Precision Health Resource of Control iPSC Lines for Versatile Multilineage Differentiation

Matthew R. Hildebrandt,1,15 Miriam S. Reuter,2,3,15 Wei Wei,1 Naeimeh Tayebi,2 Jiajie Liu,1 Sazia Sharmin,1 Jaap Mulder,1 L. Stephen Lesperance,4 Patrick M. Brauer,5 Rebecca S.F. Mok,1,6 Caroline Kinnear,2 Alina Piekna,1 Asli Romm,1 Jennifer Howe,2,3 Peter Pasceri,1 Guoliang Meng,1 Matthew Rozycki,1 Deivid C. Rodrigues,1 Elisa C. Martinez,5 Michael J. Szego,7,8,9,10 Juan C. Zúñiga-Pflücker,5 Michele K. Anderson,5 Steven A. Prescott,4,11,12 Norman D. Rosenblum,1,13 Binita M. Kamath,1,13 Seema Mital,2,3 Stephen W. Scherer,2,3,6,14,* and James Ellis1,6,*

1Developmental and Stem Cell Biology, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada  
2Genetics and Genome Biology, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada  
3The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada  
4Neurosciences and Mental Health, The Hospital for Sick Children, Toronto, ON M5G 0A4, Canada  
5Department of Immunology, University of Toronto, Sunnybrook Research Institute, Toronto, ON M4N 3M5, Canada  
6Department of Molecular Genetics, University of Toronto, Toronto, ON M5S 1A8, Canada  
7Dallal Lana School of Public Health, University of Toronto, Toronto, ON M5T 3M7, Canada  
8Department of Family and Community Medicine, University of Toronto, Toronto, ON M5C 2T2, Canada  
9The Joint Centre for Bioethics, University of Toronto, Toronto, ON, Canada  
10Unity Health Toronto, Toronto, ON M5T 3M6, Canada  
11Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON M5S 3G9, Canada  
12Department of Physiology, University of Toronto, Toronto, ON M5S 1A8, Canada  
13Department of Pediatrics, University of Toronto, Toronto, ON M5G 1X8, Canada  
14McLaughlin Centre, University of Toronto, Toronto, ON M5G 0A4, Canada  
15Co-first author  
*Correspondence: stephen.scherer@sickkids.ca (S.W.S.), jellis@sickkids.ca (J.E.)  
https://doi.org/10.1016/j.stemcr.2019.11.003

SUMMARY

Induced pluripotent stem cells (iPSC) derived from healthy individuals are important controls for disease-modeling studies. Here we apply precision health to create a high-quality resource of control iPSCs. Footprint-free lines were reprogrammed from four volunteers of the Personal Genome Project Canada (PGPC). Multilineage-directed differentiation efficiently produced functional cortical neurons, cardiomyocytes and hepatocytes. Pilot users demonstrated versatility by generating kidney organoids, T lymphocytes, and sensory neurons. A frameshift knockout was introduced into MYBPC3 and these cardiomyocytes exhibited the expected hypertrophic phenotype. Whole-genome sequencing-based annotation of PGPC lines revealed on average 20 coding variants. Importantly, nearly all annotated PGPC and HipSci lines harbored at least one pre-existing or acquired variant with cardiac, neurological, or other disease associations. Overall, PGPC lines were efficiently differentiated by multiple users into cells from six tissues for disease modeling, and variant-preferred healthy control lines were identified for specific disease settings.

INTRODUCTION

The development of induced pluripotent stem cells (iPSC) led to rapid development of many stem cell-based models of disease (Takahashi and Yamanaka, 2016). Despite exponential growth in the application of iPSCs across multiple tissue- and organ-based systems, there remains no consistent consensus about which control lines should be used in disease-modeling studies. Over the past decade, choices for control cells have ranged from: (1) human embryonic stem cells (hESCs) that are considered healthy despite a medical history being unavailable, (2) iPSCs from healthy but unrelated individuals (Schwartzentruber et al., 2018), (3) iPSCs from unaffected family members who may have been phenotyped for the disease of interest, but with unknown broader health profile (Lan et al., 2013), and (4) isogenic pairs of iPSC lines derived through CRISPR-Cas9 gene editing (Deneault et al., 2018), or through non-randome X chromosome inactivation status in female cells (Tchieu et al., 2010). Hundreds of sources of unrelated and related healthy iPSC lines exist and are widely available from individual labs, biobanks, and large iPSC-focused consortia, such as HipSci (Streeter et al., 2017). Although there are genetically diverse lines to reflect heterogeneity found within the human population, all control lines are potentially compromised by genetic variants that may predispose to a phenotype or mask it (Hollingsworth et al., 2017). At present, disease modeling has focused on penetrant monogenic disorders that may be relatively unaffected by the presence of concurrent variants. However, we anticipate an emerging need for healthy controls with few disease variants as modeling of complex diseases builds toward assessing the impact of modifier genes or multigenic disorders that may involve
Figure 1. Active Neurons Generated from PGPC iPSCs Display Similar Dendrite Morphology and Network Circuitry

(A) Differentiation scheme to generate excitatory cortical neurons by induction of Ngn2. Transduced iPSCs were dissociated to single cells and plated in the presence of ROCK inhibitor. At D1, medium was changed to CM1 and Ngn2 was induced by incubating with doxycycline until D7. Puromycin was added from D2 to D4 to remove any non-transduced cells. Culture medium was changed to CM2 on D3. Ara-C was added from D6-8 to remove any remaining dividing cells. On D8, neurons were re-seeded for downstream assays in BrainPhys medium.

(legend continued on next page)
multiple variants including noncoding variants in gene regulatory regions.

iPSCs carry additional variants compared with donor sequences (D’Antonio et al., 2018; Gore et al., 2011). This has made apparent the need for whole-genome sequencing (WGS) to identify the full set of potential disease-susceptibility variants present in such control lines (D’Antonio et al., 2018; Kilpinen et al., 2017; Popp et al., 2018). Although there are some common reprogramming-associated variants (Yoshihara et al., 2017), most variants appear to be present in the original mosaic source of cells reprogrammed (Abyzov et al., 2017). Some of these variants could affect downstream differentiations and baseline phenotypes of differentiated lineages (Hoekstra et al., 2017). Furthermore, most control lines are recruited for specific studies limited to a single tissue type or disease, and therefore their versatility for multilineage-directed differentiation into many functional cell types required for broad disease modeling research is not firmly established.

One way to limit the presence of potentially confounding variants is to reprogram cells from selected donors who have minimal variant load. In both the initial Personal Genome Project (PGP) and Personal Genome Project Canada (PGPC) publications, one aim was to generate iPSCs that would have extensive genomic characterization (Ball et al., 2012; Reuter et al., 2018). PGPC genotyped and clinically annotated the genomes of 56 apparently healthy individuals who consented to disclosure of their genome sequence and medical traits (Reuter et al., 2018). In addition to comprehensive annotation of all classes of constitutional genetic variants, these analyses also included their assessment of the mitochondrial genomes and their pharmacogenetic diplotype. All healthy PGPC individuals harbor heterozygous variants of unknown significance in disease-relevant genes, but still had no overt disease phenotype at the time of initial assessment or at the start of this study. Here we report the iPSC resource generated from PGPC donors.

Our resource comprises multiple iPSC lines derived from two male and two female donors. One line each from both males and one female was subjected to multilineage-directed differentiation into cortical neurons, cardiomyocytes (CMs), and hepatocytes representative of the three germ layers. The morphology and function of the resulting cells were evaluated to assess the versatility of PGPC iPSC lines for in vitro studies of different tissues. To further evaluate the versatility of the resource, we shared the three best-characterized PGPC lines with pilot users for differentiation into kidney organoids, T lymphocytes, and sensory neurons. CRISPR gene editing of a known cardiomyopathy gene created an isogenic pair of lines for modeling a cardiac disorder. As variant annotation of the donors became available (Reuter et al., 2018), we performed WGS to search for iPSC line-specific variants that were distinct from donor PGPC blood variants, and surveyed off-target mutations in the gene edited line.

Results

Isolation and Pluripotency Characterization of PGPC iPSC Lines

We invited PGPC donors to participate in this iPSC study, and selected two male (PGPC3 and PGPC17) and two female donors (PGPC14 and PGPC1) (Reuter et al., 2018). We collected peripheral blood to isolate and reprogram CD34+ cells using non-integrating Sendai viruses. Approximately 120 clones from each donor were picked and qualititative metrics (colony morphology and low levels of spontaneously differentiated cells) were used to select lines for characterization. iPSC lines were maintained in feeder-free conditions and tested for Sendai virus clearance at passage (P)8 to 10. Sendai virus-negative lines were sent for karyotyping between P13 and P15. At least four karyotypically normal cell lines were found from each donor, with standard characterization results summarized in Table S1 and representative data shown in Figure S1. All cell lines stained positive for both cell surface (SSEA4 and TRA-1-60) and nuclear (OCT4 and NANOG) undifferentiated markers (Figure S1). We tested functional pluripotency by spontaneously differentiating embryoid bodies followed by staining for markers of all three germ layers—ectoderm (TUBB3), mesoderm (SMA), and endoderm (AFP) (Figure S1). All female lines had skewed X chromosome inactivation as revealed by androgen receptor assays.

(B) Representative immunocytochemistry image of iPSC-derived neurons after 6 weeks in culture labeled with DAPI and MAP2 and sparsely labeled with GFP (independent experiments = 2; technical replicates = 20). Scale bar represents 100 μm. Color channels were independently altered to adjust contrast for publication.

(C–E) Plots of (C) soma area, (D) total dendrite length, and (E) number of intersections (Sholl analysis) from 6-week-old neurons (independent experiments n = 2; technical replicates per batch = 20). (C and D) Boxplots indicate median values. (E) Mean values were plotted with error bars indicating SE. Statistically significant pairwise comparisons indicated by * are inset.

(F and G) (F) Representative raster plots of PGPC3 neurons from a single-well of recordings collected by micro-electrode arrays at different time points (two independent experiments each with eight technical replicates). Each spike is indicated by a black line, blue lines represent bursts defined by at least five spikes each separated by an inter-spike interval of no more than 100 ms. (G) Bursts are absent after treatment of week 7 neurons with CNQX (left). Bursts begin to return after compound removal and replacement with fresh basal medium (right).
consistent with preservation of an inactive X chromosome observed in isogenic female lines (Figure S1). These data confirm basic pluripotency status of our resource and cells were expanded and banked at passages ranging from P14 to P16.

We chose to focus on one cell line from the first three donors for deeper characterization as PGPC1 was recruited much later. PGPC3_75, PGPC14_26, and PGPC17_11 were selected for further phenotyping based on qualitative metrics regarding their growth rate, morphology, and relative low rate of spontaneous differentiation. RNA sequencing was analyzed online using Pluritest, and all lines cluster to the pluripotency quadrant (Figure S1). As explained in detail below, we validated the pluripotency and explored the versatility of all three lines for multilineage-directed differentiation to excitatory cortical neurons, CMs, and hepatocytes as representatives of cells derived from ectoderm, mesoderm, and endoderm respectively.

At this point the WGS data of all the PGPC participants became available and were annotated for coding variants defined by the American College of Medical Genetics (Richards et al., 2015). Two heterozygous variants of uncertain clinical significance (VUS) associated with electrophysiological alterations in cardiac disease (Table S2) were identified in PGPC3 (TRPM4) and PGPC14 (KCNE2), respectively. VUS that could affect neurologic function were found in PGPC1 female lines available only with variant annotation (Table S2) and pluripotency characterization as part of the resource.

**Ectodermal Differentiation into Active Cortical Neurons**

To evaluate PGPC iPSC-derived neurons, we infected PGPC lines and a previously published control iPSC line (WT37) (Cheung et al., 2011) with lentivirus bearing doxycycline-inducible Ngn2 to generate homogeneous populations of excitatory cortical neurons (Zhang et al., 2013). Neurons were induced with doxycycline for 1 week and selected for further phenotyping based on their pre-existing variants. The newest PGPC1 female lines are available only with variant annotation (Table S2) and pluripotency characterization as part of the resource.

