Supplementary Materials for

A human microglia-like cellular model for assessing the effects of neurodegenerative disease gene variants

Katie J. Ryan, Charles C. White, Kruti Patel, Jishu Xu, Marta Olah, Joseph M. Replogle, Michael Frangieh, Maria Cimpean, Phoebe Winn, Allison McHenry, Belinda J. Kaskow, Gail Chan, Nicole Cuerdon, David A. Bennett, Justin D. Boyd, Jaime Imitola, Wassim Elyaman, Philip L. De Jager, Elizabeth M. Bradshaw*

*Corresponding author. Email: emb2280@cumc.columbia.edu

Published 20 December 2017, Sci. Transl. Med. 9, eaai7635 (2017)
DOI: 10.1126/scitranslmed.aai7635

The PDF file includes:

Materials and Methods
Fig. S1. Higher expression of P2RY12 and TMEM119 in MDMi cells compared to monocytes.
Fig. S2. The proapoptotic gene BAX is increased in monocytes and MDM cells compared to MDMi cells.
Fig. S3. Many SNPs have differential eQTLs in MDMi cells compared to monocytes.
Legends for table S1 to S3
References (72–77)

Other Supplementary Material for this manuscript includes the following:
(available at
www.sciencetranslationalmedicine.org/cgi/content/full/9/421/eaai7635/DC1)

Table S1 (Microsoft Excel format). Differentially expressed genes in MDMi cells compared to bulk brain tissue.
Table S2 (Microsoft Excel format). Cis-eQTLs in MDMi cells.
Table S3 (Microsoft Excel format). LD-pruned list of cis-eQTLs in MDMi cells.
Materials and Methods

Quantitative Real Time-Polymerase Chain Reaction (qRT-PCR)

RNA was extracted from each sample using an Absolutely RNA 96 Microprep Kit (Agilent). Genomic DNA contamination was minimized with the addition of DNase to the samples according to the manufacturer’s instructions. RNA was reverse transcribed into cDNA using a Taqman Reverse Transcription kit (Invitrogen). qPCR was performed using TaqMan® Fast Advanced Master Mix (Applied Biosystems) and run on a ViiA 7 System (Applied Biosystems). The cycling conditions consisted of 90 °C for 10 min and 40 cycles of 90 °C for 15 s followed by 60 °C for 1 min. Samples were assayed with 2 technical replicates. mRNA values were normalized relative to B2M by the formula $2^{-\Delta Ct}$, where $\Delta Ct = Ct_{mRNA} \text{-} X \text{–} Ct_{B2M}$.

RNA-Sequencing

Monocytes, MDM and MDMi from 5 healthy subjects were profiled. Extracted total RNA was provided to The Broad Institute's Genomics Platform for quality control and RNA-Seq. Quality control was performed on the BioAnalyzer and all samples were of high quality. The samples were then normalized by RiboGreen for a total input of 250ng for the TruSeq strand specific large insert library construction protocol. This protocol uses poly A selection and some modifications of Illumina's TruSeq protocol to obtain fragment sizes in a range of 450-550bp. The libraries were sequenced on an Illumina HiSeq2000 with 101bp paired end reads to a target coverage of 50M reads per sample.

The RNA-Seq data were processed by our parallelized and automated pipeline. This pipeline includes trimming the beginning and end bases from each read, identifying and
trimming adapter sequences from reads, detecting and removing rRNA reads, and aligning reads to the reference genome. We used the Trinity Package (72) to align reads to the transcriptome reference (hg19 build with gencode v14 annotation) and then applied RSEM (73) to estimate gene expression for all transcripts. The outcome of the RNA-Seq pipeline was fragments per kilobase of transcript per million mapped reads (FPKM) values. FPKM values were plotted in the form of a heatmap with hierarchical clustering.

M1/M2 Macrophage and MDMi Polarization
Monocytes were isolated from frozen PBMCs using CD14 positive microbeads (Miltenyi Biotech). Using a previously described protocol (41, 74), monocyte-derived macrophages were generated by treating for 5 days with human recombinant GM-CSF (5 ng/ml) for M1 polarization or M-CSF (25 ng/ml) for M2 generation. M1 MDM and MDMi were activated for 1 h with IFNγ (20 ng/ml) and 48 h with LPS (100 ng/ml), and M2 MDM and MDMi were activated for 48 h with IL-4 (20 ng/ml) and IL-13 (20 ng/ml).

