# 2 Molecular and Functional Variation in iPSC-Derived Sensory Neurons

## *Collaboration note*

The work described in this chapter has been accepted (pending revisions) for publication by Nature Genetics (Schwartzentruber et al., 2017). Supplementary Tables are available at: https://github.com/js29/ipsdsn. Daniel Gaffney and Alex Gutteridge conceived and directed the project. Stefanie Foskolou performed all differentiations from iPSCs to sensory neurons. I performed nearly all data analyses described herein. Kaur Alasoo provided some of the code used for expression quantification and QTL calling. Helena Kilpinen determined which eQTLs were tissue specific. Alex Gutteridge performed quality control and analysis of neuronal calcium imaging and electrophysiology. Anna Wilbrey prepared samples for single-cell sequencing, and Alex Gutteridge conducted preliminary analyses of the scRNA-seq data.

## 2.1 Introduction

Cellular disease models are critical for understanding the molecular mechanisms of disease and for the development of novel therapeutics. In principle, induced pluripotent stem cell (iPSC) technology enables the development of these models in any human cell type. Initial uses of iPSCs for disease modelling have focused mostly on highly penetrant, rare coding variants with large phenotypic effects (Itzhaki et al. 2011; G.-H. Liu et al. 2011; Wainger et al. 2014; G. Lee et al. 2009; Cao et al. 2016). However, there is growing interest in using iPSCs to model the effects of the common genetic variants of modest effect size that drive complex disease (Warren, Jaquish, and Cowan 2017). A key question is to what extent variability in directed differentiation is a barrier to studying the effects of common disease-associated variants in iPSC-derived cells. In addition, because cultured cells are imperfect models of primary tissues, not all common disease-associated genetic variants also alter cell phenotypes in iPSC-derived systems.

Here, we present the first large-scale study of common genetic effects in a neuronal cell type differentiated from human stem cells, iPSC-derived sensory neurons (IPSDSNs). Peripheral sensory nerve fibres innervate the skin and other organs and are brought together at the dorsal root ganglia (DRG) before synapsing with the spinal cord around the dorsal horn. The development of efficient protocols to differentiate iPSCs into nociceptive (pain-sensing) neurons (Young et al. 2014) provides the opportunity to model common genetic effects on human sensory neuron function, which may underlie individual differences in pain sensitivity and chronic pain. We investigate how power to detect common genetic effects is affected by the variability introduced by differentiation and demonstrate how initial iPSC growing conditions influence cell phenotypes in IPSDSNs. We identify quantitative trait loci (QTLs) for gene expression, RNA splicing, and chromatin accessibility and identify a number of overlaps between molecular QTLs and common disease associations. In generating this gene regulatory map we establish effective techniques for using IPSDSN cells to model molecular phenotypes relevant to common diseases.

## 2.2 Results

### 2.2.1 Sensory neuron differentiation and characterisation

We obtained 107 IPS cell lines derived from unrelated apparently healthy individuals by the HIPSCI resource (Kilpinen, Goncalves, et al. 2016), and followed an established small molecule protocol (Young et al. 2014) to differentiate these into sensory neurons of a nociceptor phenotype (Figure 1a). We performed a total of 123 differentiations; 13 of these were done with an early version of the protocol (P1) which was subsequently refined (P2) to reduce the number of differentiation failures and to yield a higher proportion of neuronal cells in the final cultures. One RNA-seq sample failed sequencing, and four others were outliers based on principal components analysis and were excluded. This left a set of 119 differentiations with gene expression data from 100 unique iPSC donors; all subsequent analyses focused on the 106 P2 protocol samples, except for QTL calling, where we used all samples to maximize discovery power.

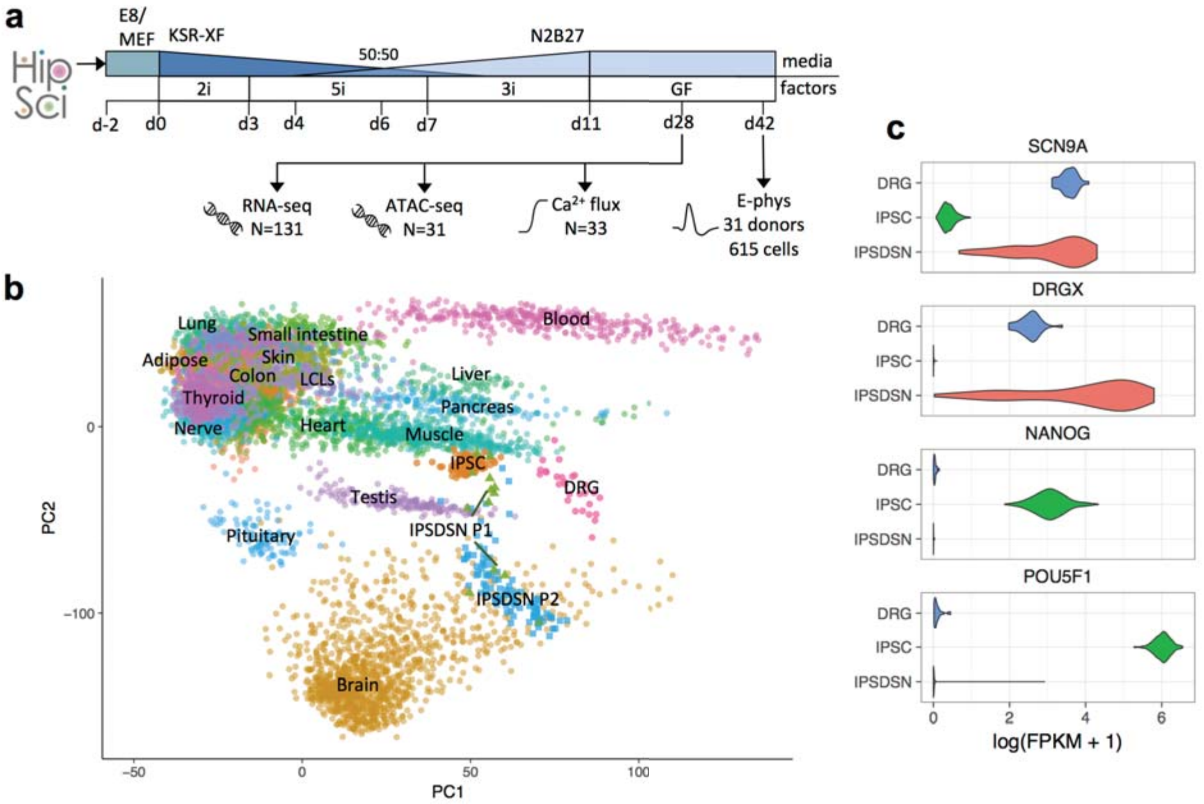


**Figure 1** Characterization of molecular phenotypes in iPSC-derived sensory neurons. (**a**) Schematic of IPSDSN differentiation and assays. iPSCs were received in Essential 8 (E8) medium (N=82) or on mouse embryonic fibroblasts (MEFs, N=49), and transferred to KSR-XF medium. Over 11 days, different inhibitor combinations were added (2i, 5i, 3i, see Methods), and N2B27 medium phased in, followed by transfer to growth factor medium at day 11 for neuronal maturation. (**b**) PCA plot projecting IPSDSN, iPSC, and DRG samples onto the first two principal components defined based on RNA-seq FPKMs in GTEx tissues. Some GTEx tissues are unlabeled due to overlapping labels. (**c**) Expression of sensory neuronal marker genes (*SCN9A*, *DRGX*) and key iPSC genes (*NANOG*, *POU5F1*).

We clustered our gene expression data with 239 iPSC samples from the many of same donors, as well as 28 post-mortem DRG tissue samples from 10 different donors, and 44 primary tissues from the GTEx project (Mele et al. 2015) (Figure 1b). Globally, IPSDSN samples showed greatest similarity to iPSCs (gene expression correlation Spearman ρ=0.89), followed by DRG (ρ=0.84), and then brain samples from GTEx. However, because different gene expression quantitation methods were used in GTEx, we cannot be certain of relative similarities between GTEx tissues and the samples we uniformly processed (DRG, IPSDSNs, iPSCs). The similarity to iPSCs may reflect lack of maturity in IPSDSNs, which is a well-recognized problem with iPSC-derived cells (Warren et al. 2017; Sala, Bellin, and Mummery 2016; Soldner et al. 2016; Pashos et al. 2017). We also note that because the same iPSCs were differentiated to IPSDSNs, both donor genetic background and cell culture effects may contribute to the observed similarity. Despite this, key sensory neuronal marker genes were highly expressed in IPSDSNs, while pluripotency genes were not (Figure 1c).

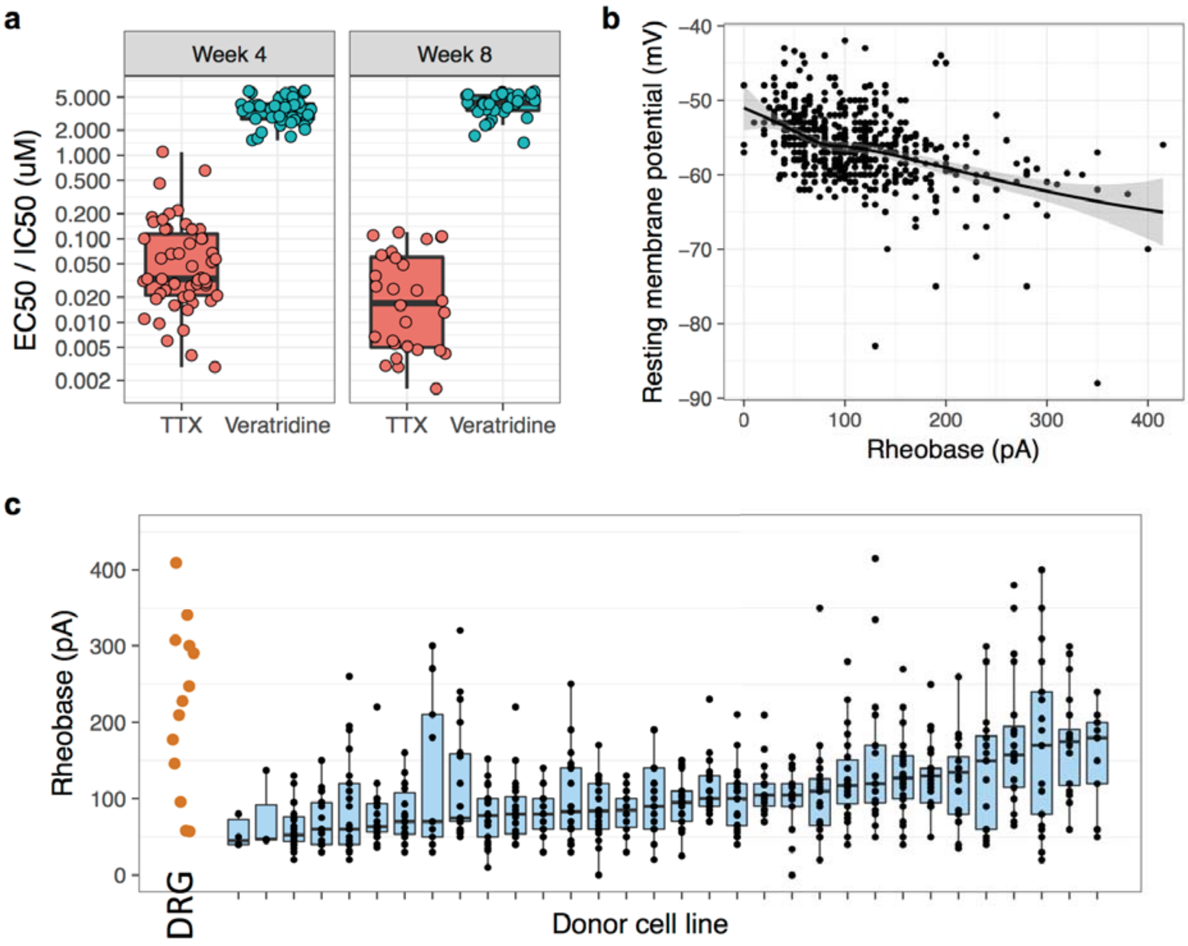


**Figure 2**: Functional characterization of IPSDSNs. (**a**) Ca2+ flux measurements on IPSDSN cultures (n=31) shows that neuronal firing is enhanced by veratridine (EC50/IC50 > 1) and is reduced by tetrodotoxin (EC50/IC50 < 1). (**b**) Rheobase is inversely related with resting membrane potential in individual neurons (n=616). (**c**) The distribution of rheobase values for the 31 samples with electrophysiology recordings, as well as literature values for DRG (leftmost bar).

Using $Ca^{2+}$ flux measurements on a subset of differentiated cultures (n=31) we confirmed that the cells consistently responded to veratridine (a sodium ion channel agonist) and tetrodotoxin (a selective sodium ion channel antagonist), as expected (Figure 2a). We also performed patch-clamp electrophysiology recordings for 616 individual neurons from 31 donors, with a median of 21 cells measured per line. The rheobase is the minimum current input that will cause an individual neuron to fire an action potential, and we used this as a measure of the overall membrane excitability. Rheobase showed the expected inverse relationship with resting membrane potential (Figure 2b). The distribution of rheobases was comparable to those obtained from primary DRG cells, but showed significant variation between donors (Figure 2c).

We next investigated whether variation in excitability was reflected in differences in gene expression of cells derived from the same donor. We examined the correlation between expression of individual genes and mean rheobase, which were measured in sister cultures from the same donor and differentiation batch. After correcting for multiple testing, no individual genes were significantly correlated with rheobase at FDR < 0.1. Similarly, none of the first five gene expression principal components were correlated with mean rheobase.