Intracellular Ca$^{2+}$ transients were measured by loading D31 and D34 CMs with Fluo-4 AM dye. Fluorescence intensity ratios were plotted against time to calculate the Ca$^{2+}$ transient amplitude and rate (Figures 2D–2F). All three PGPC CMs had similar average beat rates and amplitudes. To measure contractility of PGPC17_11 in a complementary method and to determine extracellular electrophysiology, an xCELLigence Real-Time Cell Analysis (RTCA) Cardio ExtraCellular Recording (ECR) system was used. In brief, contracting CMs were recorded every 3 h for ~25 days after re-seeding (Figure 2G). Contractility of CMs was evaluated via impedance readouts as beats per minute (bpm) and beating amplitude (BAmp) defined as the cell index value between lowest and highest

Stem Cell Reports | Vol. 13 | 1126–1141 | December 10, 2019 | 1129
Figure 2. PGPC iPSCs Differentiate to Beating Contractile Cardiomyocytes
(A) Differentiation scheme to generate CMs using STEMdiff Cardiomyocyte Differentiation Kit. iPSCs were dissociated to single cells, plated in 12-well plates, and allowed to reach 85%–90% confluency before beginning differentiation.
(B) D16 CMs were dissociated to single cells for reseeding and a proportion was labeled with anti-cTNT-fluorescein isothiocyanate and subjected to flow cytometry (independent experiments ≥ 3).
(C) Representative images of immunocytochemistry staining of D30 PGPC17 CMs labeled with DAPI, anti-MLC2V (both), and anti-cTNT (left) or anti-α-actinin (right) (independent experiments = 2). Scale bars represent 100 μm. Color channels were independently altered to adjust contrast for publication.
(D) Representative traces of spontaneous Ca²⁺ transients of PGPC CMs at D31 measured by relative fluorescence intensity (independent experiments n ≥ 3; technical replicates per batch ≥ 2).
(E and F) Plots of (E) beat rate and (F) Ca²⁺ transient amplitudes.
(G) Representative xCELLigence data of D40 PGPC17 CMs showing impedance changes (BAmp: defined as the cell index value between lowest and highest points within a beat waveform) reflecting CM beat waveform and absolute extracellular voltage tracings over a 20-s recording (independent experiments = 3; technical replicates ≥ 3).
Figure 3. Enzymatically Active HLCs Are Generated from PGPC iPSCs
(A) Hepatocyte-like cell differentiation scheme. iPSCs were dissociated to single cells and maintained in ROCK inhibitor for 24 h to support survival. From D1 to D4 cells are transferred to STEMdiff definitive endoderm (DE) differentiation kit. From D5 to D9 cells were switched to serum-free differentiation (SFD)-based medium with activin A for 4 days and basic fibroblast growth factor (bFGF) for 2 days adding B27.
points within a beat waveform. Beat rate averaged 36 bpm (range 32–49 bpm) with average amplitude 0.04 a.u. (range 0.027–0.05 a.u.). Extracellular field potential spike amplitudes defined as the difference between the lowest and highest recorded voltages ranged from 0.12 to 0.55 mV. These experiments demonstrate differentiation of three PGPC lines into beating CMs and highlight the potential value of using PGPC17 for CRISPR gene editing for cardiac disease modeling.

**Endodermal Differentiation into Enzymatically Active Hepatocytes**

For endodermal differentiation we generated hepatocyte-like cells (HLCs) (Figure 3A). Differentiated cells were characterized at multiple stages to monitor quality and efficiency. At D4, over 95% of cells co-expressed definitive endoderm (DE) markers CXCR4 and cKIT (data not shown). DE cells were induced to generate foregut (FG) progenitors as indicated with the increase in FG markers FOXA2 and GATA6 (normalized to iPSCs) compared with DE (Figure 3B). FG progenitors were further specified to hepatoblasts (HBs) followed by maturation to HLCs by D25 where clear upregulation of respective mRNAs was assessed by qPCR (normalized to fetal liver) (Figure 3B). Over 95% of HLCs tested positive via flow cytometry for hepatocyte markers including albumin (ALB), alpha fetoprotein (AFP), alpha-1-antitrypsin (A1AT), and CYP3A7 (Figure 3C), and further supported by immunostaining for AFP, ALB, CYP3A7, and HNF4A (Figure 3D). Measuring functional activity of HBs (D14) and HLCs (D25) was performed using a p450-glo assay (Promega). As expected, HLCs had significantly more enzymatic activity of CYP3A7 as measured by luminescence as compared with HBs. Treatment with 1 μM ketoconazole inhibited enzymatic activity of CYP3A7 to levels observed in HBs (Figure 3E). These results demonstrate differentiation of the three PGPC lines into hepatocytes that produce active enzymes.

**Utility of the Resource—Mesodermal Differentiation into Kidney Organoids and T Cells**

To test the utility of PGPC lines as a resource, we made them available to pilot users. Unlike monolayer differentiations described above, human kidney organoids are 3D structures generated from iPSCs consisting of multiple cell types and resembling early embryonic human kidney tissue (Takasato et al., 2015) (Figure S2A).

This protocol entailed a 7-day monolayer culture with directed differentiation toward posterior streak mesoderm (PSM) and subsequently to anterior and posterior intermediate mesoderm (AIM and PIM, respectively). This was accomplished by applying the canonical WNT-signaling activator, CHIR99021 (CHIR), followed by a switch to fibroblast growth factor 9 (FGF9) and heparin. Timing of the FGF9/heparin switch (between D3 and D5 of differentiation) determined the relative proportion of AIM versus PIM and thus fewer or more nephrons. For all experiments, we made this factor switch on D5. During this course, the PSM marker T (brachyury) was transiently induced followed by AIM marker GATA3 and PIM marker HOXD11 as measured by qPCR and expression was normalized to iPSCs at D0 (Figure 4A). After 7 days of monolayer differentiation cells were aggregated, transferred to Transwell membranes, pulsed with CHIR, and then treated with FGF9/heparin for an additional 5 days. Aggregates began reorganization and formed nephron-like structures. mRNA level analysis of D25 organoids (normalized to D7 pre-aggregated cells) showed induction of markers of different nephron segments, endothelial, and stromal cells (Figure 4A). Immunofluorescence imaging of D18 cross-sections of organoids showed glomerular structures—positive for podocyte marker Wilms tumor 1—as well as tubular structures, both proximal—labeled with lectin (LTL)—and distal—positive for E-cadherin (Figure 4B). These results show that 3D kidney organoid structures are produced by the PGPC lines.
Figure 4. Generation of Kidney Organoids and T Lymphocytes

(A) Heatmaps indicating log2 fold change of marker gene expression normalized to iPSCs (left) or D7 (right) (independent experiments [PGPC3/17 = 1, PGPC14 ≥ 2]; technical replicates ≥ 3).

(B) WT1, LTL, SCAD

(C) Percentage of CD34+ (preMACS) cells

(D) T-cell Differentiation

E) Culture day 8 + 14

F) Culture day 8 + 42

(legend continued on next page)
To evaluate the potential to generate hematopoietic stem/progenitor cells (HSPCs) and mature T cells, we compared the PGPC lines with iPSC11 (Alstem Cell Advancements) in an embryoid body differentiation protocol with feeder-free adaptation (Figure S2B). All three PGPC lines gave rise to CD34+ HSPCs with a similar proficiency as iPSC11 cells (Figure 4C) and were magnetic-activated cell sorted at D8 (Figure S2C). PGPC14 and 17 were most enriched for CD34+ HSPCs and cocultured on OP9-DL4 cells. Next, multi-color flow cytometry was used to simultaneously measure different cell populations at D8+14 and D8+42 to assess the ability of HSPCs to differentiate to T lymphocytes, a hallmark of definitive hematopoietic potential of HSPCs (Figure 4D). At D8+14 (shown in Figure 4E), early T-lineage progenitor cells (proT cells marked as CD34+ CD7+) could be observed transitioning to more developmentally matured T-lineage cells (CD34− CD7+), with a subpopulation co-expressing a pan-T cell marker (CD7+ CD5+). PGPC14 exhibited a prolonged proT stage (70%), while PGPC17 showed a more rapid transition (17%) compared with iPSC11 (53%). Simultaneous assessment of mature T-lineage markers on culture D8+42 detected the presence of double-positive (DP) CD4+ CD8+ T-lineage cells, T cell receptors (TCR+) and a T cell-specific marker (CD3+). At this time point, the PGPC lines showed similar propensity as iPSC11 to generate TCR α/β (39%–45%) cells, but only PGPC17 produced rare TCR γ/δ bearing (0.1%) DP cells. We conclude that PGPC lines differentiate into HSPC that mature into T cells but with different maturation dynamics that may require line-specific protocol optimization.

Utility of the Resource—Sensory Neuron Protocol Optimization and Subtype Identification
PGPC17_11 was selected to optimize differentiation into peripheral sensory neurons (PSNs) using a small-molecule inhibitor protocol adapted from (Chambers et al., 2012) (Figure 5A). Whole-cell patch-clamp recordings were used to assess excitability. At 2 weeks post-induction, all PSNs responded to sustained current injection with transient spiking but, by 4 weeks, half the neurons had switched to repetitive spiking (Figure 5B). The action potential waveform also experienced significant changes, which included an increase in amplitude (Figure 5C) and a decrease in width (Figure 5D) among both transient and repetitive spiking 4-week-old neurons compared with 2-week-old neurons. Among 4-week-old neurons, repetitive spiking neurons had a significantly lower rheobase (current threshold) than transient spiking neurons (Figure 5E). Additional membrane properties are described in Figures S3A–S3F.

To further characterize phenotype, we imaged the Ca2+ responses evoked by brief application of various agonists. Neurons exhibiting a robust Ca2+ response to KCl application were considered healthy and their responses to capsaicin, GABA, and ATP were tested (Figure 5F). At 2 weeks post-induction, 44.6% of neurons responded to the TrpV1 agonist capsaicin, but that number fell to 10% by week 4 (p < 0.00001). TrpV1 is a marker of peptidergic nociceptors, but is broadly expressed among immature PSNs and is developmentally downregulated (Cavanaugh et al., 2011). Our data suggest that iPSC-derived PSNs follow a similar developmental program. Low TrpV1 expression at 4 weeks suggests that repetitive spiking PSNs represent predominantly non-peptidergic nociceptors (Zeisel et al., 2018), whereas the transient spiking neurons are most likely mechanoreceptors. The proportion of neurons responsive to GABA increased over time (p = 0.0001), as did the proportion responsive to the purinergic receptor agonist ATP (p = 0.042). These results demonstrate that PGPC17 was successfully differentiated into active neurons with a non-peptidergic nociceptor or mechanoreceptor phenotype.

Utility of the Resource—WGS Analysis
To identify iPSC-specific variants we obtained whole-genome sequences of each PGPC line to compare with their respective donor blood sequences (Table 1). On average, we identified 1,502 novel nucleotide variants (range: 1,169–1,981) and 0.5 novel copy-number variants (range: 0–1) per clone. Twenty variants (range: 18–24) affected exonic gene regions: 14 non-synonymous (range: 12–16) and 1.75 loss of function (range: 1–3). PGPC1-73 had a likely pathogenic stopgain variant in the chromatin remodeler BPTF, which may disrupt normal gene expression, and particularly neuronal differentiations. We did not identify any other known pathogenic sequence variants in reprogrammed cell lines. Three...
Figure 5. PGPC17-Derived Sensory Neurons Are Predominately Non-peptidergic

(A) Sensory neuron differentiation scheme. The LDN and SB drug combination was applied between D0 and D5 with the CHIR, DAPT, SU, and NGF combination starting on D2 through D11. Starting on D4, N2 medium was added in increasing 25% increments replacing mTeSR1. Dividing cells were eliminated using Ara-C on D10. N2 medium was changed twice weekly thereafter.

(B–E) Representative spiking patterns to sustained somatic current injection (B). At 2 weeks, all cells spike transiently (independent experiments = 3; technical replicates ≥ 5), whereas at 4 weeks there is a significant increase in the proportion of repetitively spiking neurons (independent experiments = 4; technical replicates ≥ 2) compared with transiently spiking neurons (independent experiments = 4; technical replicates ≥ 2) (p < 0.0001, chi-squared test). The action potential waveform experienced a significant increase in amplitude (C) and a significant decrease in width (D) between 2 and 4 weeks post-induction. Rheobase was significantly lower in repetitive spiking neurons than in transient spiking neurons at 2 or 4 weeks post-induction (E). *p < 0.05 based on Mann-Whitney U tests.

(F) Ca²⁺ revealed a significant decrease in the proportion of neurons responsive to capsaicin between 2 and 4 weeks (p < 0.00001, chi-squared test), whereas the proportion of neurons responsive to GABA (p = 0.0001) or ATP (p = 0.042) significantly increased.
loss-of-function variants, although not associated with human disease, were in genes with high haploinsufficiency scores and known function in embryonic development (PGPC14_26: TRIM71 and FRMD4A; PGPC17_11: ROBO2, Table S3). For PGPC14_26, we also identified an intronic 16-kb deletion of uncertain significance in IL1RAPL1, a gene associated with impaired synaptogenesis and neurodevelopmental deficits. Re-annotation of the PGP donor and cell-line derivative sequences will be important as variant databases mature (Costain et al., 2018).

|                | SNVs/Indels | CNVs     |
|----------------|-------------|----------|
|                | All         | Exonic   | Non-synonymous | Loss of Function | All | Exonic |
| **PGPC_1**     |             |          |                |                 |     |        |
| Genome-wide    | 1,622       |          |                |                 |     |        |
| All genes      | 684         | 24       | 16             | 3               | 1   | 1      |
| OMIM genes     | 133         | 5        | 4              | 1               | 0   | 0      |
| Constrained genes\(^a\) | 204         | 7        | 7              | 3               | 0   | 0      |
| **PGPC_3**     |             |          |                |                 |     |        |
| Genome-wide    | 1,981       |          |                |                 |     |        |
| All genes      | 847         | 19       | 12             | 1               | 0   | 0      |
| OMIM genes     | 172         | 3        | 3              | 0               | 0   | 0      |
| Constrained genes\(^a\) | 216         | 6        | 5              | 0               | 0   | 0      |
| **PGPC_14**    |             |          |                |                 |     |        |
| Genome-wide    | 1,235       |          |                |                 |     |        |
| All genes      | 499         | 18       | 13             | 2               | 1   | 0      |
| OMIM genes     | 104         | 3        | 2              | 0               | 1   | 0      |
| Constrained genes\(^a\) | 150         | 8        | 8              | 2               | 1   | 0      |
| **PGPC_17**    |             |          |                |                 |     |        |
| Genome-wide    | 1,169       |          |                |                 |     |        |
| All genes      | 466         | 23       | 14             | 1               | 0   | 0      |
| OMIM genes     | 95          | 6        | 6              | 1               | 0   | 0      |
| Constrained genes\(^a\) | 113         | 5        | 5              | 1               | 0   | 0      |
| **PGPC17_11 MYBPC3_KO** |             |          |                |                 |     |        |
| Genome-wide    | 917         |          |                |                 |     |        |
| All genes      | 382         | 17       | 9              | 1               | 0   | 0      |
| OMIM genes     | 85          | 4        | 2              | 0               | 0   | 0      |
| Constrained genes\(^a\) | 35          | 7        | 4              | 0               | 0   | 0      |

PGPC1_73, PGPC3_75, PGPC14_26, and PGPC17_11 were compared with the sequence data obtained from whole blood. PGPC17_11 MYBPC3_KO was compared with the PGPC17_11 reprogrammed line. CNV, copy-number variant; indel, insertion/deletion; SNV, single-nucleotide variant. \(^a\)pLI>0.9 (http://exac.broadinstitute.org/).

