for the first time. Additionally, the donor identification work done using DNN-I or the L-SVM in Table 3-5 and Figure 7c can serve as a QA/QC step to detect simple culture errors such as switched wells or a flipped plate and to match identity of this implant to other replicates from the same donor. This will be especially important in a facility that manufactures thousands of autologous therapeutics and must confirm the identity of each patient's dose. However, to determine a donor's identity (as with the L-SVM) a reference differentiation and image data set is needed for each patient line generated. This can be partially mitigated by using longitudinal imaging sets during the first culture, however these models will necessarily be less robust as the training data will be limited.

The two most important features to identifying the developmental outliers in Figure 7a and Figure 7b can be seen in Figure S7e and were the standard deviation of the maximum intensity (Intensity 8) of the iPSC-RPE and the Zernike n5- l3 polynomial (Shape 10). The deviation of the max intensity parameter agrees well with absorbance results showing that AMD3 Clone A derived iPSC-RPE had a lower absorbance than iPSC-RPE derived from other clones of AMD3. While zernike polynomials have been useful for detecting invasive cancer cell shapes (40) and in classifying tumors (41); which leads the authors to hypothesize it was critical to detecting a difference between AMD1 clone A, that had 894 onco-exome sequence alterations than other AMD-iPSC-RPE samples. The authors acknowledge that the cluster analysis performed in Figure 7b is not a conclusive proof that the development of oncogene sequence alterations can be identified with only

QBAM imaging. However, in a cell therapy manufacturing setting cluster analysis could be used to screen individual therapy replicates to determine if there are outliers that might need additional scrutiny. Also, because the assay is non-invasive, this information will provide surgeons with additional confidence in the quality of the actual transplant being delivered to the patient which has not been possible previously.

In conclusion, the methodology presented here shows that QBAM imaging can be used to assess the pigmentation development of healthy and diseased iPSC-RPE non-invasively. DNNs can analyze these images and accurately predict cell TER and VEGF-Ratio across 10 different iPSC-RPE preparations. Additionally, QBAM images contain sufficient information to allow DNNs to accurately segment cell borders of live RPE cells. Once segmented, 100s of features can be calculated per cell and, using these features, cell function can be predicted, outlier samples can be identified, and donor identity can be confirmed. All of this information can be obtained on the tissue that is being implanted into the patient with an automated bright-field microscope, without the need for expertise from a clinician, in just minutes. Thus, QBAM has potential application in a biomanufacturing setting where thousands of manufactured RPE units could be non-invasively tested and qualified for clinical use by a technician.

Materials and Methods

A complete materials and methods section is provided in the supplemental information. Below is an abbreviated methods section.

Quantitative Bright-field Absorbance Microscopy

We developed a method of determining absorbance values from bright-field images called Quantitative Bright-field Absorbance Microscopy (QBAM). The basic principle of QBAM imaging is absorbance, which is an absolute measurement of light attenuation. Absorbance is calculated using QBAM by capturing an image with the light off, then a second image with the light on but no sample on the stage, and finally an image of the sample. These three images are then used to calculate absorbance of every pixel in the image of the sample.

To make the method microscope independent, the method includes a protocol of benchmarking a microscope to ensure accuracy of the measurement. Statistical criteria were developed to ensure accurate measurements of absorbance based on metrics obtained from benchmarking. The statistical criteria define the number of images that must be captured to ensure an accuracy of 0.01 absorbance units, and are calculated in real time as images are captured. If image data does not meet the criteria, then additional images are captured until there is sufficient information to obtain an accurate measurement. A thorough description of this method is provided in the supplemental Methods information.

Culture, Assays, and Imaging of iPSC-RPE

After differentiation, Healthy-1 and Healthy-2 iPSC-RPE were placed on fresh transwell plates for maturation. Beginning 2 weeks or 1 week from the start of the culture, iPSC-RPE were exposed to Aphidicolin, Hedgehog Pathway Inhibitor 4 (HPI4), or neither drug. On the same day that cells were exposed to one of the three experimental conditions, cells were imaged weekly using QBAM for the duration of cell culture (8 weeks of total culture). A total of 12 images were captured in each well (4x3 overlapping grid, 1-15% overlap). On the same day that cells were imaged, trans-epithelial resistance (TER) was measured and supernatant was collected to measure cytokine release (vascular endothelial growth factor, VEGF).

Deep Neural Networks

Three main deep neural networks were created for this work, and the development and training of each of them is described in detail in the supplemental methods. The first deep neural network (DNN-F) was trained to predict TER and VEGF release, where the input image was a 1024 x 1024 pixel, three color QBAM image and the output was a TER value and a VEGF value. The second deep neural network (DNN-S) was trained to identify borders of cells in QBAM images, where the input of the network was a 256x256 pixel, three color QBAM image and the output was a cell border segmentation. The third network (DNN-I) was trained to take two QBAM images of iPSC-RPE and determine if the images came from the same donor. The training data for

this network included iPSC-RPE from multiple donors prepared multiple times from iPSC reprogramming to maturation, and the network learned to identify iPSC-RPE from different preparations of the same donor.

Feature Extraction and Traditional Machine Learning

QBAM images segmented using DNN-S were used to extract cell features using the Web Image Processing Pipeline (WIPP) (19). A complete list of features is provided in Table S1. Features were then used as inputs to traditional machine learning algorithms (multi-layer perceptrons, partial least squared regression, linear support vector machines, random forest). Details on preprocessing and training parameters are described in detail in the supplemental methods. Each of the traditional machine learning algorithms were used to predict the same metrics of the deep neural networks (TER, VEGF, donor identity, not segmentation). The advantage of traditional machine learning algorithms is that it is easier to identify specific cellular features that correlate to predicting cell function or identity.

Quantification and Statistical Analysis

All significance between groups indicated for albino lines and for Fig 3 a, d, e were done using a linear mixed effect model controlling for repeated measures from a single well over time and for multiple images being taken per well. These models were assessed using the multcomp and the nlme packages in R. R^2 values, confidence intervals, and Kolmogorov–Smirnov, F-1 and F-2 statistics were calculated in base R. All p-values were set to statistical significance of 0.005 or less.

Study Approval

All iPSC lines were obtained from human subjects in IRB 11-EI-0245 protocol. A total of 15 iPSC-RPE cell lines were used in this paper, obtained from 10 different donors. The iPSC-RPE lines were obtained from three types of patients: healthy, AMD patients, and OCA patients. The iPSC-RPE from healthy patients were derived from iPSC lines Healthy-1 and Healthy-2 respectively. The iPSC-RPE from AMD patients were referred

to in the paper according to donor number and clone number. For example, AMD1A means the cells came from AMD donor #1 and clone A. The different clones for each donor were replicates, where each clone was completely replicated from generation of iPSCs to iPSC-RPE differentiation. The AMD clones were previously reported (25). A summary of the number of clones per donor is as follows: AMD1 had Clone A and Clone B, AMD2 had Clone A, Clone B, and Clone C, and AMD3 had Clone A, Clone B, and Clone C. The iPSC-RPE obtained from OCA patients (also referred to as albino patients), came from five different patients (a single clone each) and indicated by OCA8, OCA26, OCA103, OCA9, and OCA71. All lines were provided via the NEI EYEGene initiative. No personally identifying information about patients was given from the EYEGene bank.