### 2.2.2 Quantifying differentiation variability using single-cell RNA-seq

Our samples appeared to differ in the fraction of cells with a neuronal morphology in microscopy images. A previous study using the same differentiation protocol showed that not all individual cells express neuronal marker genes after differentiation (Young et al. 2014). To further characterize this heterogeneity, we sequenced 177 IPSDSN cells differentiated in three batches from one iPSC line, and clustered them based on expression profiles using SC3 (Kiselev et al. 2017). The data were best explained by two clusters, with 63% of cells forming a tight cluster expressing sensory-neuronal genes (e.g. *SCN9A*, *CHRNB2*), and the remaining 37% of cells forming a looser cluster expressing genes typical of a fibroblastic cell

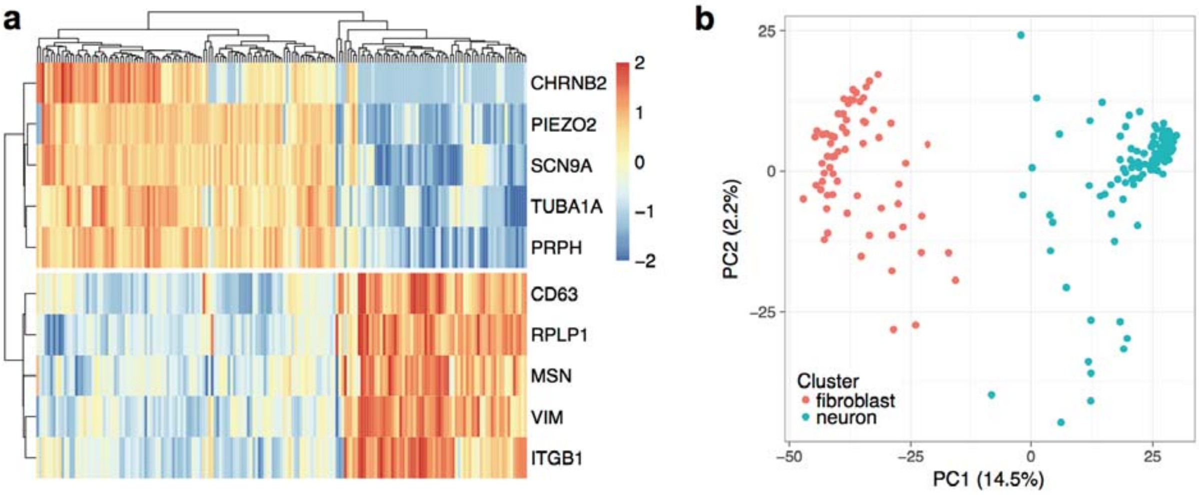
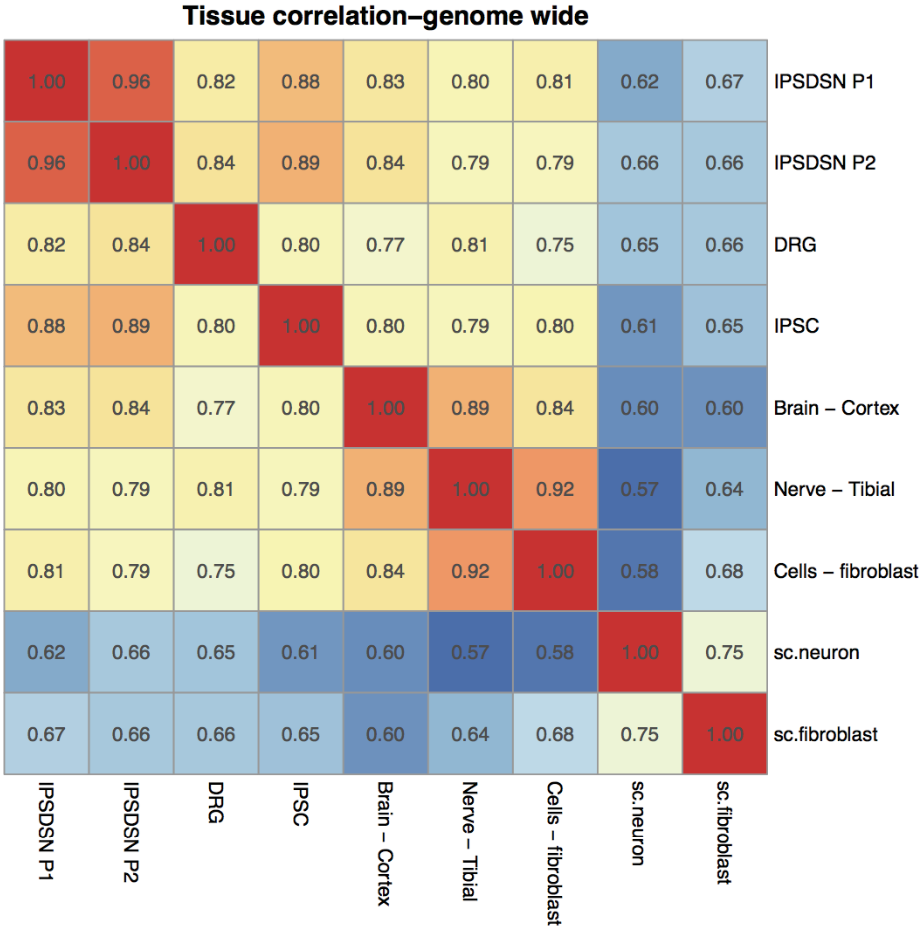


**Figure 3**: Single-cell sequencing of 177 IPSDSN cells. (**a**) Heatmap of RNA-seq data for ten marker genes of the two cell clusters identified by SC3. Color scale denotes normalized, Z-scaled gene expression counts for each gene. (**b**) PCA plot of PC1 vs. PC2 for 177 single cells, based on quantile normalized expression across all genes, with colour indicating SC3 cluster label.

type (e.g. MSN, VIM) (Figure 3a). The two cell types also separated cleanly in a principal components plot (Figure 3b), indicating that the cells do not fall on a smooth gradient from more neuronal to less, but rather have differentiated to distinct cell states. Comparing gene expression from each cluster to other tissues showed that the neuronal cluster was most similar to DRG (Spearman's ρ=0.654), followed by iPSCs (ρ=0.609) and GTEx brain (mean ρ=0.599) (Figure 4) while the fibroblast-like cluster was most similar to GTEx transformed fibroblasts (ρ=0.683), DRG (ρ=0.662), and iPSCs (ρ=0.653). The similarity of these cells to GTEx fibroblasts could suggest a general similarity of adherent cultured cells, although the neuronal cluster had lower similarity to GTEx fibroblasts (ρ=0.579) than many other tissues.



|  | IPSDSN P1 | IPSDSN P2 | DRG | IPSC | Brain – Cortex | Nerve – Tibial | Cells – fibroblast | sc.neuron | sc.fibroblast |
|---|---|---|---|---|---|---|---|---|---|
| IPSDSN P1 | 1.00 | 0.96 | 0.82 | 0.88 | 0.83 | 0.80 | 0.81 | 0.62 | 0.67 |
| IPSDSN P2 | 0.96 | 1.00 | 0.84 | 0.89 | 0.84 | 0.79 | 0.79 | 0.66 | 0.66 |
| DRG | 0.82 | 0.84 | 1.00 | 0.80 | 0.77 | 0.81 | 0.75 | 0.65 | 0.66 |
| IPSC | 0.88 | 0.89 | 0.80 | 1.00 | 0.80 | 0.79 | 0.80 | 0.61 | 0.65 |
| Brain – Cortex | 0.83 | 0.84 | 0.77 | 0.80 | 1.00 | 0.89 | 0.84 | 0.60 | 0.60 |
| Nerve – Tibial | 0.80 | 0.79 | 0.81 | 0.79 | 0.89 | 1.00 | 0.92 | 0.57 | 0.64 |
| Cells – fibroblast | 0.81 | 0.79 | 0.75 | 0.80 | 0.84 | 0.92 | 1.00 | 0.58 | 0.68 |
| sc.neuron | 0.62 | 0.66 | 0.65 | 0.61 | 0.60 | 0.57 | 0.58 | 1.00 | 0.75 |
| sc.fibroblast | 0.67 | 0.66 | 0.66 | 0.65 | 0.60 | 0.64 | 0.68 | 0.75 | 1.00 |

**Figure 4**: Correlation of genome-wide gene expression in different tissues and cell cultures. Gene expression for IPSDSN single cells was averaged within a cluster (sc.neuron, sc.fibroblast). For each gene the mean expression (FPKM) in the group of samples was computed, and these values were correlated genome-wide across groups. Spearman correlation values are shown in each square. P1 and P2 protocol samples are highly similar to each other, and compare similarly to other tissues. Both single cell neurons and single cell fibroblast-like cells have similarity to DRG, although single fibroblast-like cells have greater similarity with GTEx fibroblasts.

Next, we used CIBERSORT (Newman et al. 2015) to estimate the fraction of RNA from neuronal cells in our bulk RNA-seq samples, using the single cell gene expression counts with their cluster labels from SC3 as signatures of neuronal or fibroblast-like expression. The estimated neuronal content was strongly correlated with the first principal component of gene expression ($R^2$ = 0.75, Figure 5), and this corresponded well with a visual assessment of neuronal content from microscopy images (Figure 6).

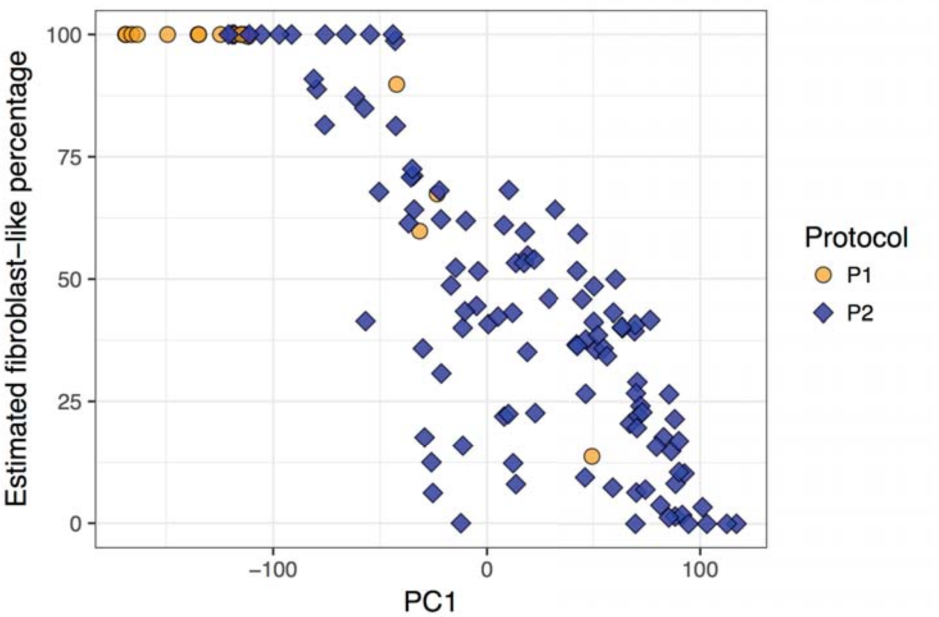


**Figure 5**: Plot of estimated fibroblast percentage for bulk RNA-seq samples versus gene expression principal component 1, after excluding 5 outlier samples. Although many samples have estimated fibroblast percentage close to 100%, these samples contained a significant fraction of neuronal cells in microscopy images.

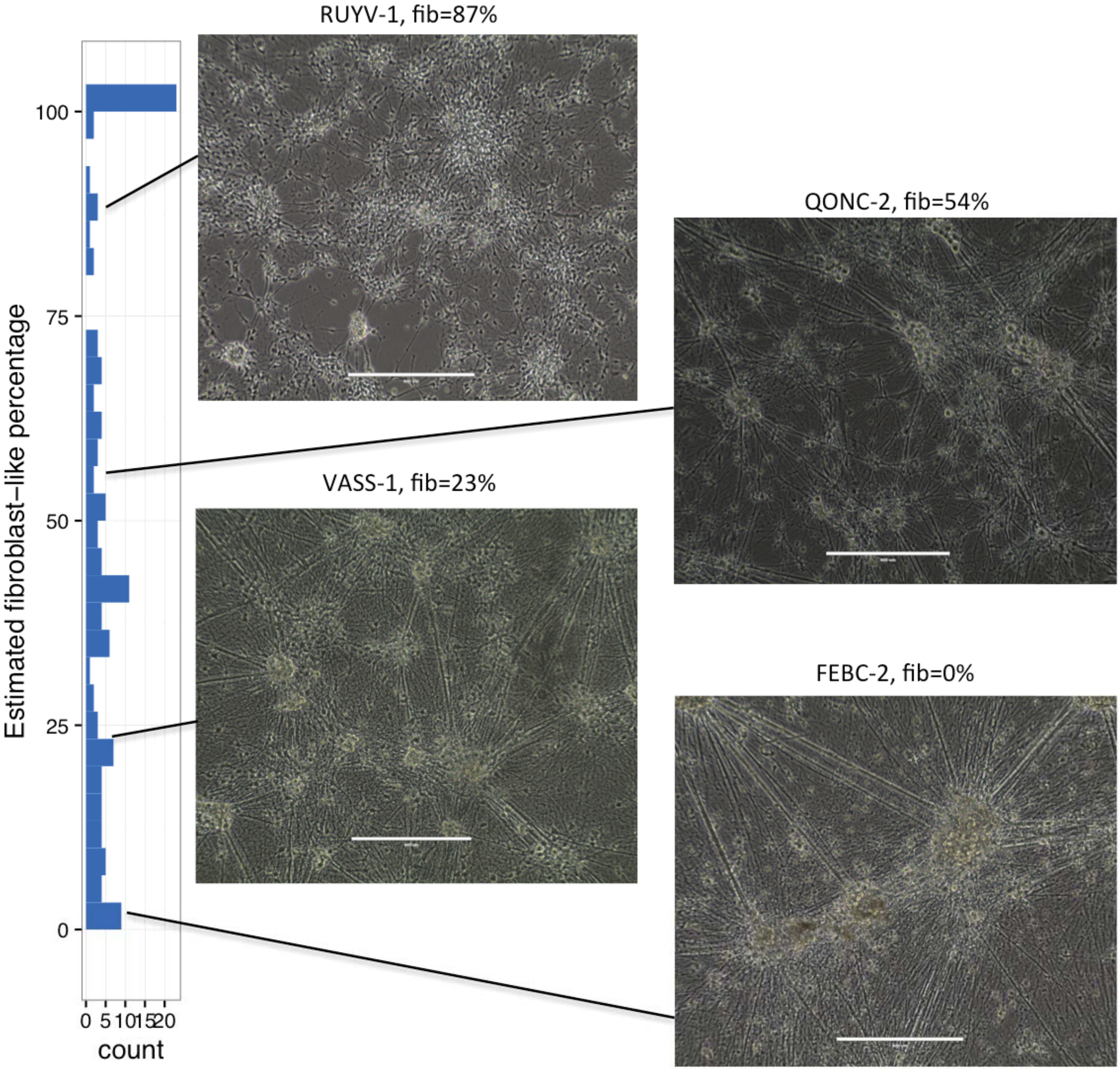


**Figure 6**: Images of four cell lines, taken 4 weeks post induction of differentiation, with their estimated fibroblast-like content shown.