Utility of the Resource—CRISPR/Cas9 Gene Editing and Phenotyping
To edit a gene for cardiac phenotyping, we targeted a region of MYBPC3 where frameshifts are associated with hypertrophic cardiomyopathy by using gRNAs for CRISPR-Cas9-directed non-homologous end-joining (Skarnes et al., 2019). We nucleofected PGPC17_11 iPSCs with a pSpCas9(BB)-2A-Puro vector containing guide RNA (gRNA) sequences targeting MYBPC3 (Figure 6A). Transfected cells were selected with puromycin treatment and
resistant colonies were isolated and expanded. A karyotypically normal sub-clone bearing an apparent homozygous frameshift mutation was identified in the MYBPC3_KO line by Sanger sequencing (Figures 6A and 6B). To characterize genetic changes, we performed WGS. On-target compound heterozygote MYBPC3 frameshifts were shown to be an 8-bp insertion at chr11:47,359,282insGTG-CAGGA, and a large >260-bp insertion at the same position in the other allele. This insertion did not map to the human genome and was not detected using our PCR-based sequencing due to the size of the insertion. To characterize potential off-target effects in the MYBPC3_KO cells, we first used benchling.com’s prediction tool to identify the top 49 off-target sites. We searched 100 base pairs up- and downstream of each predicted site and found zero novel variants within these regions. When we looked for overall novel genomic variation, 917 new single-nucleotide variants and one intergenic 32-kb deletion (chr18:12,137,685–12,169,689) were found (Table 1). None of these variants were likely pathogenic, similar to our other reported gene edited lines (Deneault et al., 2018).

To examine the consequences of the frameshifts on MYBPC3 protein, we generated CMs as described in Figure 3 and collected protein lysates from PGPC17 parental and MYBPC3_KO iPSC-CMs. Western blots were unable to detect MYBPC3 protein in the KO clone (Figure 6C). We matured CMs until D36–D44 to look for phenotypic evidence of hypertrophic cardiomyopathy as predicted by loss of MYBPC3. Indeed, xCELLigence assays detected increased B Amp in the MYBPC3_KO-CMs compared with isogenic control CMs suggestive of hypercontractility as seen in hypertrophic cardiomyopathy.
similar phenotypes. Our findings demonstrate the utility of PGPC17_1 for gene editing to produce isogenic cell lines for cardiac phenotyping.

**WGS Analysis of Publicly Available HipSci Lines**

Since we found that all our iPSC lines have pre-existing and/or novel variants of potential concern when considering experiments for different lineages, we analyzed downloaded genome sequencing data of five publicly available HipSci lines suggested as healthy controls (HPSI01114i-kolf_2, HPSI0214i.kucg_2, HPSI0214i.wibj_2, HPSI0314i-hoik_1, and HPSI0314i.sojd_3). Across all five samples, 89%-96% of the genome was covered at least 20x (quality metrics in Table S4). We interpreted likely pathogenic variants, loss-of-function constraint gene variants, and VUS as described previously (Reuter et al., 2018) (Supplemental Information).

Two likely pathogenic variants were found in kolf_2 and one in sojd_3 that were predicted to have clinical relevance if identified in humans and could also affect experimental assays. kolf_2 had a substitution of two adjacent nucleotides, disrupting exon-intron boundaries of one COL3A1 allele. One variant was within canonical splice site c.3526-1G>A, and likely to cause out-of-frame exon skipping. If splicing was preserved, the second nucleotide change would result in a likely pathogenic missense alteration p.(Gly1176Ser). COL3A1 haploinsufficiency is associated with dysfunctional connective tissue, such as in the vascular system, skin, intestine, lung, and uterus, and causes vascular type (IV) Ehlers-Danlos syndrome. The same kolf_2 line also harbored a heterozygous 19-bp deletion p.(Pro197Hisfs*12) in ARID2. The variant was likely pathogenic for Coffin-Siris syndrome, a neurodevelopmental disorder with variable skeletal and organ manifestations. These likely pathogenic variants were also confirmed in the kolf2-C1 subline (Skarnes et al., 2019). Finally, sojd_3 harbored a likely pathogenic heterozygous nonsense variant p.(Gln348*) in BCOR. This X-linked gene encodes a transcriptional corepressor with important functions in early embryonic development of various tissues. Females with heterozygous BCOR defects may exhibit oculo-facio-cardiodental syndrome. None of these likely pathogenic variants had been previously reported, and we cannot determine if they were present in the donor genomes, or arose during reprogramming, and could therefore be mosaic. We also identified several loss-of-function variants of uncertain significance in constrained genes in hoik_1 and kucg_2, mostly with known functions in early development (as in PTK2, ZNF398, UBE3C, CDC37, and TNS3; Table S3) and many VUS (Supplemental Information). We did not identify pathogenic or likely pathogenic variants in wibj_2, which suggests it is the variant-preferred line among this subset.

**DISCUSSION**

Here we generated a high-quality resource of versatile iPSC control lines for use in disease modeling studies. These cells have the benefit of both annotated genomic variants and demonstrated multilineage-directed differentiation into functional cortical neurons, CMs, and hepatocytes. Pilot users showed that the lines can be used to generate kidney organoids, T lymphocytes, or to identify specific subtypes of active sensory neurons. We also performed gene editing, which revealed a preliminary phenotype in isogenic MYBPC3 KO CMs similar to another isogenic pair (Cohn et al., 2019).

Apart from their versatility, the main advantage of these blood-derived footprint-free lines is the clinical annotation of potentially disease-associated variants that may affect cellular phenotypes. Variant analysis in the PGPC participants’ blood had revealed heterozygous variants of unknown significance in all individuals (Reuter et al., 2018). This observation suggests that it may not be possible to isolate universal control lines and reinforces the importance for WGS in characterizing control lines, especially as clinical annotation gains precision with ongoing variant discoveries. WGS has the advantage of allowing detection of coding variants, CNVs, and noncoding variants, although the latter have not yet been fully explored in these lines. Knowledge of the donors’ genomes allowed predictions on how to prioritize control lines for use as tissue specific controls. For example, as PGPC3 and PGPC14 had variants that could predispose to altered cardiac channel function, PGPC17 was deemed to be the preferred line for the study of cardiac disease. PGPC3, however, were variant preferred for neurological disorders. Consistent with a precision health approach, this strategy would allow matching of genotyped iPSC controls to the disease being modeled.

WGS has previously determined that iPSC lines have variants that differ from those in the donor. Our WGS data reveals that the reprogrammed lines have more than a thousand new SNVs each, whereas only two new CNVs were detected likely due to previous selection for normal karyotype. Most variants were of uncertain significance, with new variants of potential concern found in two of four of the blood (CD34+ cell)-derived PGPC lines (PGPC1_73 and PGPC14_26 in Table S3). Genome sequencing of the MYBPC3 KO line showed more than 900 additional SNVs compared with the unedited iPSC line. None of the new variants were near potential gRNA cut sites, suggesting that they were not off-target and were indeed novel mutations. These analyses highlight that iPSC lines harbor variants of potential concern that are not found in the donor blood. Moreover, our annotation of five healthy control lines...
from the HipSci consortium that were generated from fibroblasts discovered likely pathogenic variants in two lines and additional loss-of-function variants in constrained genes in two other lines, leaving only wibj.2 as a preferred healthy control line. Since donor WGS is not available for the HipSci lines, it is not possible to determine whether these potentially damaging variants were pre-existing or were captured during fibroblast reprogramming. In contrast, our precision health resource identified >1,000 new variants in each iPSC line, consistent with numbers reported for fibroblast reprogramming (Abyzov et al., 2017). We propose that clinical annotation of WGS data is an important quality control measure of iPSC lines, and its expanded use will identify the best source of healthy control cells to reprogram to find additional variant-preferred lines for disease modeling.

Disease modeling has generally used two to three lines from each individual to account for variability in reprogramming. To account for 1,000–2,000 novel variants in each line compared with the parental genome, this study provides another rationale for studying multiple lines from each individual. With this in mind, we generated a resource of four to five iPSC lines each from two males and two females, all with standard pluripotency characterization available. We also performed multilineage-directed differentiation on a single line from three individuals, assuming that single lines from three to four individuals can account for inter-individual variability. One highly characterized line is therefore available from three PGPC participants, and preferred lines are likely to be of high utility for gene editing studies that compare the phenotype of isogenic cells. Ultimately, users of the resource will select one or more lines from each PGPC participant depending on their research strategy. Future efforts to apply our precision health approach to the characterization of additional control lines in cell repositories should increase the numbers of variant-preferred iPSC banked for gene editing and disease modeling studies.

Overall, our resource upgrades the quality of existing healthy iPSC lines in two ways. First our identification of novel variants of potential concern after reprogramming suggest that a subset of lines may not accurately reflect the phenotype of the original donor in some tissues. To address this concern, our precision health resource provides variant-preferred lines as controls for cardiac or neurological disease modeling and for use in gene editing strategies to create isogenic pairs of mutant and control cells. Second, our exhaustive characterization of multilineage-directed differentiation by pilot users provides strong evidence that the lines can be broadly applied by the disease modeling community.

EXPERIMENTAL PROCEDURES

Reprogramming of PGPC iPSCs was performed under the approval of the Canadian Institutes of Health Research Stem Cell Oversight Committee, and the Research Ethics Board of The Hospital for Sick Children, Toronto. Blood cells were reprogrammed with Sendai virus to deliver reprogramming factors, and iPSCs were maintained in feeder-free conditions with mTeSR1 (STEMCELL Technologies); see Supplemental Information. WGS was performed on Illumina HighSeq X and analyzed as described previously (Reuter et al., 2018). A vector-based CRISPR/Cas9 approach was used to mutagenize MYBPC3, further described in Supplemental Information. Detailed descriptions of differentiations, characterizations, and functional assays are summarized in figures and Supplemental Information. Overexpression of Ngn2 induced iPSCs to differentiate to glutamatergic neurons. Extracellular electrophysiology recordings were collected with an Axion Maestro MEA reader (Axion Biosystems) micro-electrode array as described in the Supplemental Information. CMs were differentiated using STEMdiff Cardiomyocyte Differentiation Kits (STEMCELL Technologies). CM calcium imaging was captured with loading cells with Fluo-4 dye and taking images at 4 Hz for 30 s. Contractile and electrical activity was recorded with an xCELLiGence RTCA CardioECR (ACEA Biosciences). CYP3A7 was measured using a p450-Glo assay kit (Promega) as per the manufacturer’s protocol. Whole-cell electrophysiology recordings where made at room temperature with an Axopatch 200B (Molecular Devices) from borosilicate patch electrodes. Ca²⁺ imaging was performed on sensory neurons incubated in Ca²⁺-green-1 AM dye (Thermo Fisher Scientific) at room temperature. Images were acquired at 25 Hz using a NeuroCCD-SM256 imaging system (RedShirt Imaging).

ACCESSION NUMBERS

WGS datasets are available from EGA: EGAS00001003684 and RNA sequencing datasets are available from the GEO: GSE132012. iPSC lines are available upon request.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.stemcr.2019.11.003.

AUTHOR CONTRIBUTIONS

M.R.H., M.S.R., S.W.S., and J.E. designed the research project. M.R.H. and J.E. supervised the project. M.S.R. performed WGS clinical annotation and off-target analyses. M.R.H., N.T., and W.W. contributed to the CRISPR experiments. M.R.H., W.W., N.T., J.L., S.S., J.M., L.S.L., P.M.B., A.P., A.R., and G.M. contributed to iPSC isolation, characterization and differentiation. Cells studied by each lab group: iPSC by J.E. and S.W.S., cortical neurons by J.E., cardiomyocytes by S.M. and J.E., hepatocytes by B.M.K., kidney by N.D.R. and J.E., T cells by J.C.Z.-P. and M.K.A., sensory neurons by S.A.P. and J.E. C.K., D.d.C.R., J.H., P.P., M.R., E.C.M., and M.J.S. provided technical help. M.R.H., M.S.R., N.T., J.L., J.M., L.S.L., P.M.B., J.C.Z.-P., M.K.A., S.A.P., N.D.R., B.M.K., S.M., S.W.S., and J.E. wrote the manuscript with comments from all co-authors. Specific contributions of the co-corresponding
authors: S.W.S. lab obtained donor blood for reprogramming and performed WGS analyses and annotation on iPSCs; J.E. lab generated iPSCs, cortical neurons, cardiomyocytes, kidney organoids, and sensory neurons.