Author Contributions

Conceptualization, NJS, NAH, CGS, PB, and KB; Methodology, NJS, NAH, PM, SP, PB, CGS, and KB; Software, NJS, NAH, PM, SP, JC, MS, MO, and PB; Validation, NJS, NAH, PM, SP, and PB; Formal Analysis, NJS, NAH, PM, SP, JC, and MS; Investigation, NJS, NAH, PM, SP, QW, AG, and RS; Resources, NJS, NAH, PM, SP, QW, AG, RS, JC, MS, PB, CGS, and KB; Data Curation, NJS, NAH, PM, SP, and PB; Writing – Original Draft, NJS and NAH; Writing – Review & Editing, NJS, NAH, CGS, PB, and KB; Visualization, NJS, NAH, and QW; Supervision, CGS, PB, and KB. Nicholas Schaub and Nathan Hotaling discussed who would come first in the authorship order, though both contributed equally, Nick had worked on the project longer and thus was given the first Author position.

Data and Software Availability

All image data used in this study is available online from the National Institute of Standards and Technology at the following address: <https://isg.nist.gov/deepzoomweb/data/RPEimplants>

The QBAM method of collecting images on a microscope was implemented as a Micromanager plugin, and the source code for the plugin is available on Github: <https://github.com/Nicholas-Schaub/SQuIRE>

The QBAM method of converting bright-field microscope images into absorbance values was implemented as an ImageJ plugin, and the source code is available on Github: <https://github.com/Nicholas-Schaub/CARPE>

Acknowledgements

The authors would like to thank the National Eye Institute Electron Microscopy core for help with transmitted electron microscopy micrographs, the NHLBI flow cytometry core for the VEGF secretion measurement, and Dr. Arvydas Maminishkis for his insight in experimental design and imaging. We would like to give a special thanks to Davide Ortolan, Tareq Uddin, Nyusha Lin, Nikhil Etikela, Malika Nimmagadda, Peter Stewart, Helen Zhao, Marissa Coene, Tochi Ugweje, Jun Jeon, and Hannah Bush for their hard work manually segmenting cell borders.

Funding

This work was supported by the National Institute of Standards and Technology under the computational science in metrology project. Nicholas J. Schaub was supported by the National Research Council Research Associateship Program and the NIH TEAM T32 tissue engineering training grant (2T32DE007057-41).

Conflict of Interest Statement

The authors have declared that no conflict of interest exists. Commercial products used in this work are only referenced to specify the experimental procedure adequately. Such identification of commercial products is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the products identified are necessarily the best available for the purpose.

References:

1. Mandai M et al. Autologous Induced Stem-Cell-Derived Retinal Cells for Macular Degeneration. *N. Engl. J. Med.* 2017;376(11):1038–1046.
2. Aijaz A et al. Biomanufacturing for clinically advanced cell therapies. *Nat. Biomed. Eng.* 2018;2(6):362–376.
3. Giancola R, Bonfini T, Iacone A. Cell therapy: cGMP facilities and manufacturing. *Muscles Ligaments Tendons J.* 2012;2(3):243–247.
4. Jones MK, Lu B, Girman S, Wang S. Cell-based therapeutic strategies for replacement and preservation in retinal degenerative diseases. *Prog. Retin. Eye Res.* 2017;58:1–27.
5. Arora S et al. Analysis of Human Retinal Pigment Epithelium (RPE) Morphometry in the Macula of the Normal Aging Eye. *Invest. Ophthalmol. Vis. Sci.* 2013;54(15):2014–2014.
6. Mazzitello KI et al. Druse-Induced Morphology Evolution in Retinal Pigment Epithelium. *ArXiv160904496 Phys. Q-Bio*; <http://arxiv.org/abs/1609.04496>. cited January 9, 2017
7. Rashid A et al. RPE Cell and Sheet Properties in Normal and Diseased Eyes. *Adv. Exp. Med. Biol.* 2016;854:757–763.
8. Cruz L da et al. Phase 1 clinical study of an embryonic stem cell–derived retinal pigment epithelium patch in age-related macular degeneration. *Nat. Biotechnol.* 2018;36(4):328–337.
9. Thumann G, Dou G, Wang Y, Hinton DR. Chapter 16 - Cell Biology of the Retinal Pigment Epithelium A2 - Schachat, Stephen J. RyanSrinivas R. SaddaDavid R. HintonAndrew P. SchachatSrinivas R. SaddaC.P. WilkinsonPeter WiedemannAndrew P.. In: *Retina (Fifth Edition)*. London: W.B. Saunders; 2013:401–414
10. Bonilha VL. Retinal pigment epithelium (RPE) cytoskeleton in vivo and in vitro. *Exp. Eye Res.* 2014;126(Supplement C):38–45.

11. Buggenthin F et al. An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. *BMC Bioinformatics* 2013;14(1):297.
12. Meijering E. Cell Segmentation: 50 Years Down the Road [Life Sciences]. *IEEE Signal Process. Mag.* 2012;29(5):140–145.
13. Litjens G et al. A survey on deep learning in medical image analysis. *Med. Image Anal.* 2017;42:60–88.
14. Guo W, Xu Y, Feng X. DeepMetabolism: A Deep Learning System to Predict Phenotype from Genome Sequencing. *ArXiv170503094 Q-Bio*; <http://arxiv.org/abs/1705.03094>. cited June 27, 2018
15. Abadi M et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *ArXiv160304467 Cs*; <http://arxiv.org/abs/1603.04467>. cited June 27, 2018
16. Awan AA, Hamidouche K, Hashmi JM, Panda DK. S-Caffe: Co-designing MPI Runtimes and Caffe for Scalable Deep Learning on Modern GPU Clusters. In: *Proceedings of the 22Nd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*. New York, NY, USA: ACM; 2017:193–205
17. Wolf L, Hassner T, Maoz I. Face recognition in unconstrained videos with matched background similarity. In: *CVPR 2011*. 2011:529–534
18. Cho Y, Bianchi-Berthouze N, Julier SJ. DeepBreath: Deep learning of breathing patterns for automatic stress recognition using low-cost thermal imaging in unconstrained settings. In: *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. 2017:456–463
19. Bajcsy P, Chalfoun J, Simon M. *Web Microanalysis of Big Image Data*. Springer International Publishing; 2018:

20. Kamao H et al. Objective Evaluation of the Degree of Pigmentation in Human Induced Pluripotent Stem Cell-Derived RPE. *Invest. Ophthalmol. Vis. Sci.* 2014;55(12):8309–8318.
21. May-Simera HL et al. Primary Cilium-Mediated Retinal Pigment Epithelium Maturation Is Disrupted in Ciliopathy Patient Cells. *Cell Rep.* 2018;22(1):189–205.
22. Bhatia SK et al. Analysis of RPE morphometry in human eyes. *Mol. Vis.* 2016;22:898–916.
23. Ach T et al. Quantitative Autofluorescence and Cell Density Maps of the Human Retinal Pigment Epithelium RPE Autofluorescence and Cell Density. *Invest. Ophthalmol. Vis. Sci.* 2014;55(8):4832–4841.
24. Jiang Y et al. Analysis of Mouse RPE Sheet Morphology Gives Discriminatory Categories. *Adv. Exp. Med. Biol.* 2014;801:601–607.
25. Sharma R et al. Clinical-grade stem cell-derived retinal pigment epithelium patch rescues retinal degeneration in rodents and pigs. *Sci. Transl. Med.* 2019;11(475):eaat5580.
26. Isono T et al. Scan-Free Absorbance Spectral Imaging $A(x, y, \lambda)$ of Single Live Algal Cells for Quantifying Absorbance of Cell Suspensions. *PLOS ONE* 2015;10(6):e0128002.
27. Lee JY et al. Absorption-Based Hyperspectral Imaging and Analysis of Single Erythrocytes. *IEEE J. Sel. Top. Quantum Electron.* 2012;18(3):1130–1139.
28. Rieck J, Juuti-Uusitalo K, Autio R, Hytönen V, Skottman H. Quantitative pigment extraction analysis for human pluripotent stem cell derived retinal pigment epithelial cells. In: *EMBECC & NBC 2017*. Springer, Singapore; 2017:61–64
29. Abitan H, Bohr H, Buchhave P. Correction to the Beer-Lambert-Bouguer law for optical absorption. *Appl. Opt.* 2008;47(29):5354–5357.

30. Robinson CJ, Das RG, Stammers R, Rafferty B. The World Health Organization reference reagent for vascular endothelial growth factor, VEGF165. *Growth Factors* 2006;24(4):285–290.
31. Tosoni K, Cassidy D, Kerr B, Land SC, Mehta A. Using Drugs to Probe the Variability of Trans-Epithelial Airway Resistance. *PLoS ONE* 2016;11(2). doi:10.1371/journal.pone.0149550
32. Valm AM et al. Applying systems-level spectral imaging and analysis to reveal the organelle interactome. *Nature* 2017;546(7656):162–167.
33. Miyagishima KJ et al. In Pursuit of Authenticity: Induced Pluripotent Stem Cell-Derived Retinal Pigment Epithelium for Clinical Applications. *Stem Cells Transl. Med.* 2016;5(11):1562–1574.
34. Jansen F et al. Assessment of the repeatability and border-plate effects of the B158/B60 enzyme-linked-immunosorbent assay for the detection of circulating antigens (Ag-ELISA) of *Taenia saginata*. *Vet. Parasitol.* 2016;227:69–72.
35. Kokkinaki M, Sahibzada N, Golestaneh N. Human Induced Pluripotent Stem-Derived Retinal Pigment Epithelium (RPE) Cells Exhibit Ion Transport, Membrane Potential, Polarized Vascular Endothelial Growth Factor Secretion, and Gene Expression Pattern Similar to Native RPE. *Stem Cells* 29(5):825–835.
36. Brandl C et al. In-Depth Characterisation of Retinal Pigment Epithelium (RPE) Cells Derived from Human Induced Pluripotent Stem Cells (hiPSC). *Neuromolecular Med.* 2014;16(3):551–564.
37. Shan J et al. Identification of small molecules for human hepatocyte expansion and iPS differentiation. *Nat. Chem. Biol.* 2013;9(8):514–520.
38. Simm J et al. Repurposed High-Throughput Images Enable Biological Activity Prediction For Drug Discovery. *bioRxiv* 2017;108399.

39. Singh S et al. Morphological Profiles of RNAi-Induced Gene Knockdown Are Highly Reproducible but Dominated by Seed Effects. *PLOS ONE* 2015;10(7):e0131370.
40. Alizadeh E, Merrick Lyons S, Marie Castle J, Prasad A. Measuring systematic changes in invasive cancer cell shape using Zernike moments. *Integr. Biol.* 2016;8(11):1183–1193.
41. Tahmasbi A, Saki F, Shokouhi SB. Classification of benign and malignant masses based on Zernike moments. *Comput. Biol. Med.* 2011;41(8):726–735.
42. Mack AA, Kroboth S, Rajesh D, Wang WB. Generation of Induced Pluripotent Stem Cells from CD34+ Cells across Blood Drawn from Multiple Donors with Non-Integrating Episomal Vectors. *PLOS ONE* 2011;6(11):e27956.
43. Young IT. Shading Correction: Compensation for Illumination and Sensor Inhomogeneities. *Curr. Protoc. Cytom.* 2000;14(1):2.11.1-2.11.12.
44. Janesick JR, Klaasen KP, Elliott T. Charge-Coupled-Device Charge-Collection Efficiency And The Photon-Transfer Technique. *Opt. Eng.* 1987;26(10):261072.
45. Maminishkis A et al. The P2Y2 Receptor Agonist INS37217 Stimulates RPE Fluid Transport In Vitro and Retinal Reattachment in Rat. *Invest. Ophthalmol. Vis. Sci.* 2002;43(11):3555–3566.
46. Peterson WM, Meggyesy C, Yu K, Miller SS. Extracellular ATP Activates Calcium Signaling, Ion, and Fluid Transport in Retinal Pigment Epithelium. *J. Neurosci.* 1997;17(7):2324–2337.
47. Szegedy C et al. Going Deeper with Convolutions. *ArXiv14094842 Cs*; <http://arxiv.org/abs/1409.4842>. cited October 29, 2019

48. Chalfoun J et al. FogBank: a single cell segmentation across multiple cell lines and image modalities. *BMC Bioinformatics* 2014;15:431.
49. Vedaldi A, Lenc K. MatConvNet - Convolutional Neural Networks for MATLAB. *ArXiv14124564 Cs*; <http://arxiv.org/abs/1412.4564>. cited October 29, 2019
50. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: 2015:
51. Kametsky L et al. Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. *Bioinformatics* 2011;27(8):1179–1180.
52. Hall MA. *Correlation-based Feature Selection for Machine Learning*. 1999:
53. Garson GD. Interpreting Neural-network Connection Weights. *AI Expert* 1991;6(4):46–51.
54. Jaderberg M, Vedaldi A, Zisserman A. Speeding up Convolutional Neural Networks with Low Rank Expansions. *ArXiv14053866 Cs*; <http://arxiv.org/abs/1405.3866>. cited October 29, 2019