Although a majority of samples appeared by microscopy to have high neuronal content, CIBERSORT estimated relatively high fibroblast-like content for many samples (mean 49%). A factor contributing to this may be the greater RNA content of fibroblast-like cells, which had 2.3-fold more reads in the single-cell RNA-seq data. Indeed when the single cell counts were pooled, CIBERSORT estimated the fibroblast content of this “sample” as 60%, considerably higher than the 37% of single cells in the fibroblast-like cluster. A second consideration is that our scRNA-seq sample was matured for 8 weeks, whereas our bulk RNA-seq samples were matured for 4 weeks. Although gene expression changes are minor after 4 weeks maturation (Young et al. 2014), this difference in maturity means that our single cell reference profiles do not perfectly represent cells in our bulk samples. Despite this, IPSDSN samples estimated to have high fibroblast content still showed greater similarity in genome-wide gene expression with DRG than with any GTEx tissue, including fibroblast cell lines (Figure 7). Although these similarities are reassuring, we note that technical factors could contribute to the greater similarity with DRG, as different gene expression quantification tools were used for GTEx (RNASeQC) and for our iPSC, DRG, and IPSDSN samples (featureCounts).

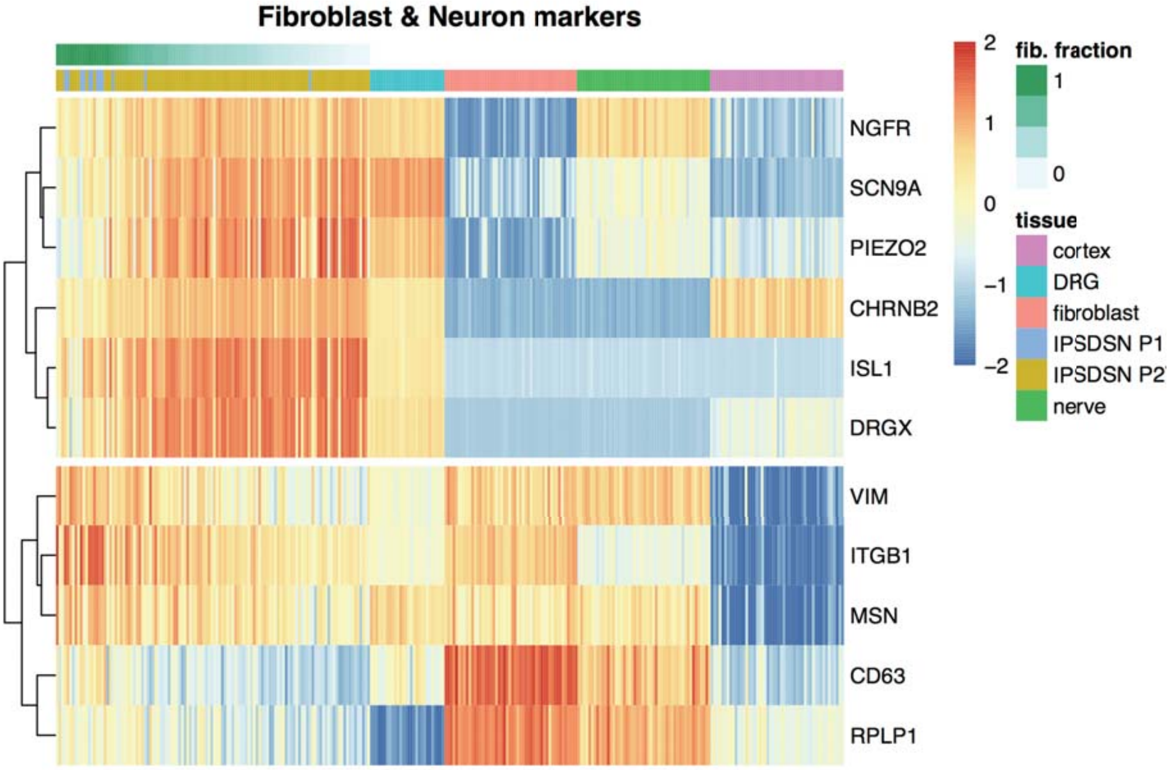


**Figure 7**: Expression of sensory neuronal and fibroblast marker genes across IPSDSN samples, in comparison with DRG (N=28) and selected GTEx tissues (N=50 each). The overall similarity with DRG is high for neuronal markers, but less so for fibroblast markers. Gene expression was determined as log10(FPKM+0.1), and then was mean-centered and Z-score normalized across samples for each gene.

### 2.2.3 Heterogeneity in IPSDSN gene expression

A central issue for genetic studies in iPSC-derived cells is heterogeneity of cellular phenotypes. This heterogeneity could arise from donor genetic background, effects of clonal selection, and effects of the cell culture environment during reprogramming and differentiation. Genome-wide gene expression was highly correlated within lines differentiated multiple times (median Spearman ρ=0.96) and reduced slightly between IPSDSNs from different donors (median ρ=0.93) (Figure 8a). However, differentiation replicates within donor cell lines did not consistently cluster together (Figure 8b), suggesting that variability due to differentiation was at least as large as that due to donor genetic background and iPSC reprogramming together.

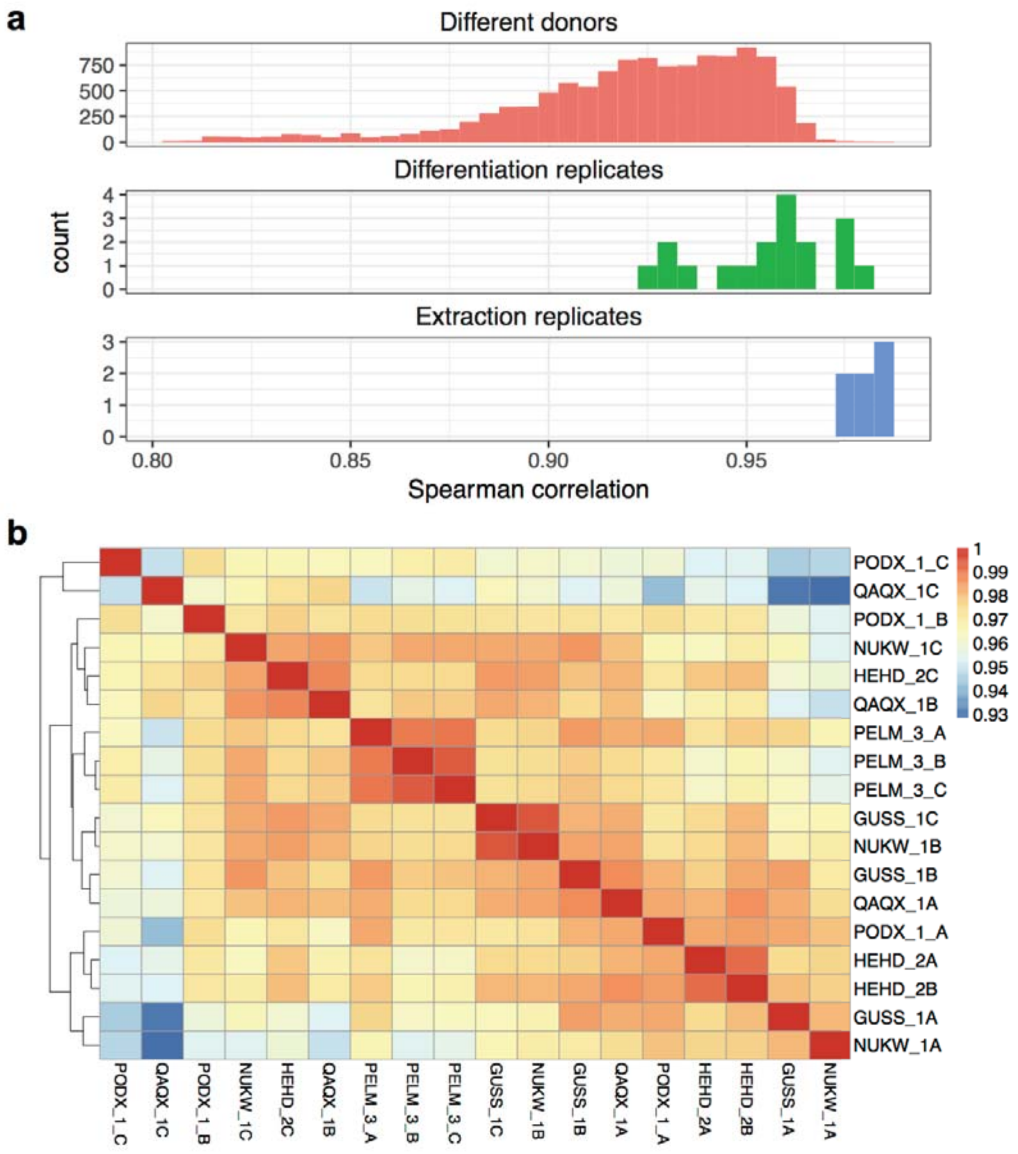


**Figure 8**: Global gene expression correlations across donors and differentiation replicates. (**a**) Histogram of pairwise spearman correlation coefficients of gene expression among RNA extraction replicates (N=7), differentiation replicates from the same donor (n=6 donors, 3 replicates each), or across donors (n=94). RNA extraction replicates were highly repeatable (spearman ρ of 0.97 - 0.98). Differentiation replicates within a donor cell line were less highly correlated (median ρ=0.96, range 0.93 - 0.98), but had higher correlation than differentiations across donors (median ρ=0.93, range 0.80 - 0.98). (**b**) Clustered heatmap of gene expression correlations between samples having differentiation replicates. Replicates for a given donor do not consistently cluster together.

Although marker genes specific to sensory neurons and nociceptors were expressed (FPKM > 1) in nearly all samples, we observed a high degree of heterogeneity in the level of expression of some genes compared with DRG and other tissues (Figure 9a), despite the fact that a cell culture system is theoretically more pure in cell type composition than a complex tissue. These observations were independent of sample size, and were robust when comparing with DRG samples from unique donors only, rather than all 28 DRG samples. Next, we examined how between-sample variability in global gene expression of IPSDSNs compared with other somatic tissues and cell lines. The distribution of coefficient of variation (CV) of gene expression in IPSDSNs fell within the range of most GTEx tissues (Figure 9b). However, the median CV of gene expression in IPSDSNs (0.37) was considerably higher than in DRG (0.23), indicating that IPSDSNs have greater between-sample variability in expression than the primary tissue they are intended to model.

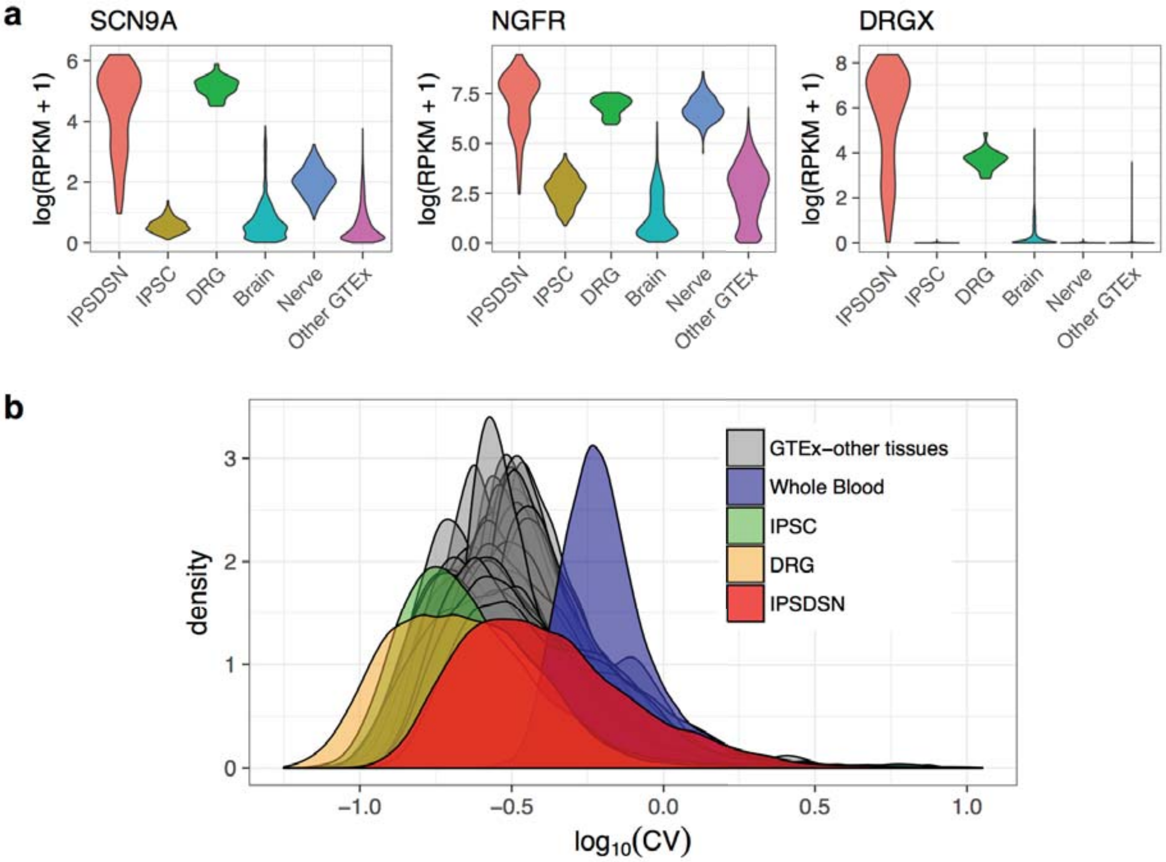


**Figure 9**: Gene expression is highly variable across IPSDSN cell lines. (**a**) Distribution of expression levels for selected sensory neuronal marker genes in IPSDSN, DRG, GTEx tibial nerve, GTEx brain, and all other GTEx tissues. (**b**) Density plot of the coefficient of variation of genes across samples, separately for each GTEx tissue, IPSDSN samples (n=106, P2 protocol only), iPSC (n=200), and DRG (n=28).

Highly variable genes in IPSDSNs were enriched for function in neuronal differentiation and development (Supplementary Table 4). Genes that were significantly upregulated between iPSCs and IPSDSNs, which will include those essential for sensory neuronal function, were also more variable than remaining genes (Figure 10). Importantly, we did not observe similar levels of expression variability of neuronal or developmental gene groups in DRG, iPSCs, or GTEx nervous tissues (Figure 11).