High-Throughput qPCR (Fluidigm Biomark)
A comprehensive analysis of gene expression in MDMi was performed using Fluidigm’s Biomark high throughput qPCR chip. This system utilizes microfluidic technology to perform high throughput gene expression measurements with real time PCR. Monocytes were plated at a density of 1 x 10^6 cells per well in 6-well plates and differentiated to MDMi for 10 days as described above. RNA was extracted from each MDMi sample using an Absolutely RNA 96 Microprep Kit (Agilent) and equalized to 1 ng/μl. Individual samples and gene expression assays
were pipetted into the microfluidic chip. Thermal cycling was performed, the chip was imaged at the end of each cycle and analysis software generated PCR curves for each of the wells. We examined the expression of 94 genes (plus one housekeeping gene, GAPDH) in MDMi from 96 healthy individuals. One individual was excluded from analysis due to low sample quality, therefore MDMi from 95 individuals were included in the analysis. Data was analyzed using the $2^{-\Delta Ct}$ method, where $\Delta Ct = CtmRNA - CtGAPDH$.

**Fold-change gene analysis**

To determine microglia enriched fold-change genes RNA-seq gene expression data from MDMi, *ex vivo* microglia (39), induced microglia (26) and mouse *ex vivo* microglia (40) were compared to RNA-seq gene expression data from total brain (75). The MDMi and pMGL samples were compared to the mean from 248 individuals without a pathological diagnosis of AD from the Memory and Aging Project and the Religious Orders Study (76, 77). For the *ex vivo* human and mouse microglia, they were compared to paired bulk tissue (39, 40). A cutoff of FPKM> 1 was used and quantile normalization was applied after filtering low expressed and low variation genes. Fold-change genes were scaled to the Z-score and the cutoff threshold was set to the top 5%. The signature datasets were compared.

**Expression Quantitative Trait Locus (eQTL) Analysis**

We performed a cis-eQTL analysis in order to identify genetic variants associated with gene mRNA expression. The Fluidigm Biomark qPCR data was analyzed using the $\Delta\Delta CT$ method normalized to GAPDH. One subject was missing mRNA quantification for >50% of the genes and
was excluded from further analysis. The PCR probes for the gene *MS4A6A* failed to amplify the cDNA, so *MS4A6A* was excluded from further analysis.

The remaining 95 subjects were genotyped genome-wide as part of the PhenoGenetic project using the Illumina Infinium Human OmniExpress Exome BeadChips. Genotype quality control was performed in PLINK (http://pngu.mgh.harvard.edu/~purcell/plink/) to remove genotypes based on Hardy-Weinberg equilibrium, genotype call rate, and heterozygosity. Genotypes were imputed using BEAGLE (version: 3.3.2) and the reference haplotype panels from the 1000 Genomes Project (The 1000 Genomes Project Consortium Phase I Integrated Release Version 3) Utah residents (CEPH) with Northern and Western European ancestry (CEU). We imputed to a minor allele frequency (MAF) >0.01 and ensured that the difference between our MAF and the 1000 Genomes Project MAF for each SNP was less than 0.3. Post-imputation genotypic probabilities were used in our analysis in order to conservatively penalize SNPs based on imputation quality.

We performed our cis-eQTL analysis by calculating the Spearman’s rank correlation and nominal p-value for the association between each gene and all SNPs within 1 Mb of the transcription start site of that gene. We included batch, sex, cell viability and differentiation, and the freezer for PBMC storage prior to MDMi induction as covariates. We used Spearman’s rank correlation because it is more robust to outliers than comparable methods.

To address the multiple-hypothesis testing burden, we adjusted our p-values using the Benjamini and Hochberg False Discovery Rate (FDR) and used a FDR threshold of 5% to find 35 genes with significant eQTL effects. In our targeted analysis of 70 SNPs previously-associated with AD (42), PD (44), and MS (45) by genome-wide association studies which were also located
within 1 Mb of the transcription start sites of our genes of interest, we performed a less stringent, targeted FDR correction because these SNPs have strong prior evidence of functionality. We used R (www.r-project.org) for statistical calculations.

**Comparison of MDMi and Monocyte eQTLs**

We compared our MDMi eQTL results with our previously published monocyte eQTL results (13). In our monocyte eQTL experiment, we profiled CD14+CD16- monocyte gene expression on Affymetrix GeneChip Human Gene ST 1.0 and used the same genotype and imputation strategy to perform an eQTL analysis. Of note, 59 subjects were profiled in both the MDMi analysis (N = 95) and the monocyte analysis (N = 211), limiting our ability to compare their overlap using methods assuming independent samples from a shared population. For each FDR 0.05 significant eQTL in our MDMi analysis, we examined the replication in the monocyte eQTL results.