Figures

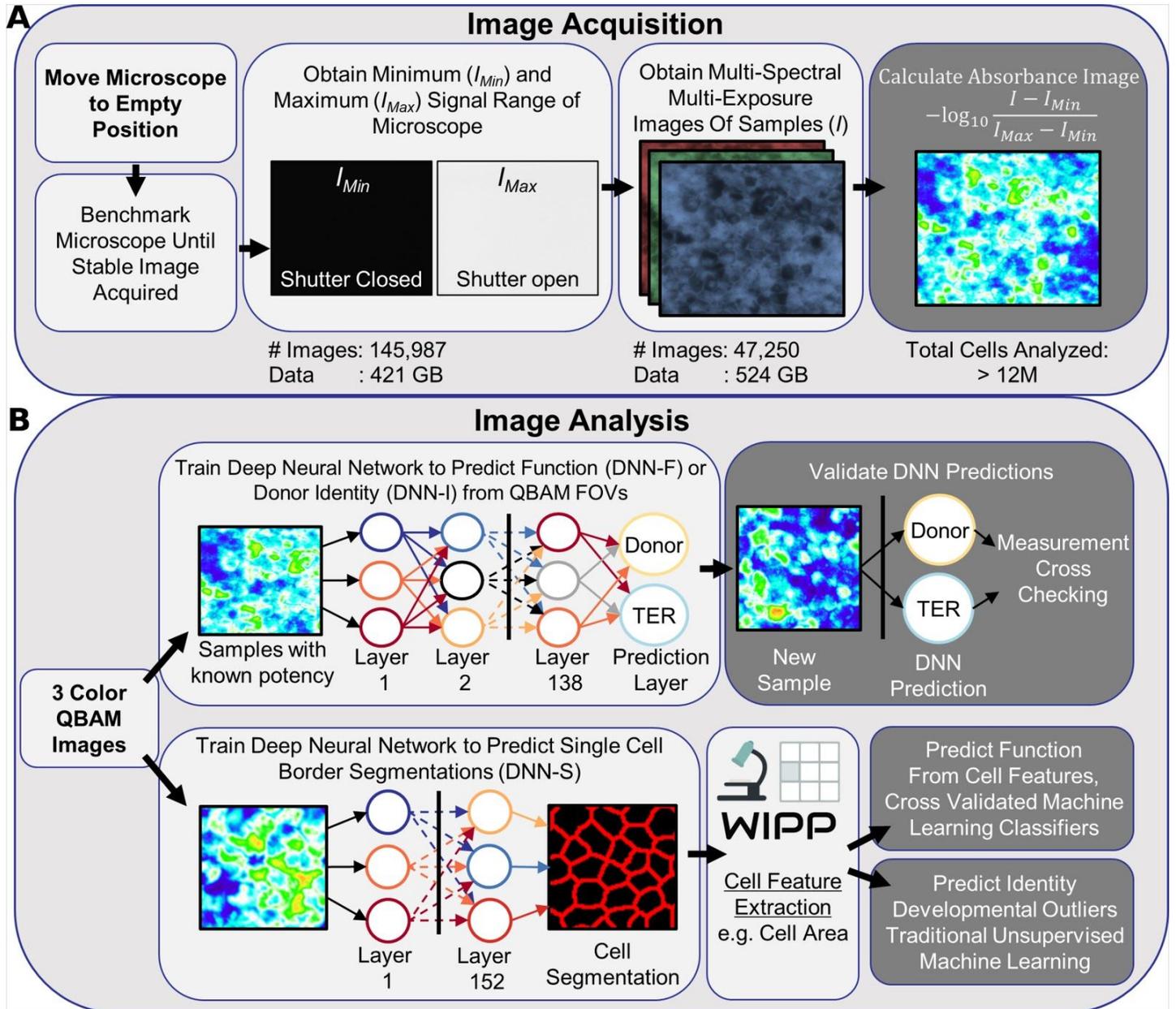


Fig. 1: Methodology for image acquisition and image analysis. (a) An overview of quantitative bright-field absorbance microscopy (QBAM) imaging. The method converts pixel values in brightfield images to absorbance values. (b) Methods of analyzing QBAM images of human iPSC-RPE to predict function, identity and developmental outliers. Three deep neural networks (DNNs) were constructed for this study: (1) Predicts

function (TER and VEGF-Ratio) from QBAM images (DNN-F) using entire fields of view (FOVs) from microscopes, (2) **Identifies** if QBAM images from different clones came from the same donor (DNN-I), (3) **Segments** individual iPSC-RPE cells in absorbance images (DNN-S). QBAM images of iPSC-RPE that were segmented with DNN-S, also had individual cell image features extracted with the web image processing pipeline (WIPP), and these features were used to predict cell function, cell identity and if cells were developmental outliers. More information on QBAM imaging is presented in Figure S1, and the DNN architectures are presented in Figure S2. Features extracted by WIPP are listed in Table S1.