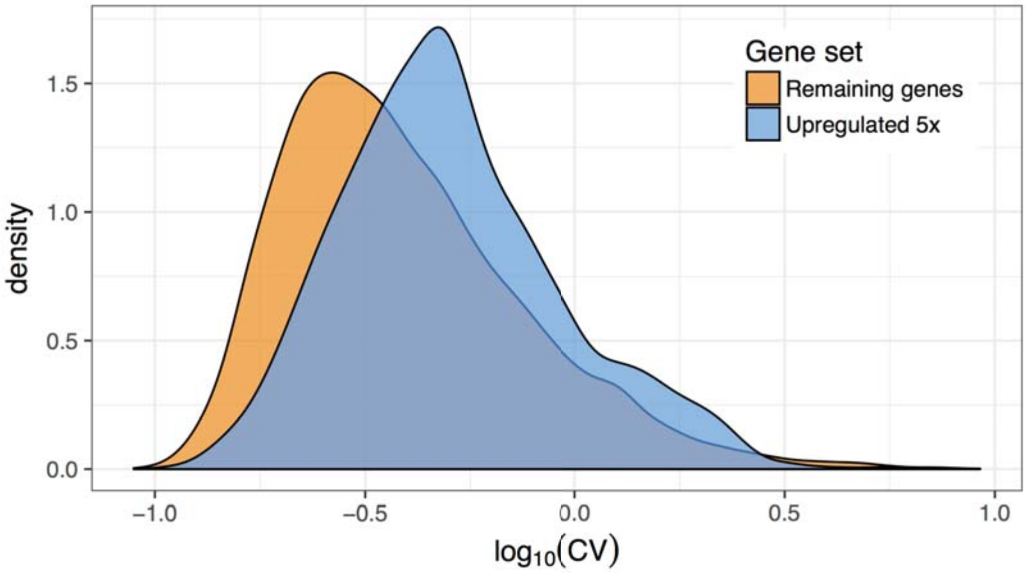


**Figure 10**: Genes upregulated at least 5-fold upon differentiation from iPSC to IPSDSN (FDR < 1%, N=4,246 genes) have increased variability relative to the remaining genes, despite similar levels of expression (median/mean FPKM of upregulated genes: 4.1 / 15.6; remaining genes: 4.6 / 11.8).

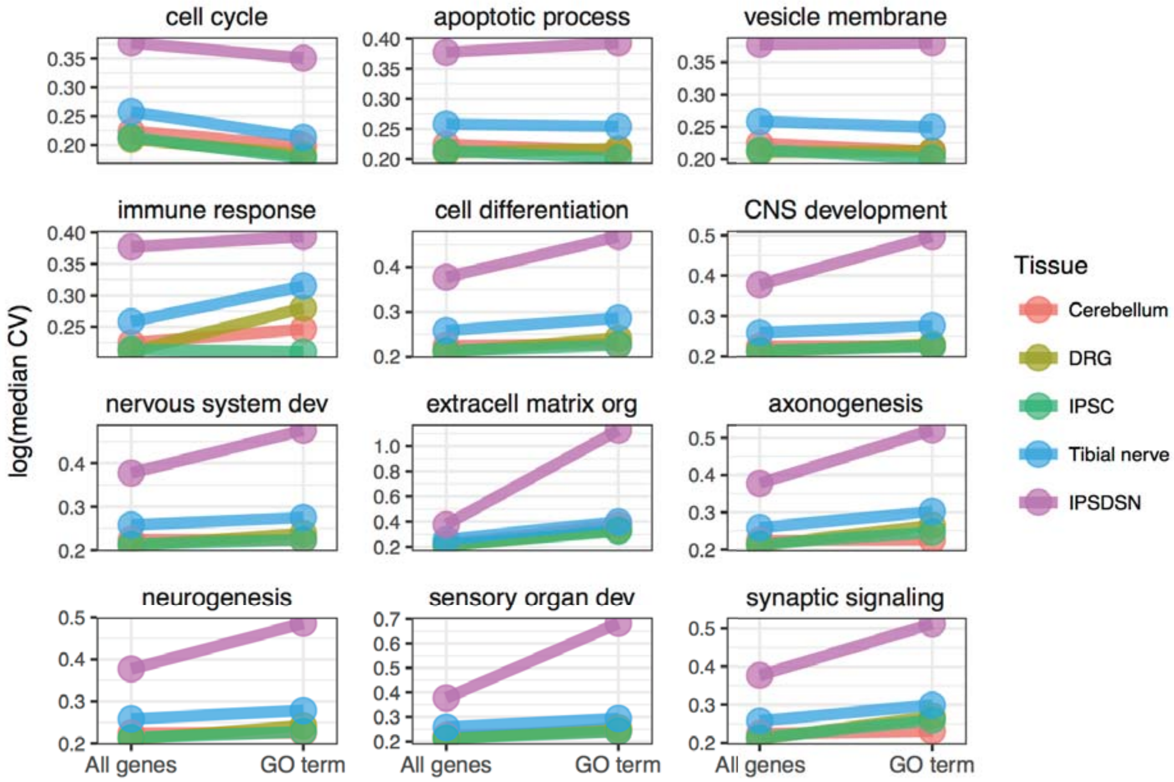


**Figure 11:** Median variability of genes in specific GO terms is compared to median variability for all genes, separately for IPSDSNs, iPSCs, DRG, and GTEx cerebellum and tibial nerve. Genes categories related to neuronal function and differentiation have increased median variability in IPSDSNs relative to all genes, but this is much less the case for iPSCs or nervous tissue samples, and for other gene categories such as immune response or cell cycle genes.

These results highlight that expression of neuronal genes varies substantially more in IPSDSNs than in somatic nervous tissue, probably as a result of variability in differentiation. Consistent with this, variance components analysis (Figure 12) showed that as much or more variation was explained by differentiation batch (median 24.7%) as donor/iPSC line of origin (median 23.3%), which would include both donor and reprogramming effects.

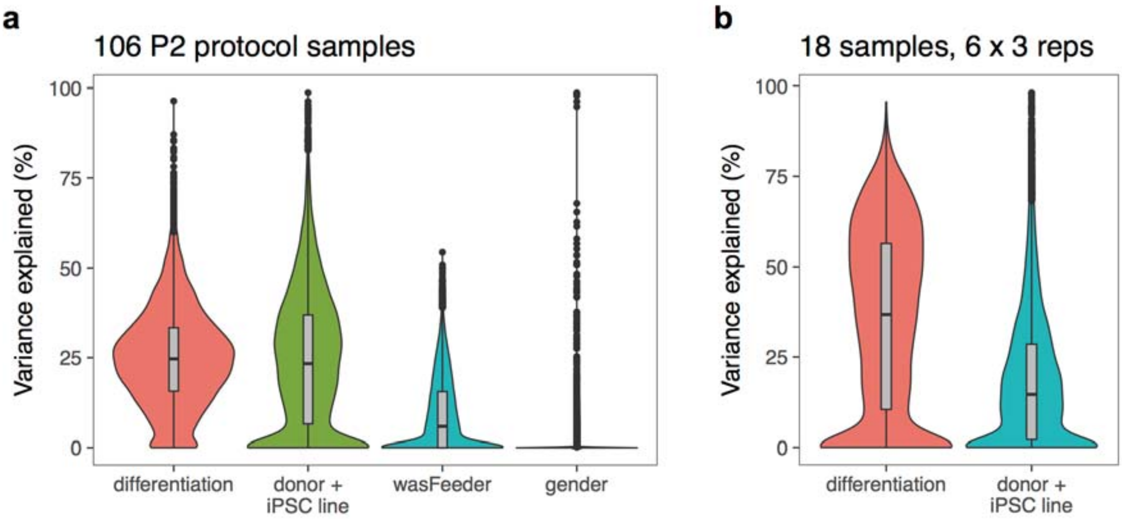


**Figure 12:** Variance components analyses of IPSDSN gene expression. (**a**) Variance partitioning for 119 samples (13 P1 protocol, 106 P2 protocol). (**b**) Variance partitioning for 18 samples, from 6 iPSC lines differentiated 3 times each. All 18 samples were from E8-medium iPSC lines derived from females, and were differentiated with the P2 protocol.

### 2.2.4 iPSC culture conditions influence cell fate

Intriguingly our variance components analysis suggested that, although the cell lines for this analysis were differentiated using an identical protocol, starting iPSC cell culture conditions influenced gene expression patterns in the IPSDSNs produced four weeks later (Figure 13a). Of the 106 successful P2 protocol differentiations, 27 were from iPSCs maintained on mouse embryonic fibroblast (MEF) feeder cells (feeder-iPSCs), while the remaining 79 were grown in Essential 8 medium (E8-iPSCs). The first principal component (PC) of iPSC gene expression clearly differentiated feeder- and E8-iPSCs (Figure 13a), indicating that culture conditions are among the largest global effects on transcription. Similarly, PC1 of gene expression in IPSDSNs distinguished samples originating from feeder- and E8-iPSCs; moreover, IPSDSNs from E8-iPSCs had higher neuronal content (Figure 13b, 28% higher for E8-iPSCs, t-test p=1.84x10$^{-5}$). A possible technical explanation for these results is that protocol implementation and batch effects changed subtly over the course of the project. However, the difference in neuronal content between IPSDSNs derived from E8 or feeder-iPSCs remained when sample derivation date was included as an explanatory covariate (linear regression p=6.5x10$^{-4}$, 36% higher for E8-iPSCs, Figure 13c).

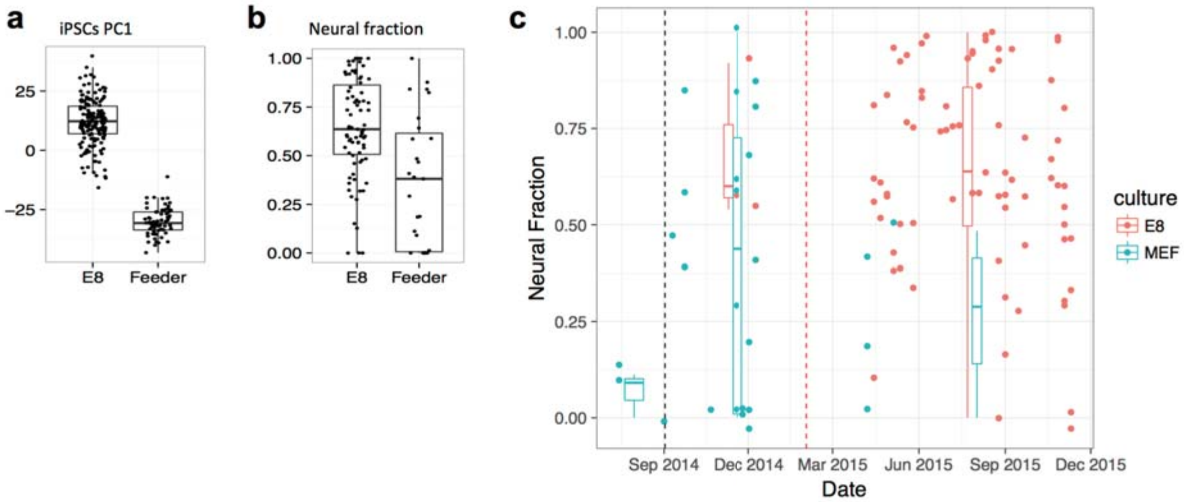


**Figure 13**: Neural fraction and global gene expression in IPSDSNs is influenced by iPSC culture conditions. (**a**) Global gene expression differences between feeder- and E8-iPSCs are captured in PC1. (**b**) Estimated neural fraction of samples differs in IPSDSNs derived from feeder- and E8-iPSCs. (**c**) Neural fraction in IPSDSN samples varies with date and iPSC culture condition. Only P2 protocol samples are shown, and each dot represents the estimated neural fraction for one sample. We used linear models in R to estimate the effects of culture conditions or differentiation date on neural fraction. We either used date as a continuous variable, or split differentiation date into 2 bins (red line) or 3 bins (black and red lines). In all models including both factors, culture conditions were more strongly associated with neural fraction than was date, and the association of culture conditions remained significant. In contrast, the association of date with was not significant with date as a continuous variable or split into 2 bins (p>0.3), and was marginally significant with date split into 3 bins (p=0.038).

Next, we determined genes that were differentially expressed between E8- and feeder-iPSCs and IPSDSNs (Figure 14). Genes more highly expressed in feeder-iPSCs were strongly enriched for mesenchyme development, stem cell differentiation, and Wnt and TGF-β signalling, while genes more highly expressed in E8-iPSCs showed less clear enrichment (Supplementary Tables 5-7). Notably, inhibition of TGF-β/SMAD signalling is a key step in sensory neuronal differentiation. Top differentially expressed genes include early developmental regulators such as *EMX1* (15-fold higher in E8-iPSCs), important for specific neuronal cell fates, and *BMP2* (13-fold higher in feeders), which has been shown to suppress differentiation to sensory cell fates by antagonizing Wnt/beta-catenin (Kléber et al. 2005) (Figure 14b). In addition, *SCN9A* and *TAC1*, key markers of sensory neurons, were expressed at low levels in iPSCs, with 2.2-fold and 2.9-fold higher expression in E8-iPSCs. We also considered genes differentially expressed between IPSDSNs derived from E8- and feeder-iPSCs (Figure 14c). Genes more highly expressed in IPSDSN samples from feeder-iPSCs were overrepresented in extracellular matrix components, pattern specification, organ morphogenesis, and Wnt signalling (Supplementary Tables 8-10), and include *FGFR2*, *BMP7*, and *WNT5A* (Figure 14d). Genes more highly expressed in IPSDSN samples from E8-iPSCs were overrepresented in ion channel complexes, peripheral nervous system development, and synapse organisation, and include *SCN9A*, *DRGX*, and *CACNA1A*. These differences likely reflect the increased neuronal content of samples from E8-iPSCs. Together these results suggest that iPSCs are primed towards different cell fates depending on the iPSC culture medium.