ACKNOWLEDGMENTS

The research was supported by grants from the University of Toronto McLaughlin Centre (MC-2014-06) and Ontario Brain Institute Province of Ontario Neurodevelopmental Disorders Network (to J.E. and S.W.S.; IDs-11-02), GlaxoSmithKline—Canada Institutes of Health Research (CIHR) Chair in Genome Sciences (to S.W.S.), Ted Rogers Center for Heart Research Strategic Innovation grant (to J.E. and S.M.), Heart and Stroke Foundation Chair (to S.M.), CIHR Team Grant (to B.M.K.; THC-135232), Tier I Canada Research Chair and CIHR Foundation Grant (to N.D.R.; SOP-155609), CIHR Chronic Pain Network-SPOR (to S.A.P. and J.E.; 2017-007), Medicine by Design New Ideas grants (to J.E. and M.K.A., MBNCL1-2017-03; and J.C.Z.-P. and M.K.A., CITP-2016-20), NSERC RGPII grant (to M.K.A., 05333-14), and Fellowship support from SickKids RestraComp award (to J.M and S.S.). We thank B. Thiruvahindrapuram, T.Nalpathamkalam, W.W. Sung, Z. Wang, and G. Kaur for bioinformatic support; The Center for Applied Genomics (TCAG; a CGEn node), the SickKids-uhn Flow and Mass Cytometry Facility and the SickKids Imaging Facility for technical support. The Ngn2/rtTA lentiviral constructs were gifts from T.C. Südhof, and we thank N.N. Kashi and K. Linda for technical advice. We thank the Personal Genome Project Canada, Cheryl Cytrynbaum, Ny Hoang, and Barbara Kellem for genomic data and collecting samples for reprogramming, and the PGPC blood donors for volunteering to participate in this research. M.R.H. became an employee of Stemcell Technologies Inc during peer review of this report.

Received: June 11, 2019
Revised: November 4, 2019
Accepted: November 7, 2019
Published: December 5, 2019

REFERENCES

Abyzov, A., Tomasini, L., Zhou, B., Vasmatzis, N., Coppola, G., Amenduni, M., Pattini, R., Wilson, M., Gerstein, M., Weissman, S., et al. (2017). One thousand somatic SNVs per skin fibroblast cell set baseline of mosaic mutational load with patterns that suggest proliferative origin. Genome Res. 27, 512–523.

Ball, M.P., Thakuria, J.V., Zaranek, A.W., Clegg, T., Rosenbaum, A.M., Wu, X., Angrist, M., Bhak, J., Bobe, J., Callow, M.J., et al. (2012). A public resource facilitating clinical use of genomes. Proc. Natl. Acad. Sci. U.S.A 109, 11920–11927.

Cavanaugh, D.J., Chesler, A.T., Braz, J.M., Shah, N.M., Julius, D., and Basbaum, A.I. (2011). Restriction of transient receptor potential vanilloid-1 to the peptidergic subset of primary afferent neurons follows its developmental downregulation in nonpeptidergic neurons. J. Neurosci. 31, 10119–10127.

Chambers, S.M., Qi, Y., Mica, Y., Lee, G., Zhang, X.J., Niu, L., Bilsland, J., Cao, L., Stevens, E., Whiting, P., et al. (2012). Combined small-molecule inhibition accelerates developmental timing and converts human pluripotent stem cells into nociceptors. Nat. Biotechnol. 30, 715–720.

Cheung, A.Y.L., Horvath, L.M., Grafodatskaya, D., Pascrei, P., Weksberg, R., Hotta, A., Carrel, L., and Ellis, J. (2011). Isolation of MECP2-null Rett syndrome patient hiPSC cells and isogenic controls through X-chromosome inactivation. Hum. Mol. Genet. 20, 2103–2115.

Cohn, R., Thakar, K., Lowe, A., Ladha, F.A., Pettinato, A.M., Romano, R., Meredith, E., Chen, Y.-S., Atamanuk, K., Huey, B.D., et al. (2019). A contraction stress model of hypertrophic cardiomyopathy due to sarcomere mutations. Stem Cell Reports 12, 71–83.

Costain, G., Jobling, R., Walker, S., Reuter, M.S., Smell, M., Bowdin, S., Cohn, R.D., Dupuis, L., Hewson, S., Mercimek-Andrews, S., et al. (2018). Periodic reanalysis of whole-genome sequencing data enhances the diagnostic advantage over standard clinical genetic testing. Eur. J. Hum. Genet. 26, 740–744.

D’Antonio, M., Benaglio, P., Jakubosky, D., Greenwald, W.W., Matsui, H., Donovan, M.K.R., Li, H., Smith, E.N., D’Antonio-Chronowaska, A., and Frazer, K.A. (2018). Insights into the mutational burden of human induced pluripotent stem cells from an integrative multi-omics approach. Cell Rep. 24, 883–894.

Deneault, E., White, S.H., Rodrigues, D.C., Ross, P.I., Faheem, M., Zaslavsky, K., Wang, Z., Alexandrova, R., Pellecchia, G., Wei, W., et al. (2018). Complete disruption of autism-susceptibility genes by gene editing predominantly reduces functional connectivity of isogenic human neurons. Stem Cell Reports 11, 1211–1225.

Gore, A., Li, Z., Fung, H.-L., Young, J.E., Agarwal, S., Antosiewicz- Bouget, J., Canto, I., Giorgetti, A., Israel, M.A., Kiskinis, E., et al. (2011). Somatic coding mutations in human induced pluripotent stem cells. Nature 471, 63–67.

Hoekstra, S.D., Stringer, S., Heine, V.M., and Posthuma, D. (2017). Genetically-informed patient selection for iPSC studies of complex diseases may aid in reducing cellular heterogeneity. Front. Cell. Neurosci. 11, 1–8.

Hollingsworth, E.W., Vaughan, J.E., Orack, J.C., Skinner, C., Khouri, J., Lizarraga, S.B., Hester, M.E., Watanabe, F., Kosik, K.S., and Imitola, J. (2017). IPhemap: an atlas of phenotype to genotype relationships of human iPSC models of neurological diseases. EMBO Mol. Med. 9, 1742–1762.

Klipinen, H., Goncalves, A., Leha, A., Afzal, V., Alasoo, K., Ashford, S., Bala, S., Benasdekk, D., Casale, F.P., Culley, O.J., et al. (2017). Common genetic variation drives molecular heterogeneity in human iPSCs. Nature 546, 370–375.

Lan, F., Lee, A.S., Liang, P., Sanchez-Freira, V., Nguyen, P.K., Wang, L., Han, L., Yen, M., Wang, Y., Sun, N., et al. (2013). Abnormal calcium handling properties underlie familial hypertrophic cardiomyopathy pathology in patient-specific induced pluripotent stem cells. Cell Stem Cell 12, 101–113.

Popp, B., Krumbiegel, M., Grosch, J., Sommer, A., Uebe, S., Kohl, Z., Plotz, S., Farrell, M., Trautmann, U., Kraus, C., et al. (2018). Need for high-resolution genetic analysis in iPSC: results and lessons from the ForIPS consortium. Sci. Rep. 8, 1–14.

Reuter, M.S., Walker, S., Thiruvahindrapuram, B., Whitney, J., Cohn, I., Sondheimer, N., Yuen, R.K.C., Trost, B., Paton, T.A., Pereira, S.L., et al. (2018). The Personal Genome Project
Canada: findings from whole genome sequences of the inaugural 56 participants. Can. Med. Assoc. J. 190, E126–E136.
Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W.W., Hegde, M., Lyon, E., Spector, E., et al. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet. Med. 17, 405–423.
Schwartzentruber, J., Foskolou, S., Kilpinen, H., Rodrigues, J., Alasoo, K., Knights, A.J., Patel, M., Goncalves, A., Ferreira, R., Benn, C.L., et al. (2018). Molecular and functional variation in iPSC-derived sensory neurons. Nat. Genet. 50, 54–61.
Skarnes, W.C., Pellegrino, E., and McDonough, J.A. (2019). Improving homology-directed repair efficiency in human stem cells. Methods 164–165, 18–28.
Streeter, I., Harrison, P.W., Faulconbridge, A., Flice, P., Parkinson, H., and Clarke, L. (2017). The human-induced pluripotent stem cell initiative—data resources for cellular genetics. Nucleic Acids Res. 45, D691–D697.
Takahashi, K., and Yamanaka, S. (2016). A decade of transcription factor-mediated reprogramming to pluripotency. Nat. Rev. Mol. Cell Biol. 17, 183–193.
Takasato, M., Er, P.X., Chiu, H.S., Maier, B., Baillie, G.J., Ferguson, C., Parton, R.G., Wolvetang, E.J., Roost, M.S., De Sousa Lopes, S.M.C., et al. (2015). Kidney organoids from human iPS cells contain multiple lineages and model human nephrogenesis. Nature 526, 564–568.
Tchieu, J., Kuoy, E., Chin, M.H., Trinh, H., Patterson, M., Sherman, S.P., Aimiwu, O., Lindgren, A., Hakimian, S., Zack, J.A., et al. (2010). Female human iPS cells retain an inactive X-chromosome. Cell Stem Cell 7, 329–342.
Yoshihara, M., Araki, R., Kasama, Y., Sunayama, M., Abe, M., Nishida, K., Kawai, H., Hayashizaki, Y., and Murakawa, Y. (2017). Hot-spots of de novo point mutations in induced pluripotent stem cells. Cell Rep. 21, 308–315.
Zeisel, A., Hochgerner, H., Lönnnerberg, P., Johnsson, A., Memic, F., van der Zwan, J., Haring, M., Braun, E., Borm, L.E., La Manno, G., et al. (2018). Molecular architecture of the mouse nervous system. Cell 174, 999–1014.
Zhang, Y., Pak, C., Han, Y., Ahlenius, H., Zhang, Z., Chanda, S., Marro, S., Patzke, C., Acuna, C., Covy, J., et al. (2013). Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron 78, 785–798.
Supplemental Information

Precision Health Resource of Control iPSC Lines for Versatile Multilineage Differentiation

Matthew R. Hildebrandt, Miriam S. Reuter, Wei Wei, Naeimeh Tayebi, Jiajie Liu, Sazia Sharmin, Jaap Mulder, L. Stephen Lesperance, Patrick M. Brauer, Rebecca S.F. Mok, Caroline Kinnear, Alina Piekna, Asli Romm, Jennifer Howe, Peter Pasceri, Guoliang Meng, Matthew Rozycki, Deivid C. Rodrigues, Elisa C. Martinez, Michael J. Szego, Juan C. Zúñiga-Pflücker, Michele K. Anderson, Steven A. Prescott, Norman D. Rosenblum, Binita M. Kamath, Seema Mital, Stephen W. Scherer, and James Ellis
Figure S1. Quality control and pluripotency characterization of PGPC lines. Related to Table S1.
(A) Normal karyotype, positive pluripotency marker immunocytochemistry labeling (OCT4, SSEA4, NANOG, and TRA-1-60) of iPSC colonies, and germ layer marker labeling of embryoid body cultures (TUBB3, SMA, and AFP) were found in all PGPC cell lines. Colour channels were independently altered to adjust contrast for publication. (B) PluriTest plot of PGPC1_73, 3_75, 14_26, and 17_11 RNAseq data localized all four lines sequenced to date to the pluripotency quadrant. (C) Female iPSC lines were assessed for X chromosome inactivation via the androgen receptor assay. All lines showed clear separation between the two amplicons in the undigested control. After digestion with methylation-sensitive enzymes all lines show complete skewing towards either active X-chromosome as shown by the presence of a single peak cluster. Full skewing list found in Table S1.
Figure S2.

A

| ROCKi | CHIR 8 μm | Heparin 1 μg/ml |
|-------|-----------|-----------------|
|       |           |                 |

mTeSR1

APEL-2

* aggregation and 1 hour pulse with CHIR 5 μM

B

| BMP4 | VEGF, IL6, IL11 | IL7, FLT3L, SCF |
|------|-----------------|-----------------|
| ROCKi | bFGF            | SB-431542       |

IGF1, EPO, SCF

OP9-DL4 coculture

StemPro34

alpha-MEM

C

unstained  preMACS  postMACS

IPSC11

PGPC3.75

PGPC14.26

PGPC17.11

CD34

RFK
Figure S2. PGPC differentiation to kidney and T-cells. Related to Figure 4.