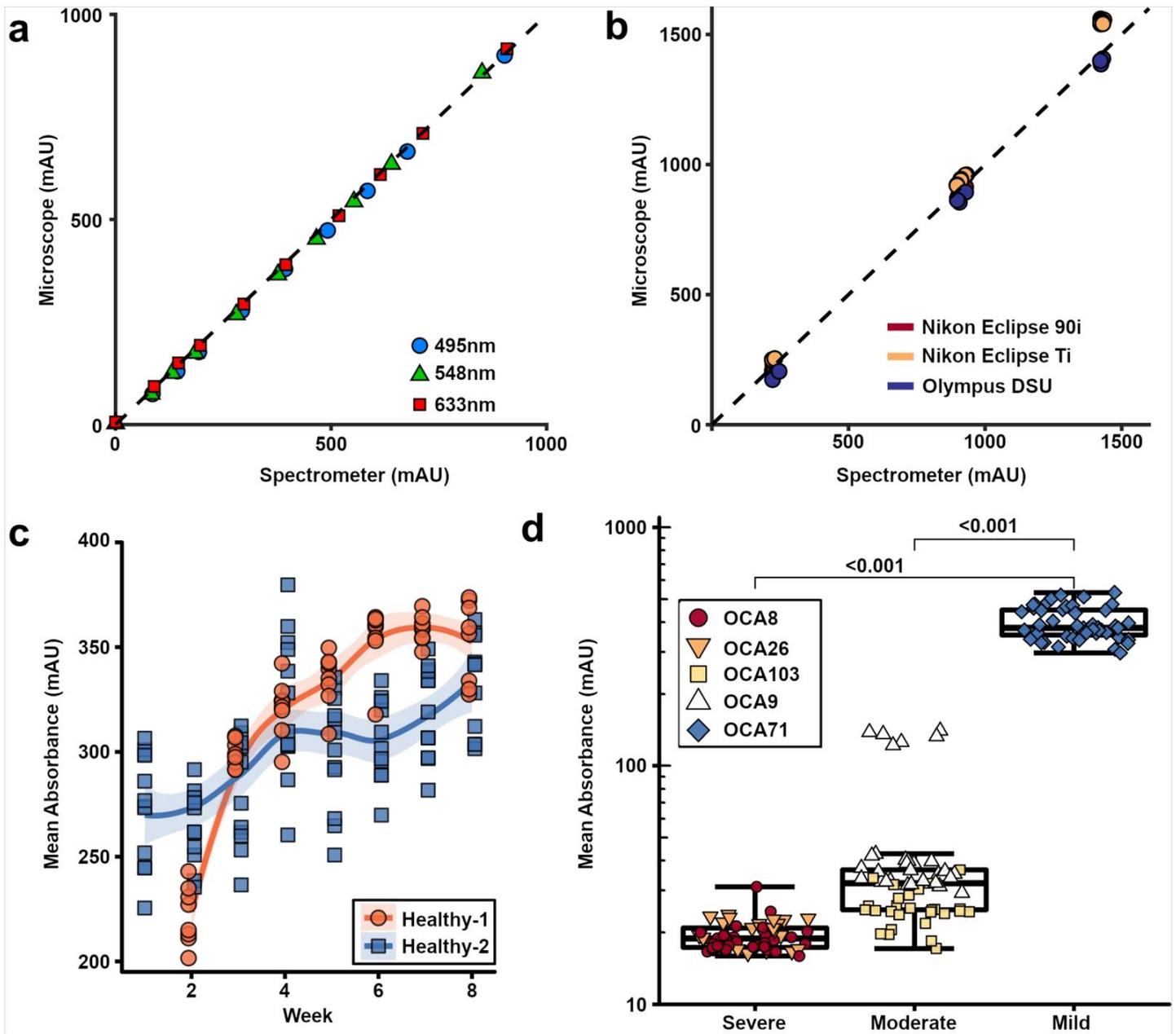


Fig. 2: Assessment of QBAM reproducibility, accuracy, and robustness. (a) ND filters were analyzed with a spectrometer and compared to absorbance values from QBAM images. $n = 3$ replicates per point, error bars - 3SD (smaller than size of data point) (b) Three different ND filters were imaged on three different microscopes using different color filters to determine the comparability of absorbance values between different configurations (e.g. filters, cameras, etc.). $n = 3$ replicates per point, error bars - 3SD (smaller than size of data

point) (c) iPSC-RPE from two healthy patients were imaged over time with QBAM (n=12 wells per donor) to observe changes in pigmentation as iPSC-RPE mature. Each data point represents the mean of 12 images captured of one well. Shaded region represents 95% standard error (d) iPSC-RPE from patients with oculocutaneous albinism (OCA) were imaged to determine if QBAM was able to recapitulate clinical presentation (OCA patients have iPSC-RPE with low pigment). Each data point represents one field of view of each sample. Whiskers represent 3 times the inner quartile range, box limits equal 25% and 75% quantile. n = 9 replicates for Severe, n = 10 replicates for Moderate and n = 8 replicates for Mild. A Linear mixed effect model controlling for multiple images being taken per well was performed for albino cells.

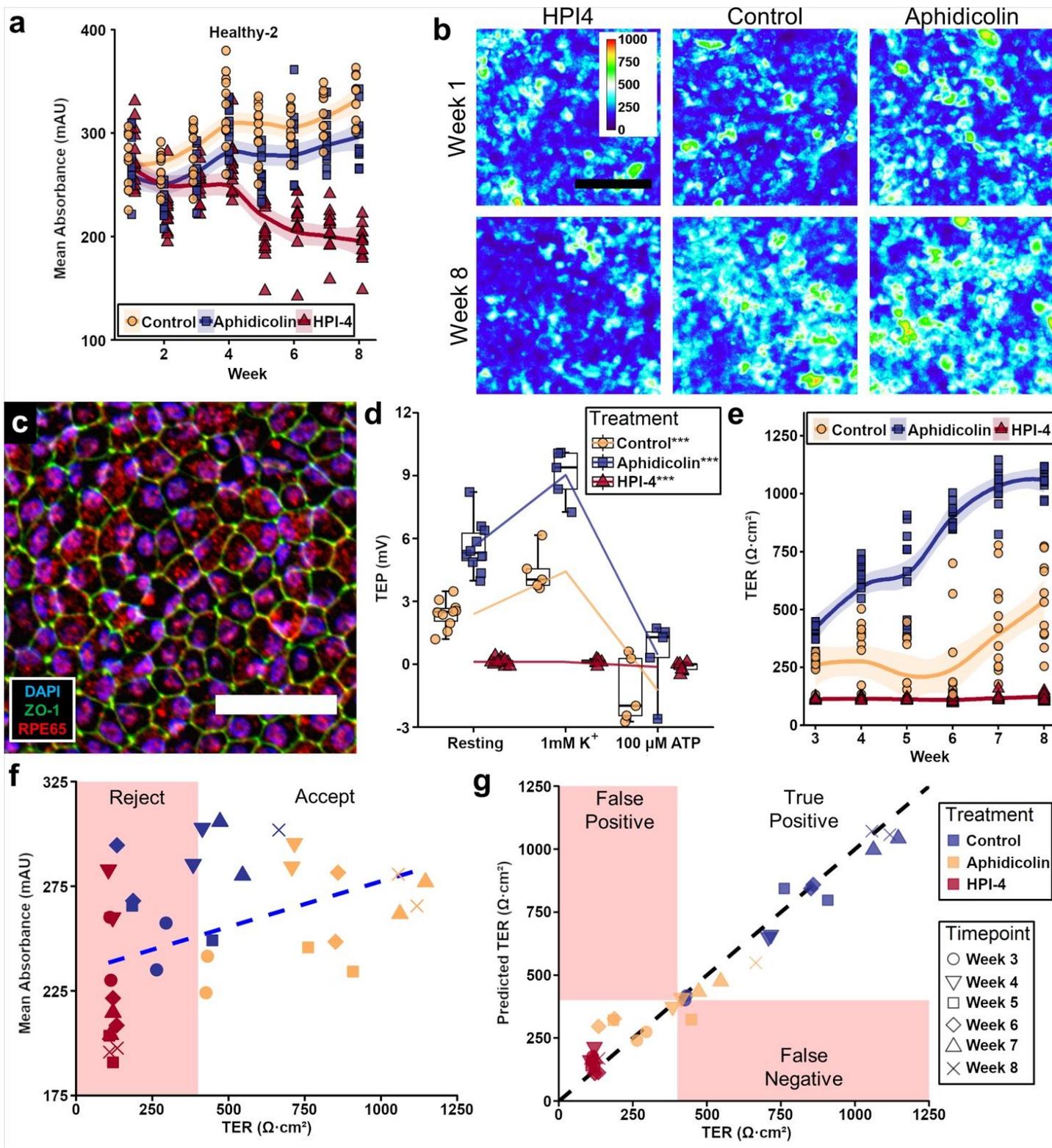


Fig. 3: Prediction of Healthy-2 iPSC-RPE function from QBAM images. (a) A plot of the mean absorbance from 12 images collected in each well over time. Shaded region represents 95% standard error. (b) Representative QBAM images of live iPSC-RPE prior to treatment (Week 1, top row) and after 8 weeks of maturation (bottom row) in the presence of a maturation promoter (Aphidicolin), a maturation inhibitor (HPI4), or neither (Control). (c) Fluorescent labeling of a Control sample from Healthy-2 iPSC-RPE after 8 weeks of culture, where blue is cell nuclei (DAPI), green is cell borders (ZO-1), and red is an RPE specific maturation marker (RPE65). (d) Evaluation of iPSC-RPE transepithelial potential (TEP) in response to an ATP challenge. Asterisks (***) in legend indicate significance $P < 0.005$. Whiskers represent 4 times the inner quartile range and box limits equal the 25% and 75% quantile. (e) A plot of transepithelial resistance (TER) over time for every replicate starting at Week 3. Shaded region represents 95% standard error. (f) A plot of TER against mean image absorbance ($R^2=0.19$, blue dashed line is linear regression). (g) A plot of TER predictions from a deep neural network (DNN-F) against the measured TER ($R^2=0.97$, black line represents a perfect prediction from the DNN). The scale bar in (b) is $100\mu\text{m}$ and the color calibration bar has units of mAU. The scale bar in (c) is $50\mu\text{m}$. See also Figure S3 for additional functional testing, and Figure S4 for DNN-F prediction of VEGF secretion. $n=12$ replicate wells per treatment and 12 images per replicate for all graphs. Linear mixed effect models controlling for repeated measures from a single well over time and for multiple images being taken per well were performed for a, d, and e.