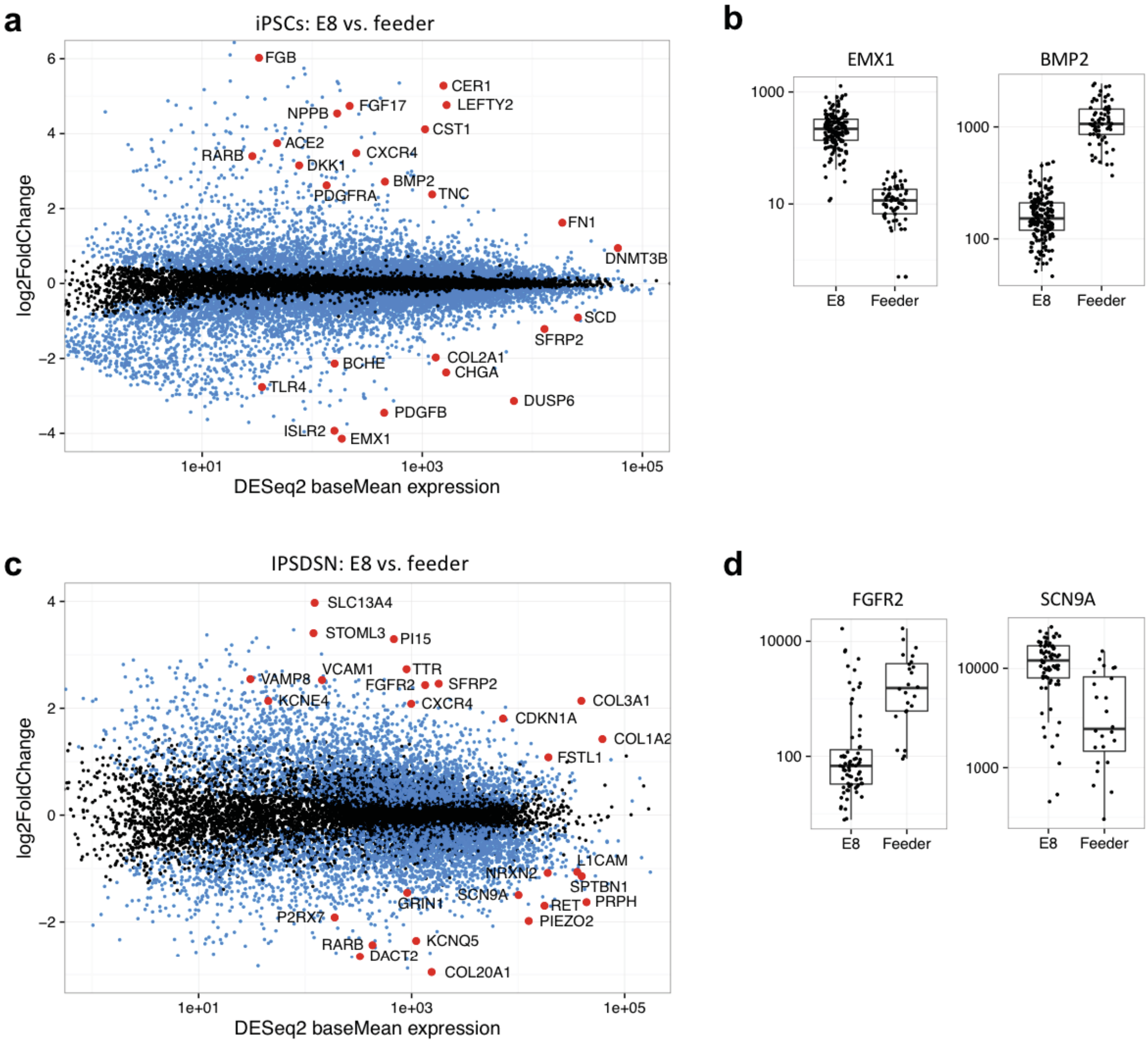


**Figure 14**: (**a**) Differentially expressed genes (FDR 1%, blue and red points) between iPSC samples grown on feeders (n=68) vs. E8 medium (n=171). (**b**) Barplots of selected genes differentially expressed between feeder- and E8-iPSCs. (**c**) Differentially expressed genes (FDR 1%) between IPSDSNs from feeder- (n=27) and E8-iPSCs (n=79). Neuronal differentiation genes, such as *RET* and *L1CAM*, are more highly expressed in samples from E8-iPSCs. (**d**) Barplots of selected genes differentially expressed between IPSDSNs derived from feeder- and E8-iPSCs.

Since iPSC culture conditions influenced differentiation outcomes, we examined gene expression variability within subsets of IPSDSN samples. IPSDSNs differentiated from feeder-iPSCs had somewhat higher global gene expression variability, yet those from E8-iPSCs were still highly variable relative to DRG and iPSCs (Figure 15), with neuronal and developmental gene sets enriched for highly variable genes (Supplementary Table 11). Among the 79 IPSDSNs from E8-iPSCs, samples with high fibroblast content had somewhat higher variability, but those with low fibroblast content still showed high variability relative to DRG and iPSCs.

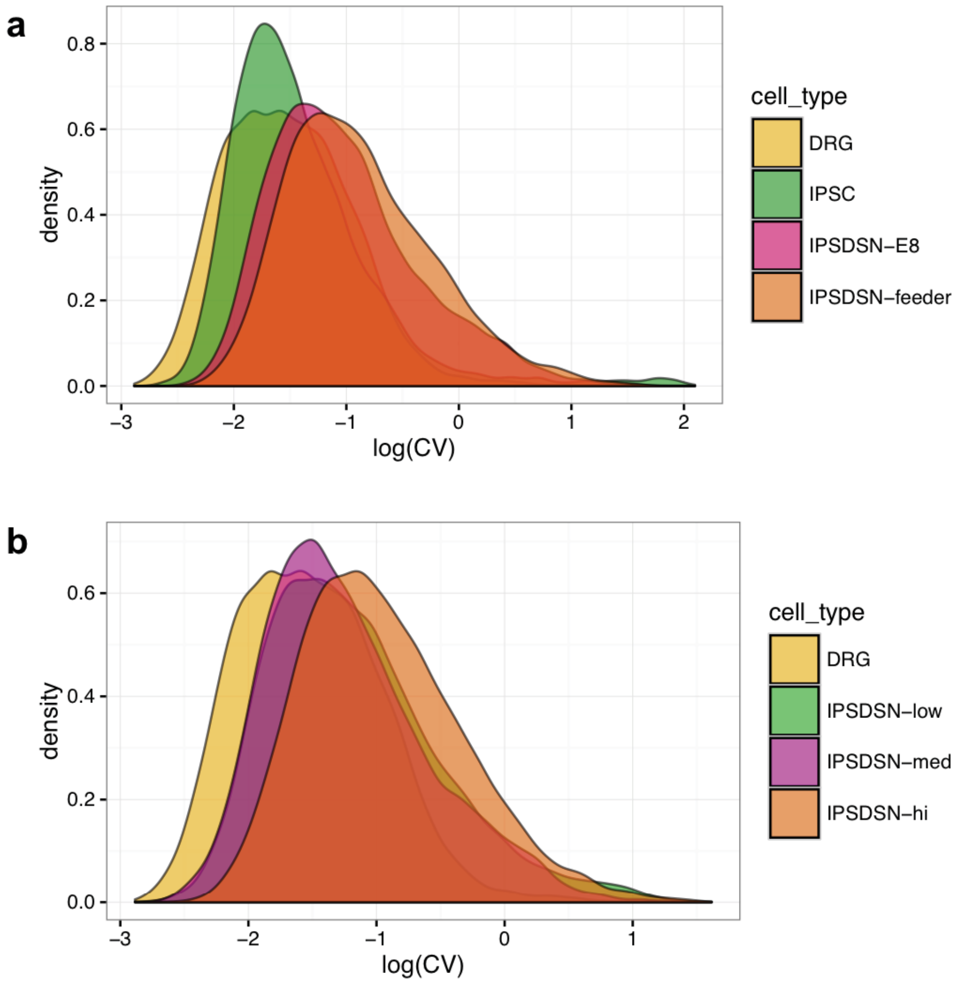


**Figure 15:** The natural log of the coefficient of variation across samples is plotted for different subgroups of samples. **(a)** Samples from feeder-iPSCs (N=27) have slightly higher variability than those from E8-iPSCs (N=79), but samples from E8-iPSCs still have much higher variability than iPSC or DRG. **(b)** Comparing DRG and P2 protocol IPSDSN samples from E8-iPSCs only, separated based on fibroblast-like content: low (estimated < 20%, N=24), medium (20-50%, N=31), and high (> 50%, N=24). High fibroblast-like samples have slightly higher CV across all genes, but this accounts for only a fraction of the increased variability seen relative to primary DRG.

### 2.2.5 Genetic variants influence gene expression, splicing and chromatin accessibility in sensory neurons

Using the linear model FastQTL (Ongen et al. 2016), we mapped 1,403 expression quantitative trait loci (eQTLs)  at FDR 10%, of which 746 were expressed at a moderate level (FPKM > 1). We noted that we discovered many fewer eQTLs than in GTEx tissues of comparable sample size (Figure 16a). This suggested that power for eQTL discovery was lower in IPSDSNs than somatic tissues, possibly due to additional variability introduced by differentiation. Using the allele-specific method RASQUAL (Kumasaka, Knights, and Gaffney 2016) we detected 3,778 genes with expression-modifying genetic variants, termed eGenes, at FDR 10% (Supplementary Table 12), with 2,607 of these expressed at FPKM > 1. Notably, it was only using the additional information from allele specific signals that we achieved approximately similar statistical power to GTEx tissues with equivalent sample sizes.

To identify eQTLs that were not already reported in GTEx (v6), we used a protocol described previously for the HIPSCI project (Kilpinen, Goncalves, et al. 2016). Of all 3,778 eGenes, 954 had tissue-specific associations (Supplementary Table 15), including genes with known involvement in pain or neuropathies, such as *SCN9A*, *GRIN3A*, *P2RX7*, *CACNA1H*/Cav3.2, and *NTRK2*. Because these novel eQTLs were not seen in any GTEx tissue, this suggests that these regulatory variants may have IPSDSN-specific function.

We investigated whether the improvement in power from using allele-specific information was related to gene expression variability. Splitting genes into quartiles of expression variability revealed that power improvement was greatest among genes with high variability across samples (Figure 16b,c). The fraction of novel eQTLs increased for both fastQTL and RASQUAL as gene variability increased (Figure 16d), with RASQUAL overall finding a higher fraction of novel eQTLs. One explanation would be a higher false positive rate for RASQUAL; however, various properties of the novel eQTLs did not differ significantly from known eQTLs, including expression levels, and eQTL variant allele frequency, hardy-weinberg equilibrium, and mapping bias. In addition, the rate of novel eQTLs increased only moderately from the least to the most highly variable genes, and the trend was similar for FastQTL and for RASQUAL. It is possible that relative to GTEx, which used a linear model, RASQUAL finds true eQTLs that are more difficult to discover without examining allele-specific expression.

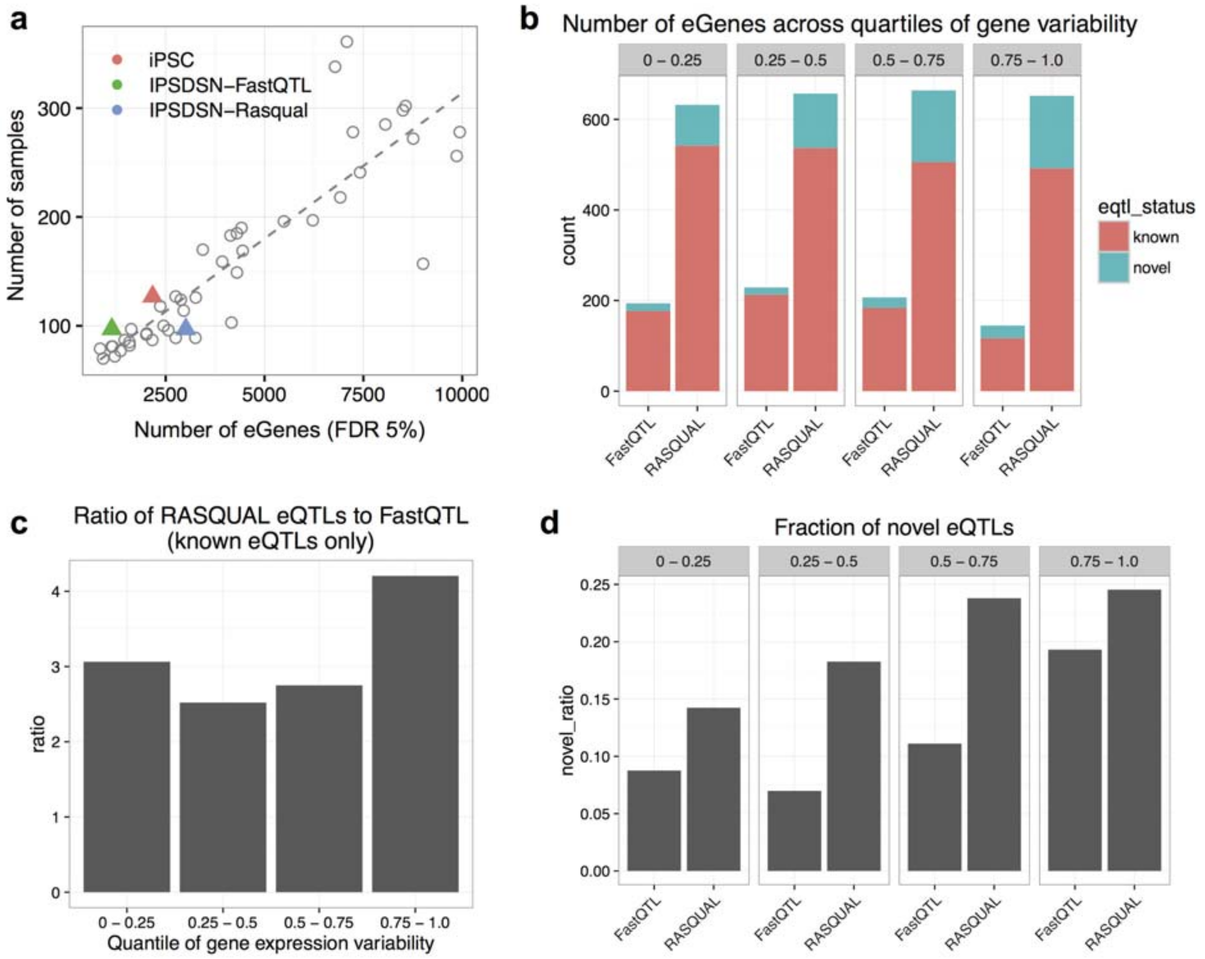


**Figure 16:** (**a**) Number of eGenes discovered for IPSDSNs using either FastQTL (IPSDSN-FastQTL) or RASQUAL (IPSDSN-Rasqual), in relation to sample size, compared with GTEx tissues. (**b**) Number of eGenes discovered by RASQUAL and FastQTL across quartiles of gene expression variability. The bottom 25% of eGenes ranked by sample-to-sample expression variability are at the left, while the top 25% are at the right. Bar colors show the number of eQTLs that overlap with a GTEx eQTL ("known") or where no GTEx tissue has $p < 0.01$ for our lead eQTL SNP ("novel"). (**c**) Restricting to known eQTLs discovered by either method, the ratio of the number of eGenes discovered by RASQUAL to FastQTL is highest in the highest quartile of gene expression variability, although power gains from RASQUAL are high across the board. **(d)** The fraction of novel eQTLs, separately for fastQTL and RASQUAL, across quartiles of gene expression variability.