(A) Kidney organoid differentiation scheme. Single-cell suspensions of iPSCs were dissociated to single cells and grown in ROCK inhibitor for 24 hours before beginning differentiation by changing medium to APEL-2 supplemented with CHIR. Medium was fully replaced every other day until D5 where CHIR-medium was switched to APEL-2 supplemented with human fibroblast growth factor (hFGF)-9 and heparin. D7 cells were dissociated and aggregated by centrifugation in 1.5 ml tubes. Three to four aggregates were carefully transferred to Transwell cell culture plates and grown in media consisted of APEL-2 with CHIR. Medium was switched to APEL-2 with hFGF9 and heparin after one hour. Media was replaced every other day thereafter. From D12 onward, APEL-2 was only supplemented with heparin. (B) T-cell differentiation scheme. iPSCs were lifted with ReLeSR to clusters of 5-10 cells and transferred to low-cluster plates in StemPro34 media containing BMP4 and ROCK inhibitor. Media was changed after 24 hours to StemPro34 containing BMP4, CHIR and bFGF. At 42 hours, media was changed again now including SB-431542. At D4 EBs have a crumpled appearance and media was changed to include bFGF, VEGF, IL-6, IL-11. At D6 media was added on top of current culture media and contained VEGF, EPO, IGF, SCF, IL-6, IL-11, and bFGF. At D8, EBs were dissociated and enriched for CD34+ cells by MACS then co-cultured with OP9-DL4 cells in Alpha MEM containing Flt3-L, SCF and IL-7. Differentiating T-lineage cells were transferred to fresh OP9-DL4 cells approximately every 4-5 days. (C) Expression of CD34 measured by flow cytometry before and after MACS enrichment for CD34+ cells.
Figure S3. Comparison of other electrophysiological properties of sensory neurons. Related to Figure 5. (A) membrane capacitance, which is proportional to membrane surface area, (B) soma diameter, (C) rheobase normalized by membrane capacitance, (D) input resistance, (E) membrane leak, based on the reciprocal of input resistance normalized by capacitance, and (F) resting membrane potential. * shows p < 0.05 based on Mann-Whitney U tests. Notably, differences in input resistance (D) were lost after normalizing by capacitance (A), which suggests that differences in input resistance were due primarily to cell size rather than membrane leakiness.
Table S1. Summary of PGPC iPSC characterizations.

| Cell line    | Gender | Age | Mycoplasm a | Sendai virus | Karyotype | X-chromosome active | Pluripotency staining                                           | Embryoid body assay | Additional information |
|--------------|--------|-----|-------------|--------------|-----------|---------------------|---------------------------------------------------------------|--------------------|------------------------|
| PGPC3_23     | M      | 48  | negative    | negative     | normal    | N/A                 | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | positive               |
| PGPC3_66     | M      | 48  | negative    | negative     | normal    | N/A                 | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | positive               |
| PGPC3_75     | M      | 48  | negative    | negative     | normal    | N/A                 | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | positive               |
| PGPC3_93     | M      | 48  | negative    | negative     | normal    | N/A                 | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | positive               |
| PGPC3_87     | M      | 48  | negative    | negative     | normal    | N/A                 | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | positive               |
| PGPC14_26    | F      | 62  | negative    | negative     | normal    | Skewed—X1          | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | enhanced characterization |
| PGPC14_70    | F      | 62  | negative    | negative     | normal    | Skewed—X1          | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | enhanced characterization |
| PGPC14_16    | F      | 62  | negative    | negative     | normal    | Skewed—X1          | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | enhanced characterization |
| PGPC14_94    | F      | 62  | negative    | negative     | normal    | Skewed—X1          | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | enhanced characterization |
| PGPC17_4     | M      | 61  | negative    | negative     | normal    | N/A                 | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | enhanced characterization |
| PGPC17_11    | M      | 61  | negative    | negative     | normal    | N/A                 | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | enhanced characterization |
| PGPC17_45    | M      | 61  | negative    | negative     | normal    | N/A                 | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | enhanced characterization |
| PGPC17_80    | M      | 61  | negative    | negative     | normal    | N/A                 | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | enhanced characterization |
| PGPC17_142   | M      | 61  | negative    | negative     | normal    | N/A                 | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | enhanced characterization |
| PGPC1_30     | F      | 46  | negative    | negative     | normal    | Skewed—X1          | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | enhanced characterization |
| PGPC1_67     | F      | 46  | negative    | negative     | normal    | Skewed—X2          | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | enhanced characterization |
| PGPC1_73     | F      | 46  | negative    | negative     | normal    | Skewed—X2          | Oct4, Nanog, SSEA4, Tra-1-60 positive                        | TUBB3, SMA, AFP     | enhanced characterization |
| PGC1_96  | F   | 46 | negative | negative | normal | Skewed—X | Oct4, Nanog, SSEA4, Tra-1-60 positive | TUBB3, SMA, AFP positive |
|----------|-----|----|----------|----------|--------|---------|----------------------------------------|-------------------------|
Table S2 (excel). Comprehensive list of genomic variants in PGPC donor and iPSC lines related to Table 1.

Table S3. Genomic coding variants of concern in PGPC donors/cell lines and publicly available iPSC lines related to Table 1.

| PGPC1 (donor) | Gene, accession number | Variant, zygosity | Disease/biological function, inheritance | Interpretation |
|---------------|------------------------|-------------------|------------------------------------------|----------------|
| PGPC1_73      | BPTF, NM_004459        | c.3115A>T, p.(Lys1039*), het | Chromatin remodeler, required for normal gene expression, neurodevelopmental disorder with dysmorphic facies and distal limb anomalies | Likely pathogenic |
| PGPC1_73      | DSCAM, NM_001389       | c.509-1G>A, p.?, het | Neural cell adhesion molecules, essential for neuronal circuit assembly | Uncertain, pLI=1.0 |
| PGPC1_73      | GCN1, NM_006836        | c.7423delT, p.(Ser2475Profs*19), het | Interacts with translating 80S ribosomes | Uncertain, pLI=1.0 |
| PGPC1_73      | ZFP1                   | chr16:75153001-75187000 duplication exon 1, het | Zinc finger protein | Uncertain, pLI=0 |
| PGPC3 (donor) | TRPM4, NM_017636       | c.2531G>A, p.(Gly844Asp), het | Progressive familial heart block, AD | VUS |
| PGPC3_75      | ZNF283, NM_001297752   | c.63C>A, p.(Cys21*), het | Unknown. | Uncertain, pLI=0 |
| PGPC14 (donor) | KCNE2, NM_172201      | c.29C>T, p.(Thr10Met), het | Long QT syndrome; Atrial fibrillation, AD | VUS |
| PGPC14_26     | TP53, NM_000546        | c.473G>A, p.(Arg158His), (mosaic?) | Li-Fraumeni syndrome, AD | Likely pathogenic |
| PGPC14_26     | TRIM71, NM_001039111   | c.1963C>T, p.(Arg655*), het | Ubiquitin ligase. Neural differentiation, brain development. Embryonic stem cell proliferation. | Uncertain, pLI=0.99 |
| PGPC14_26     | FRMD4A, NM_001318337   | c.563+1G>T, het | Cell polarity, associated with Alzheimer’s disease risk. | Uncertain, pLI=1.0 |
| PGPC14_26     | ILIRAPL1               | chrX:29319382-29335268 intronic deletion, het | Presynaptic differentiation during synapse formation, associated with nonsyndromic X-linked recessive intellectual disability | VUS |
| PGPC17 (donor) | N/A                    | N/A               | N/A | N/A |
| PGPC17_11     | ROBO2, NM_001290039    | c.2621G>A, p.(Trp874*), het | Slit receptor. Axon branching, dendritic patterning, and neuronal migration. Endothelial cell | Uncertain, pLI=1.0 |
| HPSI0114i.kolf_2 | COL3A1, NM 000990 | c.[3526-1G>A; 3526G>A], p.[?; (Gly1176Ser)], het | Ehlers-Danlos syndrome, type IV, AD | Likely pathogenic |
| HPSI0114i.kolf_2 | ARID2, NM 152641 | c.590_608delCTAAAATCATCACTTTACT, p.(Pro197Hisfs*12), het | Coffin-Siris syndrome, AD | Likely pathogenic |
| HPSI0114i.kolf_2 | ZNF398, NM 170686 | c.1420C>T, p.(Gln474*), het | Unknown. | Uncertain, pLI=0.96 |
| HPSI0114i.kolf_2 | UBE3C, NM 014671 | c.661_668del TCAAGTAT, p.(Ser221*), het | Ubiquitin ligase. Cell proliferation. Wnt/β-catenin signaling. | Uncertain, pLI=1.0 |
| HPSI0114i.kolf_2 | CDC37, NM 007065 | c.192C>A, p.(Cys64*), het | Cell cycle progression. | Uncertain, pLI=0.96 |
| HPSI0214i-wibj_2 | N/A | N/A | N/A | N/A |
| HPSI0314i.sojd_3 | BCOR, NM 017745 | c.1042C>T, p.(Gln348*), het | Oculofaciocardiodental syndrome, XLD | Likely pathogenic |
| HPSI0314i-hoik_1 | PTK2, NM_001352736 | c.2551C>T, p.(Gln851*), het | Cell growth. Integrin-mediated signal transduction. Heart and blood vessel development. | Uncertain, pLI=1.0 |
| HPSI0214i-kueg_2 | TNS3, NM 022748 | c.1239delC, p.(Arg414Glyfs*4), het | Cell migration. Focal adhesion. | Uncertain, pLI=1.0 |

Abbreviations: AD, autosomal dominant; het, heterozygous; PGPC, Personal Genome Project Canada; pLI, probability of loss-of-function intolerance (http://exac.broadinstitute.org/); VUS, variant of uncertain significance; XLD, X-linked dominant.
Table S4. Genome sequencing quality metrics of donor material, newly derived and publicly available iPSC lines.

| Sample                  | Total reads | Median read depth | Region covered 10X [%] | Region covered 20X [%] | Region covered 30X [%] |
|-------------------------|-------------|-------------------|------------------------|------------------------|------------------------|
| PGPC1 (donor)           | 933,877,478 | 38                | 96.8                   | 95.6                   | 84.4                   |
| PGPC1_73                | 904,069,536 | 33                | 96.8                   | 94.7                   | 69.0                   |
| PGPC3 (donor)           | 963,792,408 | 39                | 97.3                   | 93.7                   | 84.5                   |
| PGPC3_75                | 780,473,626 | 31                | 96.8                   | 88.9                   | 55.2                   |
| PGPC14 (donor)          | 958,002,700 | 39                | 97.0                   | 95.9                   | 87.0                   |
| PGPC14_26               | 774,673,256 | 31                | 96.9                   | 92.4                   | 56.4                   |
| PGPC17 (donor)          | 779,923,976 | 33                | 96.8                   | 90.8                   | 64.7                   |
| PGPC17_11               | 765,097,084 | 31                | 96.8                   | 89.0                   | 56.3                   |
| PGPC17_11_MYBPC3_KO     | 914,771,348 | 36                | 97.4                   | 92.9                   | 77.9                   |
| HPSI0114i-kolf_2        | 785,913,438 | 31                | 96.1                   | 89.0                   | 57.9                   |
| HPSI0214i-kucg_2        | 856,258,780 | 37                | 96.9                   | 92.7                   | 77.7                   |
| HPSI0214i-wibj_2        | 841,290,742 | 35                | 97.0                   | 95.2                   | 74.8                   |
| HPSI0314i-hoik_1        | 945,070,686 | 36                | 97.0                   | 95.6                   | 80.9                   |
| HPSI0314i-sojd_3        | 964,602,272 | 32                | 96.7                   | 94.0                   | 63.4                   |

Table S5 (excel). Comprehensive list of genomic variants in HipSci cell lines.
SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Reprogramming and cell culture
Reprogramming of iPSCs was performed under the approval of the Canadian Institutes of Health Research Stem Cell Oversight Committee, and the Research Ethics Board of The Hospital for Sick Children (REB#1000056039). CD34+ cells were isolated from blood samples collected from each of the four PGPC donors using CD34+ Reprogramming Kit (STEMCELL Technologies) according to the manufacture’s protocol. Cultured CD34+ cells were tested for mycoplasma by PCR (see Table S7 for primer pairs) during routine early passaging then reprogrammed using CytoTune-iPS 2.0 Sendair Reprogramming Kit (ThermoFisher) with modifications post transduction. Transduced cells were plated in ReproTeSR (STEMCELL Technologies) and isolated colonies were directly transferred to mTeSR1 (STEMCELL Technologies).

iPSC maintenance
All iPSC lines were maintained on Matrigel (Corning) coated dishes using mTeSR1. Daily media changes were performed except for the day following passaging. ReLeSR (STEMCELL Technologies) was used to lift clumps of iPSCs for routine weekly passaging. Accutase (InnovativeCellTechnologies) and 10 µM Rho-associated kinase (ROCK) inhibitor (Y-27632; STEMCELL Technologies) were used for single-cell dissociation purposes unless otherwise specified. Mycoplasma was routinely tested as cells were maintained.

PCR for Sendai virus
Total RNA was isolated using TRIzol (Invitrogen) according to the manufacture’s protocol. 1 µg isolated RNA was used to generate cDNA with SuperScript VILO cDNA Synthesis kit (ThermoFisher) according to the manufacture’s protocol. 10 µl of the reaction was amplified with AccuPrime Taq DNA polymerase with Sendai virus primers Table S7. Resulting DNA was run on a 2% agarose gel, stained with ethidium bromide and imaged. Passage (P)2 iPSC cDNA samples were used as a positive control for Sendai virus.

Karyotyping
Karyotyping and standard G-banding chromosome analysis with 400-band resolution was performed at The Centre for Applied Genomics (TCAG).

AR assay
Assay was performed as previously described (Cheung et al., 2011). 400 ng genomic DNA was digested with HpaI and HhaI simultaneously for 5 hours. 2 µl digest was used in a PCR with Platinum™ Taq DNA Polymerase High Fidelity (Invitrogen) using AR gene primers (Table S7). 15 µl PCR was sent to TCAG for Genetic Analysis electrophoresis and resulting traces were analyzed using Peak Scanner software. Male samples were used to confirm complete digestion.

RNA sequencing and Pluritest
RNA was collected from single wells of 6-well plates using RNA PureLink RNA mini kit (Life Technologies) as per supplied protocol. RNA samples were sent to TCAG where RNA quality was checked by Bioanalyzer then sequenced as paired end 2x125 bases on an Illumina HiSeq 2500 at a depth of 20M reads. Resulting compressed FASTQ files were uploaded to www.pluritest.org for analysis (Müller et al., 2011).