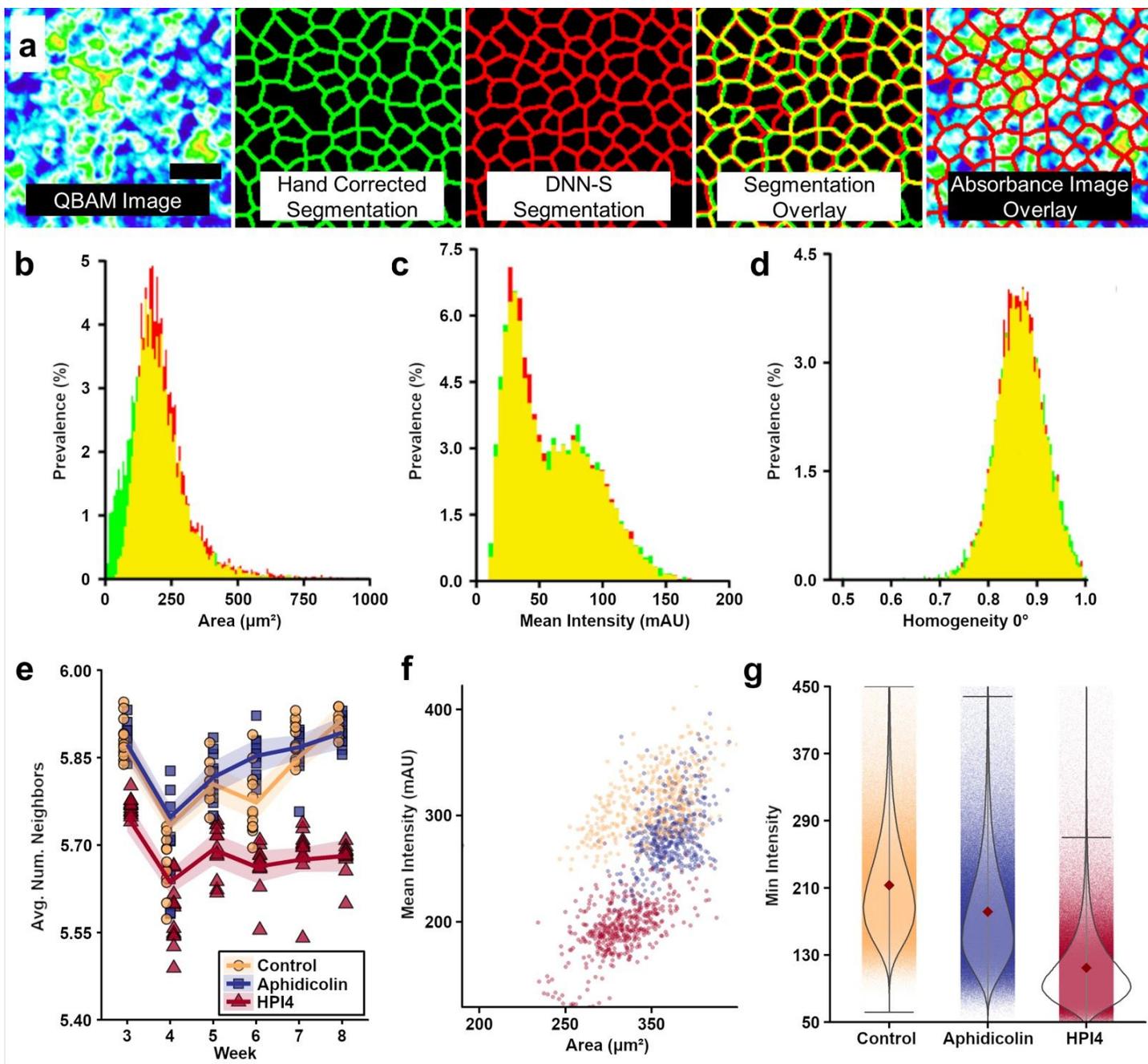


Fig. 4: DNN segmentation of iPSC-RPE in QBAM images. (a) A deep neural network (DNN-S) was constructed to segment iPSC-RPE cells in absorbance images. To train the DNN, iPSC-RPE monolayers were fluorescently labeled for cell borders (ZO-1) and registered to absorbance images for hand labeling of cell borders. Scale bar = 25 μm . (b) A comparison of 3 of 42 cell feature histograms for hand segmented (green)

and DNN segmented (red) images, where yellow is the overlap in the histograms. (c) Time course of the average number of cells bordering each cell for an entire well. Shaded region represents 95% standard error. 12 wells per time per treatment (d) A scatterplot of mean cell area and mean intensity (absorbance) for each treatment group assessed for each microscope field of view. Each dot represents 1 of 864 fields of view (12 wells per treatment, 12 images per well, 6 time points) (e) Minimum intensity (absorbance) found for individual cells as a function of treatment. Whiskers represent three times the standard deviation and Single dots behind the violin represent individual cells measured - $n = 3,871,106$ cells for control, $n = 3,831,362$ cells for Aphidicolin, and $n = 4,146,927$ cells for HPI4. A complete set of feature histogram comparisons is presented in Figure S5, and results of statistical tests are in Table S2. Red diamond indicates the mean.