Variants affecting gene splicing (sQTLs) often change either protein structure or context-dependent gene regulation, and may be more enriched for complex trait loci than are eQTLs (Y. I. Li et al. 2016). To detect sQTLs we used the annotation-free method LeafCutter (Y. I. Li, Knowles, and Pritchard 2016) to define 30,591 clusters of alternatively spliced introns. Using FastQTL (Ongen et al. 2016) we discovered QTLs for 2,079 alternative splicing clusters at FDR 10% (Supplementary Table 13). Notably, only 538 (26%) of the lead variants for these splicing associations were in linkage disequilibrium (LD) $r^2$ >= 0.5 with a lead eQTL variant in our dataset, indicating that the sQTLs extend our catalog of expression-altering variants and are not merely proxies for gene-level eQTLs (or vice versa).

| | Number | GWAS overlap |
|---|---|---|
| eQTLs | 3778 | 156 |
| sQTLs | 2079 | 129 |
| ATAC QTLs | 6318 | 172 |
| Joint ATAC/eQTLs | 177 | 14 |

**Table 1**: QTL associations. Columns show the number of associations, and the number of unique overlaps ($r^2 > 0.8$) between lead QTL SNPs and GWAS catalog SNPs after removing duplicates for each GWAS trait.

We collected ATAC-seq data for 31 samples (Buenrostro et al. 2013) and used this to identify active regulatory regions in IPSDSNs and to map 6,318 caQTLs chromatin accessibility QTLs (caQTLs) at FDR 10% (Supplementary Table 14). To identify transcription factors in IPSDSNs whose binding is altered by regulatory variants, we used the LOLA Bioconductor package (Sheffield and Christoph 2015) to test for enrichment of our lead QTL SNPs, relative to GTEx lead SNPs, in ENCODE ChIP-seq peaks and JASPAR transcription factor motifs (Supplementary Tables 16,17). Tissue-specific eQTLs were highly enriched within SMARCB1 and SMARCC2 peaks (odds ratios 5.8 and 14.1; $p < 5x10^{-5}$), which are both members of the neuron-specific chromatin remodeling (nBAF) complex (Lessard et al. 2007). Also enriched were REST/NRSF (OR=5.7, $p=1.1x10^{-4}$) and SIN3A (OR=3.9, $p=1.0x10^{-4}$), which bind neuron-restrictive silencer elements during development, but have suggested roles in the development and maintenance of neuropathic pain (Willis et al. 2016). Considering all IPSDSN eQTLs, we found enrichments for ELK1 and ELK4, as well as c-Fos, a target of ELK1 and ELK4 which is widely expressed but is known to have specific functions in sensory neurons (Hunt, Pini, and Evan 1987; Kohno et al. 2003). Notably, DNA sequence motifs for REST, ELK1 and ELK4 are also among the most highly enriched motifs in our ATAC-seq peaks (Supplementary Table 18).

### 2.2.6 Sensory neuron eQTLs and sQTLs overlap with complex trait loci

While we were interested in comparing our set of QTLs with GWAS for pain, the largest GWAS for pain to date included just 1,308 samples and found no associations at genome-wide significance (Peters et al. 2013). We therefore considered all GWAS catalog associations with $p < 5x10^{-8}$ that were in high LD ($r^2 > 0.8$) with a QTL in our dataset, with two purposes in mind: to determine whether any GWAS traits are enriched overall for overlap with sensory neuron QTLs, and to find individual cases where a QTL is a strong candidate as a causal association for the GWAS trait. Overall, IPSDSN eQTLs were significantly enriched for overlap with GWAS catalog SNPs ($p < 0.001$) relative to 1000 random sets of SNPs matched for minor allele frequency (MAF), distance to nearest gene, gene density, and LD (Pers, Timshel, and Hirschhorn 2015), and the overlap was consistent with that seen for eQTL studies in other tissues (Figure 17). Although nociceptive neurons are specialized for sensing and relaying pain signals, they share characteristics with other neurons; thus, we might expect enrichment for traits known to involve the nervous system more generally. However, among the 41 traits with at least 40 GWAS catalog associations, we could not detect any trait with significantly greater overlap with our QTL catalog than other traits after correcting for multiple testing (Supplementary Table 19).

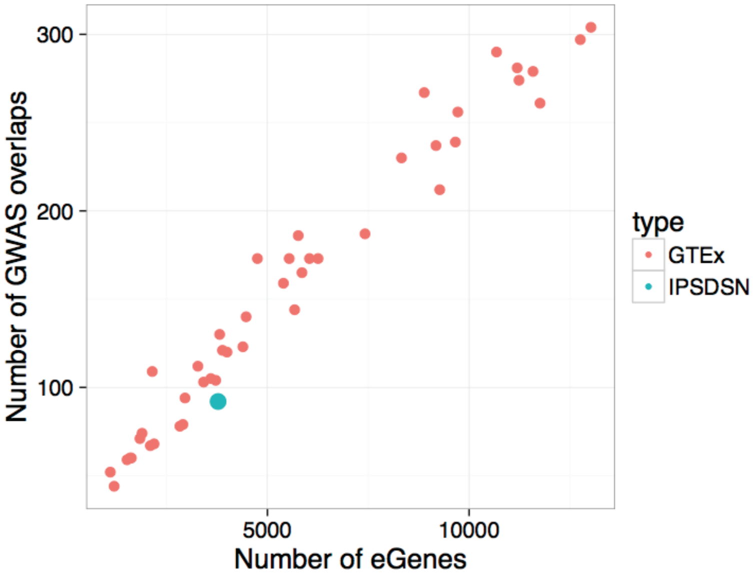


**Figure 17:** The number of overlaps at LD $R^2 > 0.8$ between eQTLs in IPSDSNs and GWAS catalog SNPs is within the range seen for similarly powered tissues in GTEx. The number of eQTL-GWAS overlaps is heavily dependent on the number of eQTLs discovered, which is reflected in the tight linear relationship with the number of eGenes.

Across all traits, we found 156 genes with an eQTL overlapping at least one GWAS association, and similarly 129 sQTLs and 172 caQTLs with GWAS overlap (full catalog in Supplementary Tables 20-22). We examined individual associations, in conjunction with ATAC-seq peaks and LD information, to identify candidate causal variants influencing both a molecular phenotype and a complex trait. For most of these associations we do not expect that sensory neurons are the most relevant cell type; rather the overlaps may reflect either general neuronal mechanisms or non-cell-type-specific functions. We thus focused on traits where neurons are likely to be a relevant cell type.

Among overlapping associations we found a number that relate to neuronal diseases, such as Parkinson’s disease, multiple sclerosis, and Alzheimer’s disease. One striking overlap is between an eQTL for *SNCA*, encoding alpha synuclein, and Parkinson’s disease, for which a likely causal variant has recently been identified (Soldner et al. 2016). The lead GWAS SNP and our lead eQTL are both in perfect LD with rs356168 (1000 genomes MAF 0.39), which lies in an ATAC-seq peak in an intron of *SNCA*. Soldner et al. used CRISPR/Cas9 genome editing in iPSC-derived neurons to show that rs356168 alters both *SNCA* expression and binding of brain-specific transcription factors (Soldner et al. 2016). In IPSDSN cells we find that the G allele of rs356168 increases *SNCA* expression 1.14-fold, in line with Soldner et al. who reported 1.06- to 1.18-fold increases in neurons and neural precursors. However, despite residing in a visible ATAC-seq peak in our data, rs356168 is not detected as a caQTL (SNP p value = 0.22). eQTLs for SNCA have recently been reported in the latest GTEx release (v6p), but none of the tissue lead SNPs are in LD ($r^2 > 0.2$) with rs356168, suggesting that the effect of this SNP can be more readily detected in specific cell and tissue types, including IPSDSNs and the frontal cortex tissue and iPSC derived neurons studied by Soldner et al.

We also find multiple compelling overlaps between splice QTLs and GWAS associations (Figure 18). One known example is a strong sQTL for *TNFRSF1A* ($p=9.9x10^{-29}$) with the same lead SNP (rs1800693, MAF 0.30) as a multiple sclerosis association. This likely causal SNP is located 10 base pairs from the donor splice site downstream of exon 6, and has been experimentally shown to cause skipping of exon 6, which results in a truncated, soluble form of TNFR1 that appears to reduce TNF signalling (Gregory et al. 2012). TNFRSF1A is highly expressed (>15 FPKM) in both IPSDSNs and in DRG. We do not see an effect of this variant

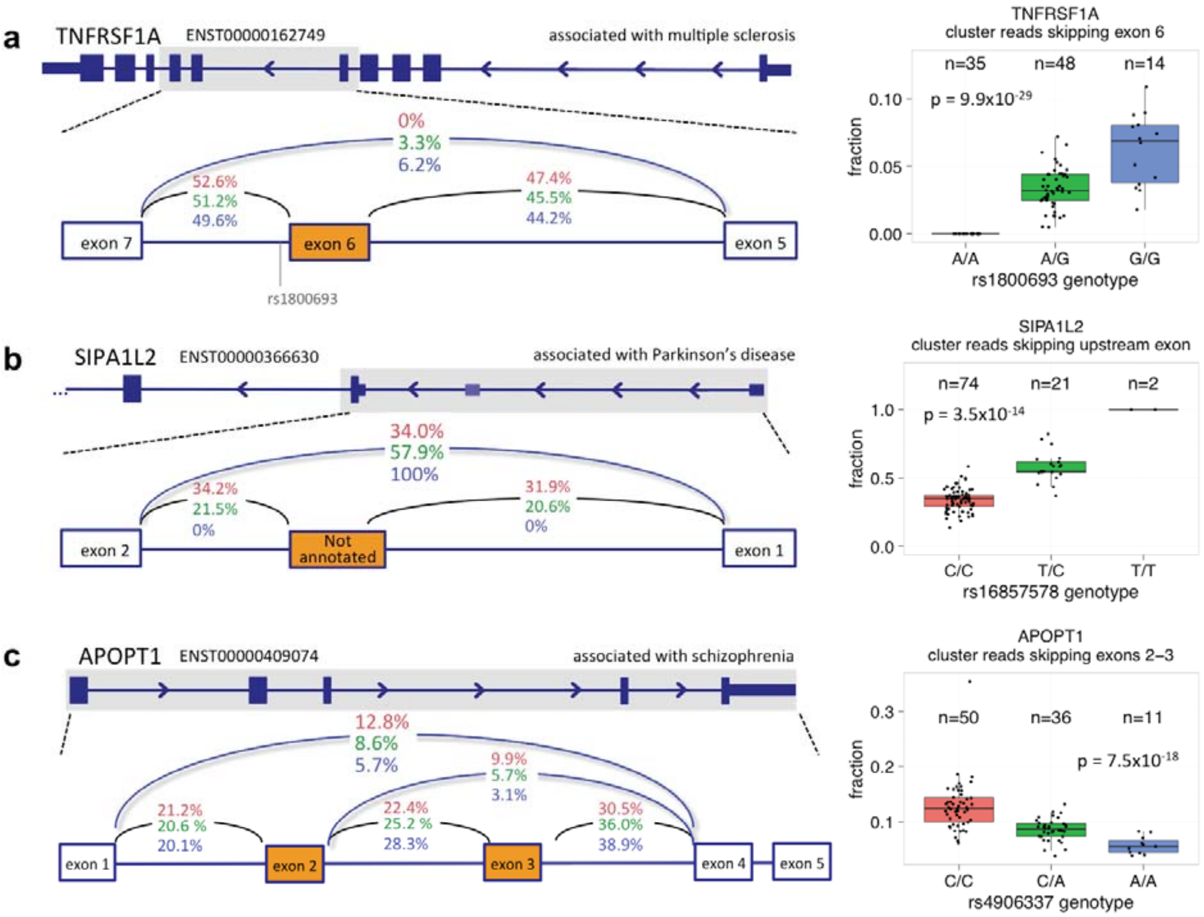


**Figure 18**: Splicing QTLs overlapping GWAS. **(a)** An sQTL for *TNFRSF1A* leads to skipping of exon 6, and overlaps with a multiple sclerosis association. **(b)** An sQTL for *SIPA1L2* leads to increased skipping of an unannotated exon between alternative promoters, and overlaps with a Parkinson's disease association. **(c)** An sQTL for *APOPT1* alters skipping of exons 2 and 3, and overlaps with a schizophrenia association. P values are from the beta approximation based on 10,000 permutations as reported by FastQTL.

on total expression levels in our cells (p > 0.5), but we observe skipping of exon 6 in about 12% of transcripts from individuals homozygous for rs1800693 (Figure 18a). Since these transcripts undergo nonsense-mediated decay (Gregory et al. 2012), the actual rate of exon skipping is likely to be higher. Given the broad role of TNF in inflammation and immunity, it is interesting that rs1800693 is associated with MS but not with other autoimmune disorders, apart from primary biliary cirrhosis (Gregory et al. 2012). Moreover, whereas TNF inhibitors are effective in many autoimmune disorders, they exacerbate MS, an effect that is mimicked by the reduction in TNF signalling produced by the *TNFRSF1A* splice variant. These observations suggest an interplay between cells of the CNS and immune system involving TNF signalling. TNF signalling has been shown to have both inflammatory and neuroprotective effects in the CNS and, despite a large body of research, the exact mechanisms and cell types responsible for the genetic risk associated with TNF receptor polymorphisms remain unclear (Probert 2015).

An sQTL for *SIPA1L2* (rs16857578, MAF 0.23) is in LD with associations for both Parkinson's disease (rs10797576, $r^2$=0.93) and blood pressure (rs11589828, $r^2$=0.94). An unannotated noncoding exon (chr1:232533490-232533583) between alternative *SIPA1L2* promoters is included in nearly 50% of transcripts in individuals with the reference genotype, but splicing in of the exon is abolished by the variant (Figure 18b). SIPA1L2, also known as SPAR2, is a Rap GTPase-activating protein expressed in the brain and enriched at synaptic spines (Spilker, Christina, and Kreutz 2010). Although its function is not yet clear, expression is seen in many tissues profiled by GTEx, with highest expression in the peripheral tibial nerve. Interestingly, the related protein SIPA1L1 exhibits an alternative protein isoform with an N-terminal extension that is regulated post-translationally to influence neurite outgrowth (Jordan et al. 2005).