Immunocytochemistry
iPSCs were plated on 24-well dishes for epifluorescence microscopy while differentiated cells were seeded on µ-Plate 24-well Black plates (IBIDI) for confocal microscopy. iPSCs, iPSC-derived cardiomyocytes and iPSC-derived hepatocytes were fixed with 4% paraformaldehyde (PFA), while iPSC-derived neurons were fixed with PFS (4% PFA in Krebs-Sucrose buffer, 50 mM KCl, 1.2 mM CaCl₂, 1.3 mM MgCl₂, 20 m Hepes pH 7.4, 12 mM NaH₂PO₄, 400 mM sucrose, 145 mM NaCl, 10 mM glucose in water). Three washes with phosphate buffered saline (PBS) was performed following incubation with 0.1% Triton X-100 incubation for 10 minutes. Cells were blocked with blocking solution for one hour. Primary antibody (for appropriate dilutions see Table. S8) was prepared in blocking solution and incubated with the cells at 4°C overnight. Three washes with PBS were done before incubating with fluorescence-tagged secondary antibodies (Table S8) for one hour in the dark at room temperature. Cells were washed twice with PBS then incubated with 1:2000 4’,6-diamidino-2-phenylindole (DAPI) for 5 minutes to counterstain nuclei followed with a final wash of PBS.

Imaging
iPSCs and sensory neurons were imaged on a Leica DM14000B epifluorescence microscope using a DFC
7000T camera and LAS X software (Leica). Cortical neurons were imaged with an Olympus IX81 spinning disc confocal using a Hamamatsu C9100-13 EM-CCD camera and Volocity software (Perkin Elmer). Cardiomyocytes were imaged with a Zeiss AxioVert 200M microscope using a Hamamatsu C9100-13 EM-CCD camera and Volocity software. Hepatocytes and kidney organoid sections were imaged with a Zeiss AxioVert 200M microscope, using an AxioCam HRm camera, and AxioVision (V4.9.2 SP3). ImageJ (V1.52i) and Photoshop (V12.0) were used for image analysis and figure preparation. Individual colour channels of fluorescent images were altered to correct for brightness.

Whole genome sequencing (WGS) and clinical annotation

Genomic DNA was isolated using Quick-DNA miniprep kit (Zymo Research) according to manufacturer’s protocol. 1 µg gDNA was submitted to TCAG for genomic library preparation and whole genome sequencing on an Illumina HighSeq X system as described in Reuter et al., 2018. For reprogrammed PGPC lines, we performed genome sequencing as previously described Reuter et al., 2018. We used Illumina provided software bcl2fastq (v2.20) to convert the per-cycle BCL basecall files to standard sequencing output in FASTQ format. FastQC was used to assess the quality of the experiment, and FastQ Screen to check the composition of the library. Reads were aligned to the reference human genome (GRCh37) using BWA mem (v0.7.12) (Li and Durbin, 2009). Duplicate reads were marked using Picard Tools (v2.5.0). Indel realignment, base quality score recalibration and germline variant detection using HaplotypeCaller were performed using GATK 3.7 following the best practices recommendation (Van der Auwera et al., 2013; DePristo et al., 2011). Resulting variant calls were annotated using a custom pipeline developed at TCAG (Yuen et al., 2015), based on ANNOVAR (Wang et al., 2010). ERDS (v1.1) (Zhu et al., 2012) and CNVnator (v0.3.2) (Abyzov et al., 2011) were used to call CNVs, and a custom annotation and prioritization pipeline was used to define rare CNVs Reuter et al., 2018. We used Mutect2 (v3.7-0, default parameters) to detect somatic variants in the reprogrammed cell lines that were not found in blood (Cibulskis et al., 2013). Variants were annotated using an ANNOVAR based pipeline (Wang et al., 2010).

We downloaded genome sequencing data (fastq files) of five publicly available iPS cell lines: HPSI0114i.kolf_2, HPSI0214i-wibj_2, HPSI0214i-kucg_2, HPSI03145i-hoik_1, HPSI0314i-sojd_3 (http://www.hipsci.org/lines/#/lines, https://www.ebi.ac.uk/ena). We performed data analysis, variant prioritization and interpretation as described previously (Reuter et al., 2018).

CRISPR-Cas9 gene editing

pSpCas9(BB)-2A-Puro (PX459) V2.0 (Addgene #62988) was cloned with oligos designed to target MYBPC3 exon 24 using benchling.com CRISPR prediction tool (Benchling, 2019). Indels were introduced into iPSCs by nucleofecting 2x10^6 cells with 5 µg plasmid using Nucleofector 2b (Lonza) with program A-23 (~24% transfection efficiency). Transfected cells were split to multiple wells of a 6-well plate. On D2, puromycin (0.5 µg/ml) was added to daily mTeSR1 changes until D5. Surviving single colonies were grown to D12-18 to ensure they were large enough to isolate and transfer to 24-well plates. Isolated clones were passaged and genomic DNA (gDNA) was harvested from residual cells after the first 6-well passage with Quick DNA Miniprep kit. gDNA was genotyped by submitting PCR products (Table S7) for Sanger sequencing by TCAG.

Subsequent editing of other genes was performed with 8x10^5 cells and 1.5 µg plasmid in 100 µl scale Neon Transfection System (ThermoFisher) with one pulse at 1500 millivolts for 30 milliseconds (~40-70% transfection efficiency). Nucleofected cells were plated in 2 wells of a 6-well plate with mTeSR1 media supplemented with CloneR (STEMCELL Technologies) to enhance cell survival.

Ngn2-Lentivirus preparation

Two T-75 flasks were seeded with 7.5x10^6 HEK293-T cells and grown in DMEM (Gibco) supplements with 10% fetal bovine serum (Gibco) and antibiotics penicillin and streptomycin. Cells were transfected 24 hours post seeding with Lipofectamine 2000 (Thermofisher) with 10 µg gag-pol, 10 µg rev, 5 µg VSV-G and 15 µg \textit{Ngn2} or \textit{rtTA} plasmids (Hotta et al., 2009). Media was removed 18 hours later and increased to 25 ml per dish. The following day, supernatant containing viral particles were collected, filtered and concentrated by 91,000 g centrifugation for 2 h at 4°C. The supernatant was discarded and 50 µl HBSS was added to the pellet and left overnight at 4°C. The following day, aliquots of 10 µl were frozen at -80°C.

Differentiation and maintenance of excitatory cortical neurons

Excitatory cortical neurons were generated as previously described by Deneault et al. 2018. In brief: 5x10^5 iPSCs/well were seeded in Matrigel-coated 6-well plate in 2 ml of mTeSR1 supplemented with 10 µM Y-27632. The following morning media in each well was replaced with fresh mTeSR1 plus 10 µM Y-27632, 0.8 µg/ml polybrene (Sigma), and minimal amounts of \textit{Ngn2} and \textit{rtTA} lentiviruses necessary to generate >90% GFP+ cells.
upon doxycycline induction. Viral volumes were determined for each virus batch by titration and flow cytometry measurement of GFP positivity. The day after transduction, virus-containing media were replaced with fresh mTeSR1, and cells were expanded to 80-90% confluency. Ngn2-iPSCs were expanded once to freeze aliquots of infected cells. Ngn2-iPSCs reliably generated neurons up to passage seven.

Ngn2-iPSCs were dissociated using Accutase and seeded in Matrigel-coated 6-well plates at a density of 5x10^5 cells per well in 2 ml of mTeSR1 supplemented with 10 µM Y-27632 (D0). D1, media in each well was changed for 2 ml of CM1 [DMEM-F12 (Gibco), 1x N2 (Gibco), 1x NEAA (Gibco), 1x pen/strep (Gibco), 1 µg/ml laminin (Sigma), 10 ng/ml BDNF (Peprotech) and 10 ng/ml GDNF (Peprotech)] supplemented with 2 µg/mL doxycycline hyclate (Sigma) and 10 µM Y-27632. D2, media was replaced with 2 ml of CM1 supplemented with 2 µg/ml doxycycline hyclate and 2 µg/ml puromycin (Sigma). D3 media was changed to CM2 [Neurobasal media (Gibco), 1x B27 (Gibco), 1x Glutamax (Gibco), 1x pen/strep, 10 ng/ml BDNF and 10 ng/ml GDNF] supplemented with 2 µg/ml puromycin and 1 µg/ml laminin. The same media change was repeated at D4 without puromycin. D6, media was replaced with CM2 supplemented with 2 µg/ml doxycycline hyclate, 10 µM Ara-C (Sigma), 10 ng/ml BDNF, 10 ng/ml GDNF, and 1 µg/ml laminin. D8 post-Ngn2-induced neurons were dissociated with Accutase filtered through a 70 µm filter and seeded for subsequent experiments, as described below.

**Seeding neurons for imaging**

Two days before seeding neurons, 24-well black µ-Plates were coated overnight with poly-L-ornithine followed by an overnight coating with 40 µg/ml laminin. 3.5x10^5 neurons were seeded in each well in CM2 media and 24 hours later 7x10^4 mouse astrocytes/well were seeded on top of neurons. Astrocytes were prepared from postnatal day 1 CD-1 mice as described (Kim and Magrane, 2011). Before use, astrocytes were checked for mycoplasma contamination. All animal work was approved by The Hospital for Sick Children Animal Care Committee and complies with the guidelines established by The Canadian Council of Animal Care. CM2 media was replaced twice weekly. 48 hours prior to fixation and staining neuron we sparsely transfected each well with 1 µg EF1a-EGFP plasmid (Djuric et al., 2015) carried by Lipofectamine 2000 in Opti-MEM. 16-18 hours post transfection media was removed and replaced with CM2 media. Neurons were prepared and imaged as described above.

**Tracing neurons**

Blinded images were analyzed using Imaris (Bitplane). Somas and dendrites were manually traced using the GFP channel. Axonal tracing was prevented by using the MAP2 channel to identify dendrites. 40 neurons were measured from each cell line (20 technical replicates; 2 separate differentiations).

**Seeding neurons for MEA recording**

48-well opaque- or clear-bottom MEA plates (Axion Biosystems) were coated with filter sterilized 0.1% PEI solution in borate buffer pH 8.4 for 1 hour at room temperature, washed four times with water, and dried overnight. 5x10^4 D8 neurons were seeded per well as 40 µl droplets in CM2 [BrainPhys (STEMCELL Technologies) supplemented with BDNF, GDNF] with 400 µg/µl laminin and allowed to sit for 1 hour at 37°C before adding 300 µl CM2 supplemented with 40 ug/ul laminin. 24 hours later, 1x10^4 mouse astrocytes/well were seeded on top of neurons in CM2. Media was changed twice a week on the same days of the week. Once a week from week four to seven, electrical activity of the MEA plates was recorded using the Axion Maestro MEA reader (Axion Biosystems). Heater control was set to 37°C. Each plate was incubated for five minutes on the warmed reader, then real-time spontaneous neural activity was recorded for five minutes using AxIS 2.0 software (Axion Biosystems). A bandpass filter from 200 Hz to 3 kHz was applied. Spikes were detected using a threshold of 6 times the standard deviation of noise signal on electrodes. Offline advanced metrics were re-recorded and analyzed using Axion Biosystems Neural Metric Tool. Electrodes were considered active if at least 5 spikes were detected per minute. Bursts and network bursts were detected by Poisson surprise and Envelope (25% active electrode threshold) algorithms, respectively. No non-active well was excluded in the analysis. After the last reading at week seven, selected wells were treated with a synaptic antagonist to AMPA receptor: 6-cyano-7-nitroquinoxaline-2,3-dion (CNQX; Sigma) at 60 µM. Plates were recorded 15 minutes after addition of the antagonists. Culture media was replaced with CM2 and after a one-hour recovery period a final recording was taken.

**Statistical analysis of neurons**

All statistical tests were conducted in RStudio (V1.1.456). Comparisons between soma area and dendritic length were conducted using Kruskal-Wallis tests after determining non-normal distribution of the data by qqplots and a Shapiro test. Comparisons between cell lines at each distance of Sholl analysis was conducted using Dunn’s test when finding statistical significance by a Kruskal-Wallis test.
Differentiation and maintenance of CMs

Cardiomyocytes were generated using STEMdiff Cardiomyocyte Differentiation Kit (STEMCELL Technologies). iPSCs were dissociated with Accutase and seeded at 8x10^6 cells per Matrigel-coated well of a 12-well dish in mTesR1 supplemented with 10 µM Y-27632. 24 hours later media was replaced with mTesR1. 48-72 hours post-seeding (determined by confluency reaching 80-90%) differentiation was initiated by adding media A supplemented with Matrigel (D1) and continued as written in the manufacturer’s protocol. D16 cardiomyocytes were dissociated with STEMdiff Cardiomyocyte Dissociation Media (STEMCELL Technologies) and seeded for subsequent experiments as described below.

Flow cytometry analysis of CMs

CMs and iPSCs were prepared for flow cytometry as per Inside Stain Kit protocol (Miltenyi Biotec). Cells were stained with recombinant (REA400) FITC-conjugated anti-cardiac Troponin T (1:10) for 10 minutes. iPSCs were used as negative controls with fluorescent-minus-one (FMO) controls used to set gating parameters. Samples were analyzed on an LSR II (BD Biosciences).

Calcium Imaging of CMs

D16 contractile cardiomyocytes were dissociated and 6x10^5 cells per well were seeded in Matrigel-coated 96-well plates in STEMdiff Cardiomyocyte Support Media (STEMCELL Technologies). The following day the media was replaced with STEMdiff Cardiomyocyte Maintenance Media and every two days thereafter. D34, cells were treated with 1 µM Fluo-4 AM (Invitrogen) in HBSS solution for 30 minutes at 37°C. Following Fluo-4 loading, cells were washed one time with HBSS solution. Imaging was carried out with a Nikon TE2000 microscope at 37°C with a 10x lens using Volocity software. Images were captured with 488 nm excitation and 516 nm emission at a rate of four images per second for 30 seconds. Image stacks were converted to videos with Volocity software (Quorum Technologies) and regions of interest were analyzed for changes in fluorescence intensity (f-f0)/f0, with the resting fluorescence value F0 set as the minimal value recorded. Background intensity was subtracted from all values, and plots were normalized to zero fluorescence. Calcium transient amplitude and beat rate were calculated with Volocity software.

xCELLigence RTCA data collection

xCELLigence E-plates were coated with fibronectin in water for one hour prior to cell seeding. Before adding cells, a baseline recording was taken while only media was present in the well. 4x10^5 dissociated CMs were seeded per well, then the plate was transferred to the device in a humidified incubator at 37°C, 5% CO2 and normoxic conditions. Twenty second sweeps were collected every three hours with a Cardio speed of 2.0 ms and ECR speed of 0.1 ms using RTCA CardioECR Software without stimulation (V1.2.0.1603; ACEA Biosciences). Recordings were paused every two days to change media to fresh maintenance media.