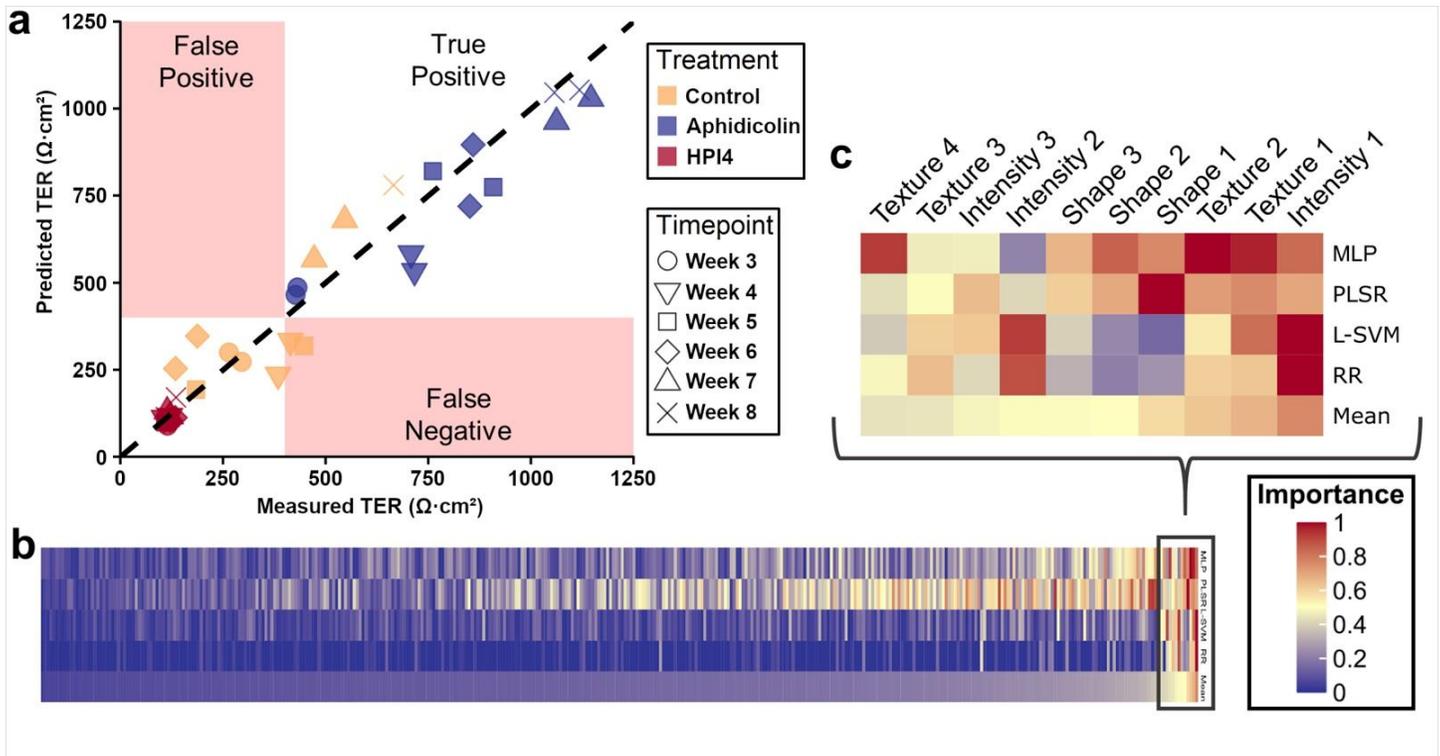


Fig. 5: Traditional Machine learning algorithms' ability to predict iPSC-RPE function (Healthy-2) from cell image features isolated by segmentations of QBAM images. (a) A plot of TER predictions from a multi-layer perceptron (MLP) against the measured TER ($R^2=0.94$, black dashed line represents a perfect prediction from the MLP). Red regions indicate areas that would be less than $400 \Omega \cdot \text{cm}^2$, which was set as the lot release criteria for these cells. $n=12$ replicate wells per treatment and 12 images per replicate. (b) Heatmap of all cell image feature importance sorted by mean importance across all features. (c) Top ten most important features from the heatmap in (b) enlarged so that individual classes of features could be identified. Red indicates most important features to predicting cell TER while blue indicates the least important features. See also Figure S6.

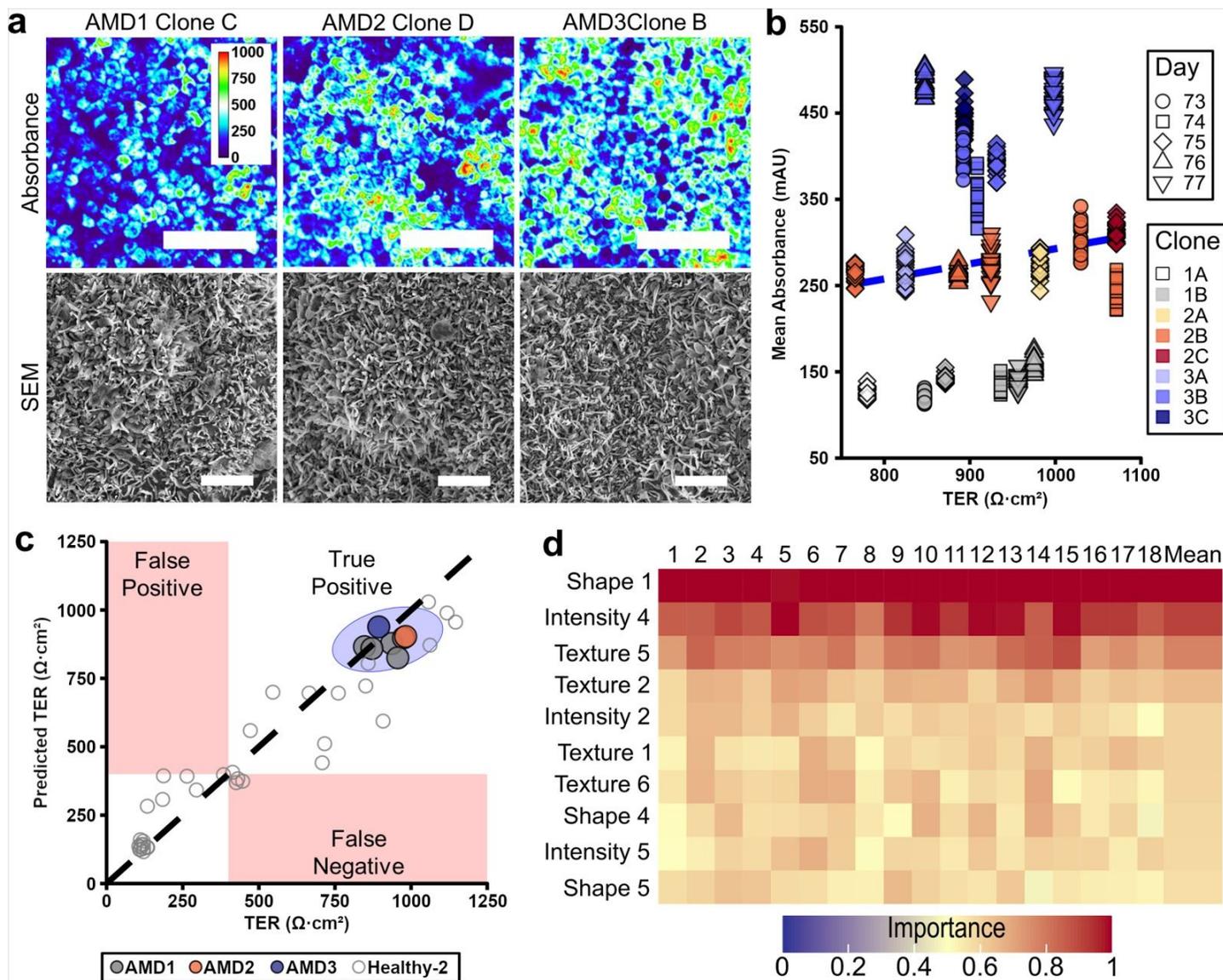


Fig. 6: Prediction of cell TER from eight AMD iPSC-RPE cell lines derived from three separate donors. (a) Representative QBAM image and SEM image of day-75 iPSC-RPE from one clone from each donor. (b) Mean absorbance as assessed by QBAM imaging versus TER for $n=8$ clones (three replicate measures) across the last five days of cell maturation, blue dotted line represents the linear best fit for these cells. (c) Random Forest prediction of iPSC-RPE function across three clones that the algorithm had not seen previously. The black dashed line represents a perfect prediction. A 95% confidence interval is shown as the blue region (d) Heatmap of the important cell image features to predicting iPSC-RPE TER across 18 different clone

combinations sorted by mean feature importance across all clone combinations. The top ten cell image features are shown of 315 analyzed. Red indicates most important features while blue represents least important features. The white scale bars in the top panels of (a) are 100 μ m and the color calibration bar is in milli-absorbance units. The white scale bars in the bottom panels of (a) are 5 μ m. See also Figure S7 and Table S4.