A complex sQTL for *APOPT1* (rs4906337, MAF 0.22) is in near-perfect LD with a schizophrenia association (rs12887734). The splicing events involve skipping either of exon 3 only or both exons 2 and 3 (Figure 18c). At least 20 variants are in high LD ($r^2$ > 0.9), including rs4906337 which is 40 bp from the exon 3 acceptor splice site, and rs2403197 which is 63 bp from the exon 4 donor splice site. No sQTL is reported in GTEx, and although eQTLs are reported for *APOPT1*, only the thyroid-specific eQTL (rs35496194) is in LD ($r^2$ = 0.94) with the schizophrenia-associated SNP rs12887734. APOPT1 is localized to mitochondria and is broadly expressed. Homozygous loss-of-function mutations in this gene lead to Cytochrome c oxidase deficiency and a distinctive brain MRI pattern showing cavitating leukodystrophy in the posterior region of the cerebral hemispheres, with affected individuals having variable motor and cognitive impairments and peripheral neuropathy (Melchionda et al. 2014).

### 2.2.7 Recall by genotype studies in iPSC-derived cells will require large sample sizes

One attractive future use of iPSCs is to experimentally characterise GWAS loci using a "recall by genotype" approach. Here, iPSC lines with specific genotypes are chosen from a large bank and differentiated into target cell types (for example, see (Warren et al. 2017)). Our observations suggested that, for certain protocols, the additional cellular heterogeneity introduced by differentiation could impact the power of these studies to detect the effects of common genetic variants. Importantly, our large set of differentiations gave us accurate genome-wide estimates of effect size and expression variability in an iPSC-derived cell type, for use as a benchmark "ground truth". We investigated the performance of iPSC-based recall by genotype studies by bootstrap resampling from a stringent (FDR 1%) IPSDSN eQTL call set. For each eQTL gene we sampled expression counts from an equal number of major and minor homozygotes for the lead SNP, sampling with replacement to achieve a specific sample size. We then estimated power as the fraction of 100 bootstrap replicates where we found a significant difference ($p < 0.05$, Wilcoxon rank sum test) in expression between the homozygotes.

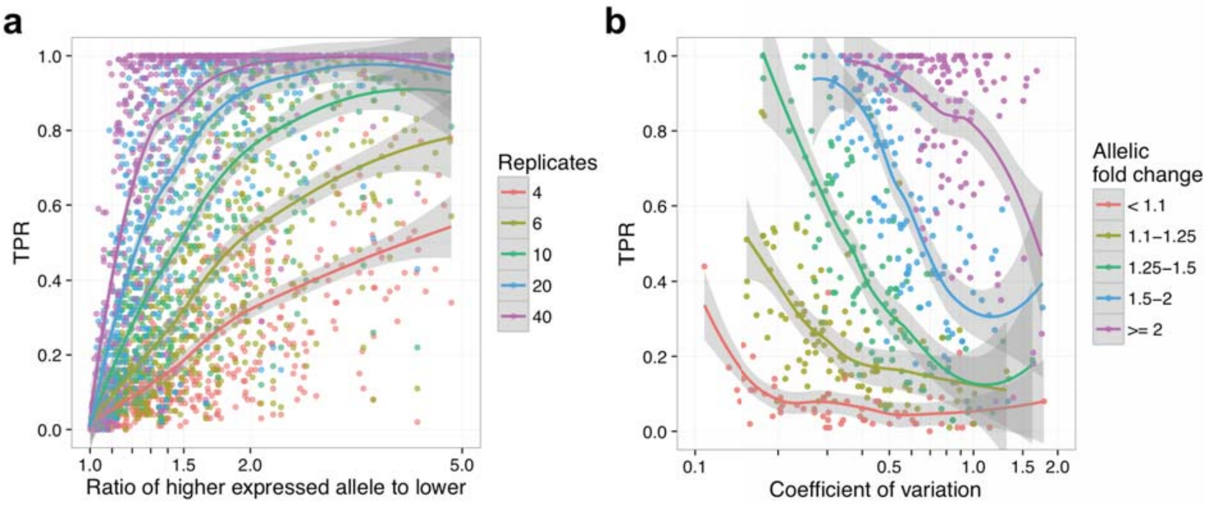


**Figure 19**: Power to detect a genetic effect in a single-variant single-gene test depends on sample size, allelic effect size, and gene expression variability. (**a**) TPR as a function of allelic fold change for five different numbers of replicates (half the total sample size). (**b**) TPR as a function of CV for five bins of allelic fold change, with 10 samples of each genotype.

Our results illustrate important trends. First, recall by genotype studies in iPSC-derived cells are likely to require relatively large sample sizes, typically 20-80 unrelated individuals, for variants with a 1.5-2-fold effect size (Figure 19a). Second, as expected, highly variable genes are more challenging (Figure 19b) with power below 40% in a sample size of 20 for even moderately variable genes (CV 0.5 - 0.75). While expression noise will not typically be known accurately a priori, an estimate of effect size may be available from previous eQTL studies in specific tissues. This could enable estimating the number of samples needed to achieve a desired power.

Note that these power estimates assume that a single gene is being tested, which is only likely to be the case when there is a very strong prior belief in the causal gene and few genes in the region. Where multiple genes are tested, power will be lower. These results also suggest that large sample sizes will be required when using genome editing to identify causal GWAS-associated variants: although genetic background can be controlled in such an experiment, differentiation noise will continue to be a major contributor to gene expression variability.

## 2.3 Discussion

iPSC-derived cells enable the molecular mechanisms of disease to be studied in relevant human cell types, including those which are inaccessible as primary tissue samples. Because the effect sizes of common disease-associated risk alleles tend to be small, observing their effects in cellular models is challenging (Raghavan et al. 2016; Soldner et al. 2016). In an iPSC-based system, this difficulty is compounded by variability between samples in the success of differentiation, as described for hepatocytes (Dianat et al. 2013), hematopoietic progenitors (Smith et al. 2013), and neurons (Hu et al. 2010; Handel et al. 2016).

Our study is the first that we are aware of to perform iPSC differentiation to a neuronal cell type  and functionally characterise the resulting cells at scale. Sample-to-sample variability in gene expression in the iPSC-derived cells was greater than in DRGs, with highly variable genes enriched in processes relating to neuronal differentiation and development. This highlights that genes likely to be of particular interest and relevance for the function of these cells are also among the most variable, a challenge which may be broadly true of iPSC-derived cells. Despite the observed variability, we detected thousands of eQTLs, sQTLs, and caQTLs in IPSDSNs, most of which were discovered only with a model that statistically combines both allele-specific and between individual differences in expression to improve power for association mapping. Some of these overlap known expression-modifying variants that are associated with disease, such as an eQTL for *SNCA* associated with Parkinson's disease. However, for most of these disease overlaps the causal variants are not known. This QTL map is thus a starting point for in-depth dissection of individual loci in iPSC-derived neurons where we have shown that a genetic effect is present.

Although our study highlights the potential power of IPSC derived cells as model systems for studying human genetic variation, our results also illustrate the limitations of this approach. First, despite expressing key marker genes and exhibiting neuronal morphology and electrophysiology, it is clear from our data that IPSDSNs are transcriptionally distinct from the other cell types we examined, including DRGs. This reflects a limitation of existing *in vitro* differentiation protocols, which produce cells that are not as functionally or transcriptionally mature as primary tissues. Second, our differentiations did not produce pure populations of neurons, nor could we measure the purity of the resulting cultures precisely. A portion of the sample-to-sample variability that we observed is likely due to this mixture of cell types, which varied across differentiations. Although mature neurons can be labeled for marker genes, they are not easily sorted by automated systems, which limits the high-throughput options available for purifying neuronal populations. As a result, the eQTLs that we discovered do not represent those of a pure sensory neuronal cell type. For many cell types, sorting is more feasible, and could provide one solution to the variable maturity and heterogeneity of differentiated cell populations.

We used single-cell RNA-seq from three differentiation batches to characterise IPSDSN heterogeneity, which showed that they cluster into neuronal cells and cells with more fibroblast-like gene expression. Using reference profiles from these clusters enabled us to estimate a proxy measure of neuronal cell purity in our bulk RNA-seq samples, and these estimates qualitatively agreed with the neuronal content in images from the cell cultures. Our method is similar to a deconvolution approach described recently using bulk and single-cell sequencing of primary human and mouse pancreas (Baron et al. 2016).

The similarity of the fibroblast-like single cells to DRG raises the important question of whether these cells are immature sensory neurons. Single-cell sequencing at multiple time points during MYOD-mediated myogenic reprogramming has suggested that some individual cells traverse a desired course, while others terminate at incomplete or aberrant reprogramming outcomes (Cacchiarelli et al. 2017). Such an approach in IPSDSNs could reveal determinants of neuronal differentiation trajectories, and may yield useful insights for protocol changes to improve the purity of differentiated neurons, or to specify more precise neuronal subtypes. More generally, replacing bulk RNA-seq with single cell sequencing across many samples could enable *in silico* sorting of cells based on their transcriptome, and better characterisation of the sources of variation within a differentiated population of cells. Further, culturing cells from multiple donors in a pool, along with an scRNA-seq readout, could reduce differentiation-related batch effects while retaining the ability to identify donor-specific genetic effects on gene expression. These advantages suggest to us that a move towards scRNA-seq will be extremely useful in iPSC-derived cell models.

For iPSC models of common disease associated variants to be used effectively, it is critical to know which candidate disease associated variants exhibit a detectable cellular phenotype in an *in vitro* model. We used *in silico* resampling to estimate the sample sizes needed to detect the effects of noncoding regulatory variants in iPSC-derived cells using a recall by genotype design. Power above 80% is only achieved with surprisingly large (40+) samples, even for alleles with a fold change of 1.5 to 2. Further, the power we report may be overestimated, due to ascertainment bias in defining a set of eQTLs as “true positives”, which fails to include true genetic effects that we did not discover in our samples. Even larger samples will be needed when multiple genes, for example in a single GWAS interval, are to be tested. These observations are consistent with a recent genome-editing experiment that required 136 differentiations in hepatocyte-like cells to discover an effect of rs12740374 on *SORT1* gene expression (Warren et al. 2017). Notably, the modest effect of this variant on expression in hepatocyte-like cells (1.3-fold increase) stands in contrast to the large effect of the variant (4- to 12-fold increase) observed previously in primary liver (Musunuru et al. 2010). Where it is possible to use a coding SNP to assess the allele-specific effect of a genome edit, as done for *SNCA* (Soldner et al. 2016), this may prove a more efficient approach to detecting causal effects of individual regulatory variants.

In summary, we have measured multiple molecular phenotypes in a large panel of iPSC-derived neurons. The catalog of QTLs we provide reveals a large set of common variants and target genes with detectable effects in IPSDSNs. These associations provide promising targets for functional studies to fine-map causal disease-associated alleles, such as by allelic replacement using CRISPR-Cas9, and our study describes the importance of considering differentiation-induced variability when planning these studies in iPSC-derived cells.

### Data Availability

Code used for processing and analysing data is available at https://github.com/js29/ipsdsn. RNA-seq and ATAC-seq for open access samples are deposited in the European Nucleotide Archive under accession ERP020576. These data for managed access samples are deposited in the European Genome Archive under accession EGAD00001003145. Summary statistics and gene expression counts are available at https://www.ebi.ac.uk/biostudies/studies/S-BSST16. Sample genotypes and accession numbers are available at http://www.hipsci.org/data.

## 2.4 Methods

### URLs

CIBERSORT, cibersort.stanford.edu.
ENCODE, www.encodeproject.org.
GTEx, www.gtexportal.org.
HIPSCI, www.hipsci.org.

## IPS cell lines

A summary of iPSC lines used is available in Supplementary Table 2, and details of processes and assays for these iPSCs generated by the HIPSCI project are available at www.hipsci.org. Briefly, 107 human iPSCs from 103 healthy donors were obtained from HIPSCI[8]. Of these, 38 were initially grown in feeder-dependent medium and the remainder were grown in feeder-free E8 medium. All HIPSCI samples were collected from consented research volunteers recruited from the NIHR Cambridge BioResource, initially under existing ethics for iPSC derivation (REC Ref: 09/H0304/77, V2 04/01/2013), with later samples collected under a revised consent (REC Ref: 09/H0304/77, V3 15/03/2013).

## Sensory neuron differentiation

All differentiations in this study were performed by a single individual, and a summary of the IPSDSN cell lines is in Supplementary Table 1. Two protocols were used, named P1 (13 differentiations) and P2 (110 differentiations). P1 protocol samples were included for QTL calling, and other analyses used P2 protocol samples exclusively. The P1 protocol (described in[7]) involved the addition of “2i” inhibitors (LDN193189 and SB-431542) for 5 days, followed by “5i” inhibitors (LDN193189, SB-431542, CHIR99021, DAPT, SU5402) for 6 days. When applying this protocol to a larger number of samples we observed excessive cell death prior to obtaining neural progenitors (days 9-12). We altered the protocol to make it more similar to (Chambers et al. 2012), by:

- using E8 rather than mTeSR1 media when maintaining iPSCs prior to differentiation;
- phasing in neurobasal media beginning at day 4, and gradually increasing this to 100% by day 11;
- beginning addition of inhibitors 5i at day 3 rather than day 5;
- stopping addition of small molecule inhibitors LDN193189 and SB-431542 beginning at day 7 rather than day 11, referred to as “3i” for the 3 inhibitors that continued to be added.

Functional assays ($Ca^{2+}$ flux, response to Veratridine) confirmed that response of the sensory neurons produced by each protocol was equivalent; however, the P2 protocol performed more consistently across cell lines and culture parameters.

### P2 protocol details

All reagents were from Life Technologies unless otherwise indicated. Clump-passaged iPSCs were single-cell seeded in E8 media on growth factor-reduced Matrigel (BD Biosciences) 48 hours prior to neural induction (day 0). KSR Media was prepared as 500ml DMEM-KO 130 ml Knockout Serum Replacement Xeno-Free, 1x NEAA, 1x Glutamax, 0.01 mM β-mercaptoethanol (Sigma). KSR media containing small molecule inhibitors

LDN193189 (100 nM) and SB-431542 (10 µM) was added to cells from day 0 to 3 to drive anterior neuroectoderm specification. From day 3, CHIR99021 (3 µM), DAPT (10 µM) and SU5402 (10 µM) were also added to further promote neural crest phenotypes. N2B27 media was progressively phased in every two days from D4. N2B27 Media was prepared as 500 ml Neurobasal medium, 5 ml N2 supplement, 10 ml B27 supplement without vitamin A, 0.01mM β-mercaptoethanol (Sigma) and 1x Glutamax. On day 7, inhibitors LDN193189 and SB-431542 were no longer used, while CHIR99021, DAPT, and SU5402 continued to be added. On day 11 cells were reseeded at 150,000 cells/cm2 in maturation media containing N2B27 media with human-b-NGF, BDNF, NT3 and GDNF (each at 25 ng/ml). Mitomycin-C treatment (1 µg/ml) was used once at day 14 for 2 hrs to reduce the non-neuronal population. Cells were differentiated in T25 flasks for RNA and nuclei isolation, and onto coverslips and 96-well plates for electrophysiology and $Ca^{2+}$-flux assays.