CM statistical analyses

Statistics for Ca²⁺ imaging or xCELLigence assays were calculated using RStudio. Normality was checked by ggqqplot and Shapiro.test. If data was normally distributed, analysis of variance was calculated followed by Tukey ANOVA followed by Tukey’s test. Where data was non-normal Kruskal-Wallis tests were used instead and pairwise comparisons made using Dunn’s test.

Differentiation to hepatocytes

HLCs were generated using a protocol adapted from (Ogawa et al., 2015). All PGPC human iPSC lines were maintained in mTeSR1 as previously described. Prior to the induction of endoderm in the monolayer cultures, human iPSCs were passaged onto a Matrigel coated surface for 1 day with Gentle Cell Dissociation Reagent (GCDR; STEMCELL Technologies) as clumps in mTeSR1 based medium supplemented with 10 µM ROCK Inhibitor Y-27632 (D0). On D1, differentiation towards definitive endoderm was carried out using STEMdiff Definitive Endoderm Kit (STEMCELL Technologies) for four days. D5, the resultant definitive endoderm (DE) cells in one of the 6-well-plate were reseeded as clumps with GCDR in 12 wells of 24-well-plate cultured in serum-free-differentiation (SFD)-based medium containing ROCKi for 24 hours (supplemented with N2, ascorbic acid, MTG, 0.1% BSA, glutamine) with B27 and 5 ng/ml basic fibroblast growth factor (bFGF) for two days and 100 ng/ml activin A for four days. The media was changed every other day. By D8, DE was specified to foregut progenitor (FP) cells.

Bipotent hepatoblasts (HBs) were generated by culturing FPs in low glucose DMEM plus B27 supplemented with 40 ng/ml bFGF, 10 µM SB-431542 and 50 ng/ml bone morphogenic protein (BMP4) from day 9 to day 15, media changed every other day. To promote expansion and maturation of the hepatoblast population, the HBs were cultured in a mixture of low glucose DMEM / Ham’s F12 (3:1) media with 0.1% BSA, 1% vol/vol
B27+retinoic acid (RA) supplement, ascorbic acid, glutamine, MTG, 20 ng/ml hepatocyte growth factor (HGF), 40 ng/ml dexamethasone (Dex) and oncostatin M (OSM) for 7 days (D15-21). On D22, cells were cultured in a mixture of high glucose DMEM / Ham’s F12 (3:1) with 0.1% BSA, 1% vol/vol B27+RA supplement, ascorbic acid, glutamine, MTG, 20 ng/ml HGF, 40 ng/ml Dex and 20 ng/ml OSM for 4 days to generate hepatocyte like cells (HLCs) at D25. The media was changed every other day.

**Flow cytometry analysis during hepatocyte-like cell differentiation**

Cells were prepared for flow cytometry by dissociating to single cells with TrypLe, washed then fixed with 4% PFA. Cells were permeabilized with Perm/Wash buffer (BD Biosciences) then labeled with primary antibodies at concentrations listed in antibody table below. Isotype controls were used to set gating parameters. Samples were analyzed on an LSR II (BD Biosciences).

**Cytochrome p450 activity**

CYP3A7 activity in HLCs was measured using the p450 Glo assay kit (Promega) according to the manufacturer’s instructions and a P450-Glomax 96 microplate luminometer (Promega). HBs were used as negative controls and 1 µM ketoconazole was used as an inhibitor of CYP3A7.

**Differentiation to kidney organoids**

Kidney organoids were made according to the protocol described by Takasato et al., 2015 but with minor modifications similar to those implemented in more recent publications (van den Berg et al., 2018; Forbes et al., 2018). Single-cell suspensions of iPSCs were made using Accutase and cells were seeded at densities of 7,500-12,000 cells/cm² on freshly prepared Matrigel-coated plates with mTetSR1 containing 10 µM Y-27632. Differentiation was started (D0) by changing medium to STEMdiff APEL-2 (STEMCELL Technologies) supplemented with 8 µM CHIR99021. Medium was fully replaced every other day until D5 where CHIR-medium was switched to APEL-2 supplemented with 200 ng/ml human fibroblast growth factor (hFGF)-9 and 1 µg/ml heparin. iPSCs and differentiated cells were harvest for mRNA analysis at indicated time-points (D0, 4 and 7). D7 cells were dissociated using Trypsin-EDTA for 1 minute at 37°C. Two milliliters of DMEM:F12 supplemented with 10% fetal bovine serum was added to inactivate Trypsin and cells were spun down at 300 g for 4 minutes. Supernatants were removed and cell pellets were resuspended in APEL-2 medium without differentiation factors. Aliquots of suspension containing 5x10⁵ cells were transferred to 1.5 ml microcentrifuge tubes and aggregated by centrifugation at 400 g for 3 minutes. Three to four aggregates were carefully transferred to 6-well Transwell cell culture plates using a wide-bore P200 pipette tip with minimal carry-over of medium. Initial Transwell medium consisted of APEL-2 with 5 µM CHIR99021. Medium was switched to APEL-2 with 200 ng/ml hFGF9 and 1 µg/ml heparin after one hour. Medium was replaced every other day thereafter. From D12 onward, APEL-2 was supplemented with heparin only.

**Immunostaining kidney organoids**

D25 organoids were fixed in 2% paraformaldehyde as described previously (Takasato et al., 2016). Organoids were embedded in paraffin and sectioned (6 µm sections) by the Core Pathology laboratory at The Centre for Phenogenomics (Toronto, ON, Canada). The immunofluorescence protocol was described previously (Rowan et al., 2018). In brief: sections were stained with the primary antibodies against Wilms Tumor 1 (WT1) (1:100) and E-Cadherin (ECAD (1:200) with appropriate secondary antibodies (Table S8). In addition, fluorescein-labeled Lotus Tetragonolobus lectin (LTL, 1:200) was used to stain the proximal tubule segments. LTL was added with secondary antibodies. After mounting, sections were imaged using an epifluorescence microscope.

**mRNA expression analysis of kidney organoid and HLC differentiations**

mRNA was isolated from iPSCs and differentiated cells using RNeasy Mini kit (Qiagen) or TRIzol (ThermoFisher). mRNA concentrations were determined using Nanodrop 2000C. mRNA was reverse transcribed using SuperScript First-Strand kit. Quantitative real-time PCR was performed using the Viia7 and PowerSYBR Green PCR technology (kidney; ThermoFisher) or Advanced qPCR Mastermix with Supergreen Lo-Rox (HLC; Wisent). Sequences of primers used were either derived from previous publications or newly designed using Nucleotide Blast (https://blast.ncbi.nlm.nih.gov/Blast.cgi). Quantitation was performed using the standard curve method with serial dilutions of pooled cDNA of all samples. mRNA expression levels were normalized to ribosomal 18S or TBP expression and expressed relative to reference group as indicated.

**T-cell differentiation**

T-cells were generated as previously described (Kennedy et al., 2012; La Motte-Mohs et al., 2005). PGPC lines and the commercially available iPSC11 (Alstem Cell Advancements) were cultured on plates coated with
Matrigel with TeSR-E8 medium (STEMCELL Technologies). iPSCs were first differentiated to generate CD34+ hemogenic endothelial (HE) progenitor cells (as described previously in Kennedy et al., 2012), which were isolated by staining with anti-human CD34-PE (BD) and anti-PE-microbeads (Miltenyi) and enriched using magnetic-activated cell sorting (MACS) columns (Miltenyi). CD34+ cells were then co-cultured with OP9-DL4 cells to direct differentiation towards T-lineage cells as described previously in La Motte-Mohs, 2005. T-lineage cells were passaged onto new OP9-DL4 cells every 4-5 days and analyzed by flow cytometry using the following T-lineage markers: CD7, CD5, CD8a, CD1a, CD4, CD45, CD3, TCRγ, TCRδ, CD34 and CD8b. FMO controls were used to set gating parameters. Flow cytometry data was collected on LSR-II (BD) and analyzed using FlowJo.

Culture conditions and neural induction of sensory neurons

Primary sensory neurons (PSNs) were generated using a protocol modified from Chambers et al., 2012. We found improved growth and survival with the addition of NGF at D2 and using N2 media with the following composition: 47.5% DMEM/F12, 47.5% neurobasal, 2% B-27 supplement, 1% N2 supplement, 1% glutamax and 1% Pen/Strep. iPSCs (PGP17_11) were dissociated to single cells with Accutase and plated on Matrigel coated plates at 3.1x10^5 - 3.7x10^5 cells/cm^2 in mTeSR1 + 10 µM Y-27632. Upon reaching confluence, neural differentiation was initiated (D0) by the addition of 100 nM LDN-193189 and 10 µM SB431542 in mTeSR1 media. On D2, media was replaced with mTeSR1 containing 100 nM LDN-193189, 10 µM SB431542, 3 µM CHIR99021, 10 µM SU5402, 10 µM DAPT, 10 ng/ml NGF and 1 µg/ml laminin. On D4, media was replaced with a 75% mTeSR1 / 25% N2 media (47.5% DMEM/F12, 47.5% neurobasal, 2% B-27 supplement, 1% N2 supplement, 1% Glutamax and 1% Pen/Strep) combination containing 100 nM LDN-193189, 10 µM SB431542, 3 µM CHIR99021, 10 µM SU5402, 10 µM DAPT, 10 ng/ml NGF and 1 µg/ml laminin. On D6, media was replaced with a 50% mTeSR1 / 50% N2 media combination containing 3 µM CHIR99021, 10 µM SU5402, 10 µM DAPT, 10 ng/ml NGF and 1 µg/ml laminin. On D8, media was replaced with a 25% mTeSR1 / 75% N2 media combination containing 3 µM CHIR99021, 10 µM SU5402, 10 µM DAPT, 10 ng/ml NGF and 1 µg/ml laminin. On D10, 100% N2 media was used containing 3 µM CHIR99021, 10 µM SU5402, 10 µM DAPT, 10 ng/ml NGF and 1 µg/ml laminin. On D12, sensory neurons were dissociated with Accutase, washed with DMEM/F12 + 0.1% BSA and triturated until no visible clumps remained. The cell suspension was then filtered through a 70 µm filter, pelleted at 1100 rpm for 5 min and resuspended in N2 media containing 10 ng/ml NGF, 10 ng/ml BDNF, 10 ng/ml GDNF, 10 µg/ml laminin and 10 µM Y-27632. Neurons were reseeded onto plastic coverslips coated with 1 mg/ml polyornithine overnight followed by 10 µg/ml fibronectin and 40 µg/ml laminin also overnight at 2x10^5 – 2.5x10^5 cells/cm^2 with media changed using the above combination twice weekly.

Patch-clamp recordings

Neurons were recorded in whole cell configuration with 70% series resistance compensation using an Axopatch 200B (Molecular Devices, Sunnyvale, CA, USA). Patch clamp recordings were performed at room temperature with traces acquired at 25 kHz and low-pass filtered at 2 kHz. Data was digitized with a power1401-3A data acquisition interface (CED Ltd., Cambridge, England). Artificial cerebrospinal fluid (aCSF) contained the following in (mM): 126 NaCl, 2.5 KCl, 2 CaCl2, 2 MgCl2, 10 D-Glucose, 26 NaHCO3, 1.25 NaH2PO4 bubbled with 95% oxygen and 5% carbon dioxide, and was continuously perfused over the coverslips at a rate of 2 ml/min. Borosilicate glass microelectrodes (World Precision Instruments, Sarasota, FL USA) with a tip resistance of 4-6 MΩ were filled with solution containing the following (in mM): 125 KMeSO4, 5 KCl, 10 HEPES, 2 MgCl2, 4 Na2ATP and 0.4 Tris-GTP (ThermoFisher Scientific); pH was adjusted with KOH to 7.2. All membrane potential values were corrected for a junction potential of 9 mV. Rheobase and spiking patterns were assessed at a holding potential of -65mV. Repetitive spiking neurons were defined as those that responded to a supra-threshold current injection with 3 or more spikes. Cells which showed a resistance change greater than 20% over the course of the recording were rejected from subsequent analysis.

Ca^{2+} imaging and analysis of PSNs

Ca^{2+} imaging was performed on neurons incubated in 5 µM Ca^{2+} green-1 AM dye (ThermoFisher Scientific, Waltham, MA, USA) at room temperature for 20 minutes in oxygenated, HEPES buffered ACSF. Images were acquired using a NeuroCCD-SM256 imaging system (RedShirt Imaging, Decatur, GA) mounted on a Zeiss Examiner A1 with dichroic filter set 46HE. Images were acquired at a rate of 25 Hz with a 10X digital gain. Image analysis was performed by drawing a region of interest around individual neurons using Neuropylex software and exporting the raw fluorescence values to SigmaPlot for subsequent analysis. Healthy cells were identified by application of 500mM KC1 using a Toohey Spritzer pressure system (Ile fluid delivery system (Fairfield, NJ) at 5 PSI with a 4-5 second pulses. The tip of the puff pipette was positioned 50-70 µm from the cell group and 5 µm above the coverslip surface while undergoing Ca^{2+} imaging. Neurons that showed a robust rise in cytosolic Ca^{2+} to KC1
puffs were considered healthy. Phenotyping was performed on these cells with the same picospritzer settings using 3 mM GABA, 20 µM capsaicin and 200 µM ATP in HEPES buffered ACSF. To differentiate ionotropic from metabotropic GABA receptors, 100 µM picrotoxin was bath applied following the initial identification of GABA responsive cells. All drugs were acquired from Sigma-Aldrich.