**Colocalization Analysis**

We performed a Bayesian colocalization analysis (47) to determine whether the AD susceptibility and PILRB eQTL associations in the rs1476679 locus, as well as the LRRK2 eQTL association in the rs76904798 PD locus were driven by a shared causal variant. We compared our MDMi eQTL results with the International Genomics of Alzheimer's Project Stage 1 results (http://www.pasteur-lille.fr/en/recherche/u744/igap/igap_download.php; n=54,162), and the Parkinson's GWAS results as reported in Nalls et. al. (44) (PDgene http://www.pdgene.org/) using the colocalization R package (http://cran.r-project.org/web/packages/coloc/). While
PDgene reports exact \( P \)-values for variants with \( P \)-values < \( 1.0 \times 10^{-5} \), it provides ranges (\( P > 0.05 \), \( P < 0.05, P < 0.0001 \)) for remaining less significant SNPs. For SNPs in the PD comparison with inexact \( P \)-values, we imputed estimated \( P \)-values from a uniform distributions \( U(\alpha, \beta) \) using the lower and upper bounds of the provided ranges as the values of the parameters \( \alpha \) and \( \beta \). We used the default colocalization prior probabilities:

1) Probability that a SNP is an eQTL = 0.0001
2) Probability that a SNP is AD-associated = 0.0001
3) Probability that a SNP is an eQTL and is AD-associated = 0.00001

Bayesian colocalization summarizes the evidence in a region to provide posterior probabilities (PP) of five mutually exclusive hypotheses between the two trait associations:

- \( H_9 \): There exist no eQTL or AD associations;
- \( H_1 \): There exists an eQTL association but no AD association;
- \( H_2 \): There exists an AD association but no eQTL association;
- \( H_3 \): There exist eQTL and AD associations driven by two independent variants;
- \( H_4 \): There exist eQTL and AD associations driven by a shared causal variant.

Bayesian colocalization supports the hypothesis that the AD association and PILRB eQTL are significant but driven by independent variants (PP: \( H_3 = 0.42 \)) rather than a shared variant (PP: \( H_4 = 0.04 \)). It also supports the hypothesis that within the LRRK2 locus, the PD causal variant and
LRRK2 MDMi eQTL are the same variant (PP:H4=0.90), rather than independent variants (PP:H3=0.06).

Unfortunately, we cannot perform a Bayesian colocalization analysis to determine the likelihood of a shared causal variant driving the MDMi and monocyte eQTLs because the colocalization method assumes that the two traits were examined in independent samples from the same ethnic group. This assumption does not hold in the present analysis where 59 subjects were profiled in both the MDMi analysis (N = 95) and the monocyte analysis (N = 211).

**Western Blotting**

Monocytes were isolated from frozen PBMCs using CD14 positive microbeads (Miltenyi Biotech). Cells were lysed in IP buffer (Thermo Scientific) with a protease inhibitor mixture (Roche Diagnostics) and a phosphatase inhibitor mixture (Sigma-Aldrich). After 20 min on ice, cell lysates were centrifuged at 12,000 rpm for 10 min and diluted in electrophoresis sample buffer. Samples were heated at 80°C for 5 min and 20 µg total protein was loaded into each well of an SDS-PAGE gel for separation by electrophoresis. Proteins were transferred on to a PVDF membrane and probed with anti-CD33 rabbit polyclonal IgG antibody (H-110) and goat anti-rabbit HRP-conjugated antibody (Santa Cruz Biotechnologies). Membranes were developed with Immobilon Western Chemiluminescent HRP substrate (Millipore). Bands were quantified by densitometric analysis using ImageJ software (Wayne Rasband, NIH, USA).
Immunocytochemistry

MDMi were cultured for 10 days as described above in optical 96-well plates (Corning 89022-352) at a density of $1 \times 10^5$ cells per well. MDMi and monocytes from the same donors (for the P2RY12 and TMEM119 staining) were fixed in 4% paraformaldehyde (PFA) for 15 min, washed twice in PBS, and blocked with BSA/PBS-T at room temperature for 1 h. Samples were incubated with primary antibodies overnight at 4°C with gentle rocking, washed thrice in PBS, followed by incubation with secondary antibody for 1 h. Samples were incubated with DAPI to stain the nuclei. Staining of MDMi was performed using adherent cells while the monocytes were stained in suspension, and briefly spun down at the end of staining before imaging. The primary antibodies used were: rabbit anti-IBA1 (Wako, 019-19741), rabbit anti-P2RY12 (Sigma, HPA014518) and rabbit anti-TMEM119 (Sigma, HPA051870). The secondary antibodies used were: goat anti-rabbit AF488 IgG and goat anti-rabbit AF555 IgG, both from Life Technologies. The cell preparations were imaged using a Zeiss LSM710 confocal microscope. Images were acquired using ZEN black software.