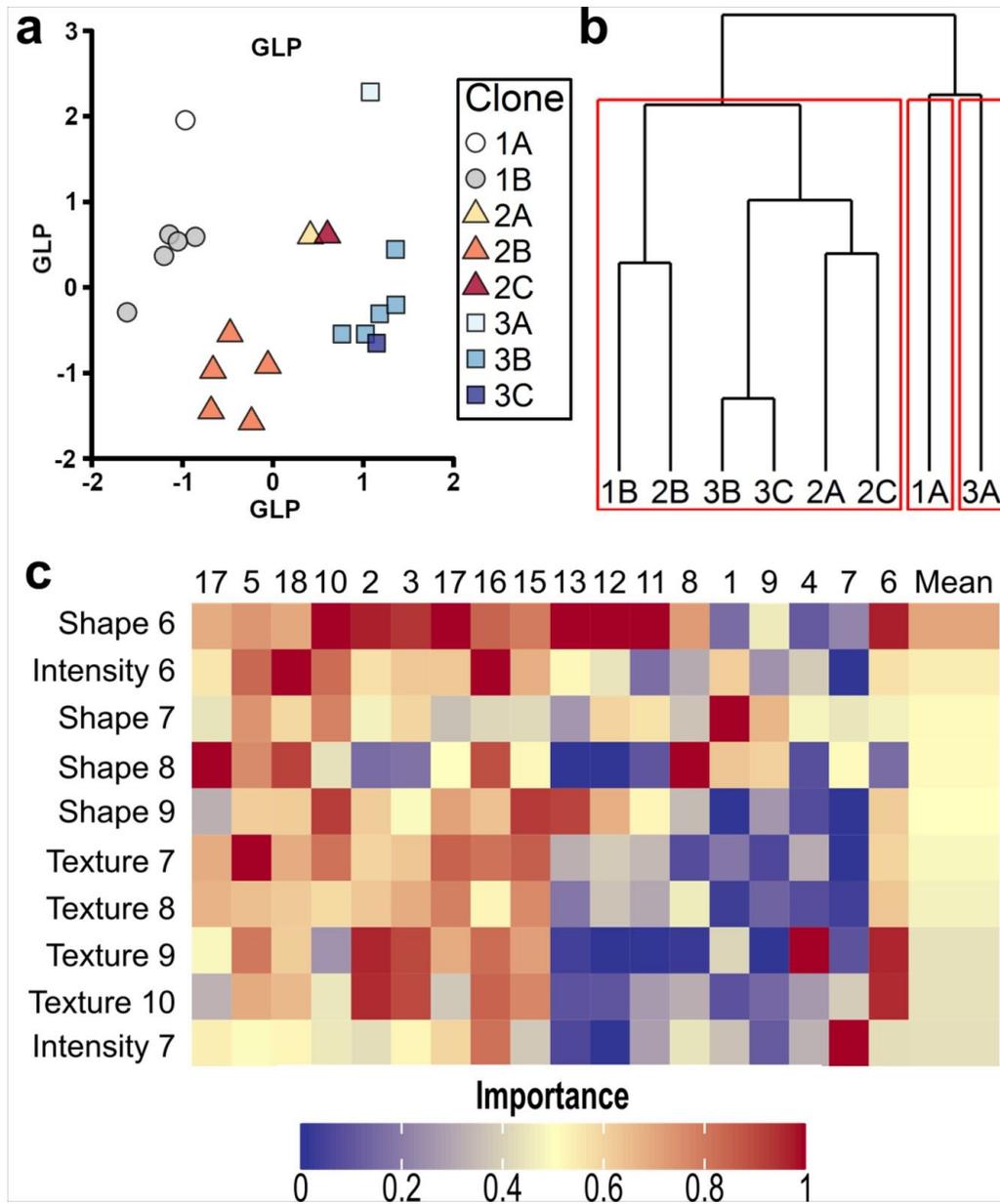


Fig. 7: Identification of iPSC-RPE monolayer developmental outliers and donor identity using only QBAM images and either DNNs or TML algorithms. (a) Principle component analysis of the first principle component versus the second principle component of cell image features from QBAM images of all clones. (b) Hierarchical cluster of all clones to show similarity of clones from three donors, red boxes indicate the three least related groups from the total population. (c) Heatmap of all 18 clone combinations of all cell image features important for the L-SVM to classifying cell identity. The top 10 cell image features are shown (of 315), red indicates the

features most important to predicted cell identity while blue indicates the features least important. n=8 clones with three replicate measures

Tables

Table 1: Analysis of all feature histograms to determine the difference between cell features in hand corrected images compared to DNN segmented images. F-1 and F-2 are pixel level accuracy measurements of the DNN relative to the hand corrected images.

Summary for all Features				
	Error (%)	KSS (P-Value)	F-1	F-2
Mean	7.94	0.31		
St. Dev.	4.42	0.32	0.66	0.71
Min	3.32	<0.001		
Max	18.47	0.996		

Table 2: Root mean squared error (RMSE) of each TML algorithm as compared to actual measures of the cells. DNN-F – deep learning network from Figure 2; PLSR – Principle least-squares regression; L-SVM – Linear support vector machine; RR – Ridge regression; RF – Random Forest.

Summary of Algorithm Regression Error						
	DNN-F	MLP	PLSR	L-SVM	RR	RF
RMSE ($\Omega \cdot \text{cm}^2$)	70.6	84.7	100.1	102.7	109.6	116.4

Table 3: Classifying deep neural network (DNN-I) binary classification performance when determining if two images were from the same donor or not. Columns represent predicted classification while rows represent actual classification.

DNN-I Classification of Donor		
	Donor	Not Donor
Donor	19.8%	9.9%
Not Donor	4.6%	65.6%

Table 4: Linear support vector machine (L-SVM) classification performance when determining which donor a QBAM image of an iPSC-RPE clone was from having never seen images from that clone previously. Columns represent predicted classification while rows represent actual classification.

L-SVM Classification of Donor		
	Donor	Not Donor
Donor	21.5%	11.8%
Not Donor	11.8%	54.9%

Table 5: Summary of the performance of DNN-I and the L-SVM in donor classification. See also Table S5.

	DNN-I	L-SVM
Accuracy	85.4%	76.4%
Sensitivity	80.9%	64.6%
Specificity	86.8%	82.3%