**Statistical analysis of PSNs**

Comparisons between time points and spiking patterns were conducted using Mann-Whitney U ranked sum tests. Chi Square tests were performed to assess developmental differences in agonist sensitivities and spiking patterns.
### Primer list

| Target                          | Forward primer                                      | Reverse primer                                      |
|--------------------------------|-----------------------------------------------------|-----------------------------------------------------|
| Sendai virus detection          | GGATCACTAGGTGATATCGAGC                              | ACCAGACAAGAGTTAAGAGATATGATATC                       |
| AR                             | FAM|CGTGCGCGAAGTGATCCAGA                               | GTTTTTGTCTGCTGCTGAGGCTAGT                          |
| MYBPC3 CRISPR/Cas9 targeting   | CACCGACTCTCTGACAGTACAGT                             | AAAACACTGTACTGTGACAGGAGTCC                          |
| MYBPC3 (gDNA)                  | CTGTTGCGGTAGTGTAGAGT                              | TAGACGCAGCATCTCGTACA                                 |
| TBP                            | TAAGAGAGCCAGCAACCAGC                               | TTDCTACACAGAGCCCAGAC                                |
| GATA6                          | GCCACTACCTGTGCAACGCT                               | CAATCCAGCCCGCCTGATGAAA                              |
| FOXA2                          | ATGCACTCGGCTTCCAGTATG                              | TGTGCTCACGGAAGAGTAGC                                |
| HNF5                           | GGGTTGCTCTAGATCCAGC                                | CCGTTGCTTACGTACCATG                                 |
| HNF4A                          | CATGGCCAAGATTGACACCT                               | TCCCCATATGTCTGCTGACAG                               |
| CYP3A7                         | GACGGGCTTCCATCCATG                             | TGGGGGTGGTGAGATAGC                                  |
| CEBPA                          | GGAGCTGAGATCCCGACAG                               | GCCTGGAGTAGACGGCCAC CTCA                           |
| ALB                            | CTTTTGGCCACAATGAAATGGTTGACACC                      | CAGCACTAGCCATTACCTACATAG                             |
| AFP                            | AGAAACCTGTCTACAAGGTCTG                                | GACAGCAAGCTGAGATTG                                 |
| A1AT                           | ACATGTGAGCACGGAGAAC                                | TTTGTTGCACTAGAGTGCT                                 |
| 18S                            | GATGGGCGGCAGGAAATAG                                 | GCGTTGATCTGCTAAATAGT                                |
| GATA3                          | GCCCTCATTAGAGGCCAAG                                | TTTGCTGCTGAGTACGCT                                 |
| HOXD11                         | GCCAGTGTGCTGCTGCTGCT                                 | CTTCTACAGACCACCACGT                                  |
| T                              | AGTACCGCAACCTGTAGGGA                               | GCAAGTGAGTGTGCTAGAATAGT                              |
| NPHS1                          | FATTCCGCGCAAACCTAGT                                 | GCCTGAGTCTCTGCTACCAT                                 |
| SLC3A1                         | TCGCTCAAGTCACAAATGC                                | GCTGAGTCTCTTGGGACATCAACAT                           |
| SLC21A1                        | GGAATGGGGTAGAGTTGCT                                 | GAACCTCAAGACCAATCCAGC                               |
| PECAM                          | CATTACGGTCACAATGACGAGT                              | ATCTGGTTCTGACCAATCAT                                 |
| FOXD1                          | FTTGGGACTCTGCGACCCAGG                               | FTTCTGAGCGCGCTAATCAT                                 |
| Target        | Species | Concentration (Technique) | Company                | Catalogue number |
|---------------|---------|----------------------------|------------------------|------------------|
| SSEA4         | Mouse   | 1:100 (ICC)               | ThermoFisher Scientific | 41-4000          |
| TRA1-60       | Mouse   | 1:100 (ICC)               | ThermoFisher Scientific | 41-1000          |
| NANOG         | Rabbit  | 1:400 (ICC)               | Cell Signaling Technology | 4903             |
| OCT4          | Rabbit  | 1:200 (ICC)               | Abcam                  | ab19857          |
| TUBB3         | Mouse   | 1:200 (ICC)               | Milipore Sigma         | MAB1637          |
| SMA           | Mouse   | 1:200 (ICC)               | ThermoFisher Scientific | 18-0106          |
| AFP           | Mouse   | 1:200 (ICC)               | R&D Systems            | MAB1368          |
| MAP2          | Guinea pig | 1:1000 (ICC)          | Synaptic Systems       | 188 004          |
| GFP           | Chicken | 1:500 (ICC)               | ThermoFisher Scientific | A-10262          |
| cTNT-FITC conjugate | Recombinant human | 1:10 (Flow)  | Miltenyi Biotec       | 130-106-687      |
| MLC2V         | Rabbit  | 1:400 (ICC)               | Abcam                  | Ab79935          |
| MYBPC3        | Rabbit  | 1:1000 (WB)               | Abcam                  | ab171153         |
| Actinin       | Mouse   | 1:200 (ICC)               | Abcam                  | Ab9465           |
| ALB           | Rabbit  | 1:50 (ICC/Flow)           | DAKO                   | F0117            |
| Alpha-1-AT    | Goat    | 1:100 (Flow)              | Bethyl                 | A80-122A         |
| AFP-APC conjugate | Mouse     | 1:200 (Flow)         | Sino Biological Inc.   | 121177-MM35-A    |
| CYP3A7        | Mouse   | 1:100 (ICC/Flow)          | NOVUS                  | BP2-37502SS      |
| HNF4A         | Rabbit  | 1:100 (ICC)               | Cell Signaling         | 3113S            |
| WT1           | Mouse   | 1:100 (IHC)               | DAKO                   | M3561            |
| ECAD          | Rabbit  | 1:200 (IHC)               | Cell Signaling Technology | 5195            |
| Lotus Tetragonolobus lectin (Fluorescein-labeled) | N/A     | 1:200 (IHC)               | Vector Laboratories   | FL-1321          |
| Anti-mouse IgG (AlexaFluor 555) | Goat    | 1:1000 (ICC)              | ThermoFisher Scientific | A21422          |
| Anti-mouse IgM (AlexaFluor 555) | Goat    | 1:1000 (ICC)              | ThermoFisher Scientific | A21426          |
| Anti-rabbit (AlexaFluor 488) | Goat    | 1:1000 (ICC)              | ThermoFisher Scientific | A11008          |
| Anti-mouse (AlexaFluor 488) | Goat    | 1:1000 (IHC)              | ThermoFisher Scientific | A21042          |
| Anti-rabbit (AlexaFluor 405) | Goat    | 1:1000 (IHC)              | ThermoFisher Scientific | A31553          |
| Anti-rabbit (AlexaFluor 568) | Donkey  | 1:1000 (IHC)              | ThermoFisher Scientific | A10042          |
| IRDye 800CW anti-rabbit | Goat    | 1:20000 (WB)              | LI-COR Biosciences     | 925-32211        |
| CD7-FITC      |        | 1:100 (Flow)              | BioLegend              | 343104           |
| CD5-PerCP/Cy5.5 |      | 1:100 (Flow)              | BioLegend              | 300620           |
| CD8a-PE/CF594 |        | 1:100 (Flow)              | BioLegend              | 301058           |
| CD1a-APC      |        | 1:100 (Flow)              | BioLegend              | 300110           |
| CD4-AF700     |        | 1:100 (Flow)              | BioLegend              | 300526           |
| CD45-APC/Cy7  |        | 1:100 (Flow)              | BioLegend              | 304014           |
| CD3-BV510     |        | 1:100 (Flow)              | BioLegend              | 317332           |
| TCRa-b-APC    |        | 1:100 (Flow)              | BioLegend              | 306718           |
| TCRg-y-FITC   |        | 1:100 (Flow)              | BioLegend              | 331208           |
| CD34-PE       |        | 1:25 (Flow)               | BD                     | 550761           |
| CD8b-PE/Cy7   |        | 1:100 (Flow)              | eBiosciences           | 25-5273-42       |
Abyzov, A., Urban, A.E., Snyder, M., and Gerstein, M. (2011). CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. Genome Res. 21, 974–984.

Van der Auwera, G.A., Carneiro, M.O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. Curr. Protoc. Bioinforma. 43, 11.10.1-11.10.33.

Benchling (2019). Benchling [Biology Software].

van den Berg, C.W., Ritsma, L., Avramut, M.C., Wiersma, L.E., van den Berg, B.M., Leuning, D.G., Lievers, E., Koning, M., Vanslambrouck, J.M., Koster, A.J., et al. (2018). Renal Subcapsular Transplantation of PSC-Derived Kidney Organoids Induces Neo-vasculogenesis and Significant Glomerular and Tubular Maturation In Vivo. Stem Cell Reports 10, 751–765.

Cibulskis, K., Lawrence, M.S., Carter, S.L., Sivachenko, A., Jaffe, D., Sougnez, C., Gabriel, S., Meyerson, M., Lander, E.S., and Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat. Biotechnol. 31, 213–219.

DePristo, M.A., Banks, E., Poplin, R., Garimella, K. V, Maguire, J.R., Hartl, C., Philippakis, A.A., del Angel, G., Rivas, M.A., Hanna, M., et al. (2011). A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat. Genet. 43, 491–498.

Djuric, U., Cheung, A.Y.L., Zhang, W., Mok, R.S., Lai, W., Pickna, A., Hendry, J.A., Ross, P.J., Pasceri, P., Kim, D.S., et al. (2015). MECP2e1 isoform mutation affects the form and function of neurons derived from Rett syndrome patient iPS cells. Neurobiol. Dis. 76, 37–45.

Forbes, T.A., Howden, S.E., Lawlor, K., Phipson, B., Maksimovic, J., Hale, L., Wilson, S., Quinlan, C., Ho, G., Holman, K., et al. (2018). Patient-iPSC-Derived Kidney Organoids Show Functional Validation of a Ciliopathic Renal Phenotype and Reveal Underlying Pathogenetic Mechanisms. Am. J. Hum. Genet. 102, 816–831.

Hotta, A., Cheung, A.Y.L., Farra, N., Garcha, K., Chang, W.Y., Pasceri, P., Stanford, W.L., and Ellis, J. (2009). EOS lentiviral vector selection system for human induced pluripotent stem cells. Nat. Protoc. 4, 1828–1844.

Kennedy, M., Awong, G., Sturgeon, C.M., Ditadi, A., LaMotte-Mohs, R., Zúñiga-Pflücker, J.C., and Keller, G. (2012). T lymphocyte potential marks the emergence of definitive hematopoietic progenitors in human pluripotent stem cell differentiation cultures. Cell Rep. 2, 1722–1735.

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.

La Motte-Mohs, R.N. (2005). Induction of T-cell development from human cord blood hematopoietic stem cells by Delta-like 1 in vitro. Blood 105, 1431–1439.

La Motte-Mohs, R.N., Herer, E., and Zúñiga-Pflücker, J.C. (2005). Induction of T-cell development from human cord blood hematopoietic stem cells by Delta-like 1 in vitro. Blood 105, 1431–1439.

Müller, F.-J., Schulte, B.M., Williams, R., Mason, D., Altun, G., Papapetrou, E.P., Danner, S., Goldmann, J.E., Herbst, A., Schmidt, N.O., et al. (2011). A bioinformatic assay for pluripotency in human cells. Nat. Methods 8, 315–317.

Ogawa, M., Ogawa, S., Bear, C.E., Ahmadi, S., Chin, S., Li, B., Grompe, M., Keller, G., Kamath, B.M., and Ghanekar, A. (2015). Directed differentiation of cholangiocytes from human pluripotent stem cells. Nat. Biotechnol. 33, 853–861.

Rowan, C.J., Li, W., Martirosyan, H., Erwood, S., Hu, D., Kim, Y.-K., Sheybani-Delouis, S., Mulder, J., Blake, J., Chen, L., et al. (2018). Hedgehog-GLI signaling in Foxd1-positive stromal cells promotes murine nephrogenesis via TGfβ signaling. Development 145, 1–13.

Takasato, M., Er, P.X., Chiu, H.S., and Little, M.H. (2016). Generation of kidney organoids from human pluripotent stem cells. Nat. Protoc. 11, 1681–1692.
Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. 38, e164–e164.

Yuen, R.K.C., Thiruvahindrapuram, B., Merico, D., Walker, S., Tammimies, K., Hoang, N., Chrysler, C., Nalpathamkalam, T., Pellecchia, G., Liu, Y., et al. (2015). Whole-genome sequencing of quartet families with autism spectrum disorder. Nat. Med. 21, 185–191.

Zhu, M., Need, A.C., Han, Y., Ge, D., Maia, J.M., Zhu, Q., Heinzen, E.L., Cirulli, E.T., Pelak, K., He, M., et al. (2012). Using ERDS to infer copy-number variants in high-coverage genomes. Am. J. Hum. Genet. 91, 408–421.