Dextran Uptake Assay

We tested the uptake ability of MDMi using FITC-labeled dextran (Sigma FD40S). Purified monocytes were differentiated into MDMi for 10 days (as described above) in optical 96-well plates after which they were incubated with 1 mg/ml FITC-labeled dextran in RPMI media for 2 h at 37 °C or 4 °C. Cells were then washed three times with PBS and fixed in 4% PFA for 15 min. Data shown is the change in MFI (delta MFI) between samples incubated at 4 °C (where uptake is blocked) and 37 °C (where uptake occurs).
High Content Analysis

To quantify CD33 expression, live MDMi were incubated with an FC receptor block for 10 min on ice and then labeled with 1:80 PE-conjugated anti-CD33 antibody (Miltenyi Biotech, Clone: AC104.3E3) on ice for 1 h. Cells were fixed with 4% PFA for 15 min, washed with PBS and were labeled with DAPI stain to visualize the nuclei. To quantify FITC-dextran internalization, an uptake assay was carried out as above, following which cells were labeled with Cell Mask (Life Technologies), to visualize the cell membrane, and DAPI. Plates were imaged on the IN Cell Analyzer 2000 (GE Healthcare). CD33 labeled cells were imaged in 6 preselected fields over 2 channels, $\lambda = 350$ nm excitation / $\lambda = 455$ nm emission for DAPI and $\lambda = 579$ nm excitation / $\lambda = 605$ nm CD33-PE at 35 ms and 100 ms exposure times, respectively. For dextran-FITC internalization, fields were selected as described above. Images were acquired over 3 channels, $\lambda = 350$ nm excitation / $\lambda = 455$ nm emission for DAPI, $\lambda = 490$ nm excitation / $\lambda = 525$ nm emission for dextran-FITC, and $\lambda = 645$ nm excitation / $\lambda = 705$ nm emission for Cell Mask far red, at 35 ms, 100 ms, and 100 ms exposure times, respectively. Image stacks were batched and analyzed using IN Cell Workstation software (GE Healthcare). For the CD33 feature extraction protocol, cells were segmented using the Multi Target Analysis algorithm, beginning by segmenting nuclei using a top hat algorithm with a minimum area of 50 $\mu m^2$ with a sensitivity setting of 60. Cells were defined by a multiscale top-hat algorithm with a 100 $\mu m^2$ minimum area and sensitivity setting at 71 on the Cy3 channel. The CD33 signal (“organelle”) was segmented using a multiscale top-hat algorithm with a granule size between 1 and 3 pixels, scale 1, and a sensitivity setting of 78. Features segmentation was limited only to cytoplasmic
compartment. Summary values included: cell count, nuclei area, cell area, cell elongation, background intensity, organelle count, organelle intensity, and organelle total area. To quantify the dextran uptake, cells were segmented using the Multi Target Analysis, under the same nuclei segmentation protocol as above. Additionally, cells were segmented using the Cy5 channel, multiscale top hat algorithm with a 250 µm² minimum area and sensitivity setting of 50. Dextran-FITC (“organelle”) was segmented using the multiscale top hat algorithm with a granule size between 1 to 7 pixels, scale 3, and sensitivity setting of 70. Dextran-FITC signal was segmented within the whole cell. Summary values included: cell count, nuclei area, nuclei intensity, organelle count, organelle intensity, and organelle area.
Fig. S1. Higher expression of P2RY12 and TMEM119 in MDMi cells compared to monocytes. 

(A) MDMi have increased RNA expression of P2RY12 compared to MDM as determined by RNA-sequencing for 5 individuals (\(**P = 0.007\)). (B&C) MDMi have higher expression of both P2RY12 and TMEM119 at the protein level as measured for 50 cells of 4 different individuals compared
to monocytes (****$P < 0.0001$). (D) Additional images of P2RY12 and TMEM119 staining in monocytes and paired MDMi. Student’s t-test.
Fig. S2. The proapoptotic gene *BAX* is increased in monocytes and MDM cells compared to MDMi cells. After differentiation, both MDM and MDMi from 6 healthy individuals were cultured for an additional 15 days in RPMI media, at day 30 expression of BAX was measured. The MDM have significantly more BAX expression than the MDMi (***P = 0.005). Also, *ex vivo*
monocytes have higher expression of BAX compared to MDMi (****P < 0.0001), demonstrating that during the conversion from monocytes, a short-lived cell type, to MDMi the cells obtain a long-lived phenotype, as do *in vivo* microglia. Student’s t-test.
**VCAM1**

Fig. S3. Many SNPs have differential eQTLs in MDMi cells compared to monocytes. The effect size for each SNP (rho) is plotted for MDMi (y-axis) and monocytes (x-axis) for each gene. SNPs that have a significant eQTL for MDMi, monocytes or both are highlighted. See separate File for Fig. S3.
Table S1. Differentially expressed genes in MDMi cells compared to bulk brain tissue.

See separate Excel File for Table S1.

Table S2. Cis-eQTLs in MDMi cells.

See separate Excel File for Table S2.

Table S3. LD-pruned list of cis-eQTLs in MDMi cells.

See separate Excel File for Table S3.