### P1 protocol details

All reagents and concentrations used were identical to the P2 protocol; the difference was timing of addition. Clump-passaged iPSCs were single-cell seeded in mTeSR1 iPSC (StemCell Technologies, Vancouver) media on growth factor-reduced Matrigel (BD Biosciences) 48 hours prior to neural induction (day 0). KSR media containing LDN193189 and SB-431542 was added to cells from day 0 to 5. From day 5, CHIR99021, DAPT and SU5402 were also added. As for the P2 protocol, cells were reseeded on day 11, and treated with Mitomycin-C on day 14.

## Single-cell RNA sequencing

Blood-derived iPSCs from a single individual, who was not a HIPSCI donor, were differentiated to IPSDSNs in 3 batches using the P2 protocol, and were matured for 8 weeks. Dissociated cells were loaded onto a Fluidigm C1 for automatic cell separation, reverse transcription and amplification.  Libraries were prepared from C1 chambers containing single cells using the Illumina Nextera XT kit.  These were quantified with the Qubit dsDNA HS assay (Thermo Fisher) and KAPA Library Quantification Kit (KAPA Biosystems) and size-checked with Agilent Bioanalyser DNA 1000.  Libraries were 96-way multiplexed and sequenced by Illumina Nextseq500 (2x75bp). Reads were aligned to GRCh38 and Ensembl 80 transcript annotations using STAR v2.4.0d with default parameters. We excluded 9 cells expressing fewer than 20% of the ~56,000 quantified genes, and then used SC3[14] to cluster the remaining 177 cells based on expression counts. We examined alternative numbers of clusters from k=2 to 5. With two clusters, marker genes clearly identified one cluster (111 cells) as neuronal, whereas the other cluster (66 cells) had high expression of extracellular matrix genes reminiscent of fibroblasts. With 3 and 4 clusters, the sensory-neuronal cell cluster remained unchanged, and the fibroblast-like cluster became further subdivided. This suggests that a majority of the cells in this sample were terminally differentiated into sensory neurons, whereas the remaining cells were more heterogeneous in their gene expression.

To display marker gene expression (Figure 2a), we used DESeq2's variance stabilizing transformation, and then R's "scale" function to mean-center and normalize expression values across cells, and plotted the result using the pheatmap R package. To compare gene expression between single-cell clusters and bulk RNA-seq samples (Figure 4), we computed the mean FPKM for each gene separately in single neurons and fibroblast-like cells. We subsetted to genes with nonzero expression in at least one GTEx tissue and in at least one of our tissues (iPSC, DRG, IPSDSN bulk, IPSDSN single cells), and computed the Spearman correlation between each pair of tissues.

## Genotypes

We obtained imputed genotypes for all of the samples from the HIPSCI project. We used CrossMap (http://crossmap.sourceforge.net/) to convert variant coordinates from GRCh37 to GRCh38, and used bcftools (http://samtools.github.io/bcftools/) to retain only bi-allelic variants (SNPs and indels) with INFO > 0.8 and MAF > 0.05 in the 97 samples used for QTL calling.

## RNA sequencing

The 131 RNA samples corresponded with 103 unique HIPSCI donors, as some samples were differentiation or RNA-extraction replicates. One sample failed in sequencing and was excluded. For QTL analyses, reads for each sample were aligned to GRCh38 and Ensembl 79 transcript annotations using STAR v2.4.0j with default parameters. Using VerifyBamID v1.1.2 (Jun et al. 2012) we identified 5 mislabeled RNA samples for which the matching genotypes could be determined, as well as two samples with no matching genotypes and which were thus excluded. For comparisons among tissues, reads were aligned to the 1000 Genomes GRCh37d5 reference with Gencode v19 transcript annotations using STAR 2.5.3a.

## Gene expression quantification, quality control and exclusions

Gencode Basic transcript annotations, GRCh38 release 79, were downloaded from www.gencodegenes.org. Gene expression was counted for uniquely mapping reads using featureCounts (v1.5.0) (Liao, Smyth, and Shi 2014) with options (-s 2 -p -C -D 2000 -d 25). A median of 45 million reads were generated per sample, with median 32.8 million reads (72%) uniquely mapping and assigned to genes. After excluding short RNAs and pseudogenes, we normalised expression counts for 35,033 genes using the R package cqn v5.0.2 (Hansen, Irizarry, and Wu 2012).

We determined pairwise correlation between samples using normalized counts for 14,215 expressed genes (CQN > 1) and the first five principal components of gene expression against each other. We excluded four outlier samples from subsequent analyses, leaving 126 samples from 97 donors. For QTL calling, replicate BAM files from same donor were merged together using samtools.

To assess gene expression replicability, we determined the spearman correlation coefficient of CQN-normalized expression between samples across all genes for (a) extraction replicates, (b) differentiation replicates, and (c) all possible pairs of samples from different donors, and plotted the histogram of correlation coefficients in Figure 8a.

## DRG samples and sequencing

Human tissue acquisition and handling was performed at Pfizer Neuroscience and Pain Research Unit in accordance with regulatory guidelines and ethical board approval. Postmortem human DRG were obtained in dissected form from Anabios or as an encapsulated sheath together with sensory/afferent axons from National Disease Research Interchange and were subsequently dissected to isolate the cell-body rich ganglion. The tissue was homogenised in QIAzol Lysis Reagent according to weight and processed according to the manufacturer's instructions for the Qiagen RNeasy Plus lipid-rich kit. RNA was prepared with the Illumina TruSeq Stranded mRNA Library Prep Kit and sequenced (2x100 bp reads) on Illumina HiSeq 2500. Reads were aligned to GRCh37 using STAR and gene counts and FPKMs obtained using featureCounts and Ensembl v75 gene annotations.

## ATAC library preparation and sequencing

### Nuclei isolation

Media was removed from T25 flasks and washed twice with 10 mL of room temperature D-PBS -/-. The adherent neuronal cultures were lifted by treating with 3 mL of Accutase (Millipore – SCR005) at room temperature for four minutes. The Accutase was quenched by adding 6 mL of 2% foetal bovine serum in D-PBS. The cells were transferred to a 15 mL conical tube and centrifuged at 300 g for 5 minutes at 4°C. The cell pellet was resuspended in 1 mL of ice-cold sucrose buffer (10 mM tris-Cl pH 7.5, 3 mM $CaCl_2$, 2 mM $MgCl_2$ and 320 mM sucrose) and pipetted briefly to break up the large clumps before incubating on ice for 12 minutes. 50 µL of 10% Triton-X 100 was added to the sucrose-treated cells and mixed briefly before incubating on ice for a further 6 minutes. Nuclei were released by performing

30 strokes with a tight dounce homogeniser on ice. Approximately 1 x $10^5$ nuclei were transferred to a 1.5 mL microfuge tube and centrifuged at 300 g for 5 minutes at 4 °C. All traces of the lysis buffer were removed from the nuclei pellet.

### Tagmentation and sequencing

The tagmentation and PCR methods used were as described in (Kumasaka, Knights, and Gaffney 2016), based on the Nextera tagmentation master mix (Illumina FC-121-1030). To remove excess unincorporated primers, dNTPS and primer dimers we used Agencourt AMPure XP magnetic beads (Beckman Coulter A63880), followed by size selection using 1% agarose TAE gel electrophoresis, selecting library fragments from 120 bp to 1 kb. Gel slices were extracted with the MinElute Gel Extraction kit (Qiagen 28604), eluting in 20 µL of Buffer EB. A total of 31 ATAC-seq libraries each prepared with a unique Nextera i5 and i7 tag combination were pooled. Index tag ratios were assessed by a single MiSeq run and were balanced before being sequenced at two per lane with paired-end reads (2x75) on a HiSeq with V4 chemistry. However, rebalancing did not appear to work correctly, as the number of reads varied from a minimum of 17 million to a maximum of 987 million. However, 22 samples had over 100 million reads, and 30 samples had over 40 million reads. Across samples, a median of 56% of reads mapped to mitochondrial DNA.

### Read alignment

We aligned reads to GRCh38 human reference genome using bwa mem v0.7.12. Reads mapping to the mitochondrial genome and alternative contigs were excluded. As for RNA-seq data, we used VerifyBamID v1.1.2 (Jun et al. 2012) to detect sample swaps. This revealed one mislabeled sample, which we then corrected. We used Picard v1.134 MarkDuplicates (https://broadinstitute.github.io/picard/) to mark duplicate fragments.

### Peak calling

We used MACS2 v2.1.1 (Zhang et al. 2008) to call ATAC-seq peaks for individual samples with parameters ‘--nomodel --shift -25 --extsize 50 -q 0.01’. We defined a consensus set of peaks as regions in which peaks overlapped in at least 3 samples. At regions of overlap, the consensus peak was defined as the union of overlapping peaks. This resulted in 381,323 peaks, with 98% of peaks ranging from 82 - 1191 base pairs.

## PCA plot clustering samples with GTEx tissues

We downloaded the GTEx v6 gene RPKM file (GTEx_Analysis_v6_RNA-seq_RNA-SeQCv1.1.8_gene_rpkm.gct.gz) as well as sample metadata (GTEx_Data_V6_Annotations_SampleAttributesDS.txt) from the GTEx web portal (http://www.gtexportal.org/home/datasets). We computed RPKMs for all genes for the 28 DRG samples, the 119 sensory neuron samples after QC exclusions, and 239 HIPSCI IPSC samples. We used genes that were quantified in all of these sample sets, and where at least 50 GTEx samples had RPKM > 0.1. We passed log2(RPKM + 1) for 8,553 GTEx samples to the bigpca R package to compute the first 5 PCs using the SVD method. We determined sample loadings for each PC using the PC weights and log2(RPKM + 1) values for GTEx samples and for our in-house samples, and plotted PC1 vs. PC2 values as Figure 1b.

## Highly variable genes in IPSDSNs and GTEx

For each of the 44 GTEx tissues, as well as IPSDSNs, DRG, and HIPSCI iPSCs, we calculated the coefficient of variation (CV) of each gene's RPKM expression among samples of the same tissue (SMTSD in GTEx metadata). In each tissue, we subsetted the genes considered to those expressed at RPKM > 1. We plotted the distribution of CVs across all genes for each tissue as Figure 3a.

We used GeneTrail2 (https://genetrail2.bioinf.uni-sb.de) to do a gene set over-representation analysis for the top 1000 most variable genes in IPSDSNs by CV (Supplementary Table 4). Similarly, gene set over-representation analysis in E8-IPSDSN subsets was done using Genetrail2 and the top 1000 most variable genes with RPKM > 1 (Supplementary Table 11).

## Variance components analysis

For Figure 12a, we selected the 119 QC-passed samples, and used DESeq2 to get FPKM values for each gene after size factor normalization. We included all genes with mean FPKM > 1, and input log2-transformed counts per sample into the variancePartition Bioconductor R package. For Figure 12b, we used 18 samples for which we had 3 differentiation replicates from each of 6 donor cell lines; all 6 iPSC lines were from females and had been cultured in E8 medium. We therefore included only donor and differentiation in the design formula.

## Estimation of neuronal purity

We used Cibersort (Newman et al. 2015) to estimate the fraction of RNA from neuronal cells in our bulk RNA-seq samples. We used the 14,786 genes with mean CQN expression > 0 in bulk RNA samples, and retrieved raw counts for these genes in our scRNA-seq data. We labeled the single cell counts as "neuron" or "fibroblast-like" based on the SC3 clustering, and used these as reference samples for CIBERSORT to generate custom signature genes. We used raw expression counts for the same genes for our 126 bulk RNA-seq samples as the mixture file for CIBERSORT to use in estimating the relative fractions of neuron and fibroblast-like cell RNA.

## Electrophysiological recordings

Six coverslips per line were placed singularly into a 12-well plate and washed 1x with 1 ml DPBS (+/+). The coverslips were then coated with 1 ml of 0.33 mg/ml growth factor reduced matrigel for > 3 hr at room temperature. D14 cells were prepared at 1.6e6/ml in 15 ml media, then diluted in NB media to create a 0.3e6/ml suspension. The coverslips were transferred into a 12-well plate and 1 ml cell suspension was added. Plates were incubated at 37°C (5% CO2) for 24hr, after which the coverslips were transferred to a 12-well plate with 2 ml media. Cells were treated with Mitomycin C (0.001 mg/ml for 2hr hours at 37°C) post-plating on days 4 and 10. Media was changed twice weekly.

Patch-clamp experiments were performed in whole-cell configuration using a patch-clamp amplifier 200B for voltage clamp and Multiclamp 700A or 700B for current clamp controlled by Pclamp 10 software (Molecular Devices). Experiments were performed at 35°C or 40°C controlled by an in-line solution heating system (CL-100, Warner Instruments). Temperature was calibrated at the outlet of the in-line heater daily before the experiments. Patch pipettes had resistances between 1.5 and 2 MΩ. Basic extracellular solution contained (mM) 135 NaCl, 4.7 KCl, 1 $CaCl_2$, 1 $MgCl_2$, 10 HEPES and 10 glucose; pH was adjusted to 7.4 with NaOH. The intracellular (pipette) solution for voltage clamp contained (mM) 100 CsF, 45 CsCl, 10 NaCl, 1 $MgCl_2$, 10 HEPES, and 5 EGTA; pH was adjusted to 7.3 with CsOH. For current clamp the intracellular (pipette) solution contained (mM) 130 KCl, 1 $MgCl_2$, 5 MgATP, 10 HEPES, and 5 EGTA; pH was adjusted to 7.3 with KOH. The osmolarity of solutions was maintained at 320 mOsm/L for extracellular solution and 300 mOsm/L for intracellular solutions. All chemicals were purchased from Sigma. Currents were sampled at 20 kHz and filtered at 5 kHz. Between 80% and 90% of the series resistance was compensated to reduce voltage errors. Rheobase was measured in current clamp mode by injecting increasing 30 milliseconds current steps until a single action potential was evoked. Intersweep intervals were 2 seconds. Current clamp data was analyzed using Spike2 software (Cambridge Electronic Device, UK) and Origin 9.1 software (Originlab).

## Correlation of iPSC and IPSDSN gene expression with cell culture conditions

We selected the 106 IPSDSN samples differentiated with the P2 protocol, as well as the 87 iPSC samples these were derived from and for which we had RNA-seq data, and we used DESeq2’s variance stabilising transformation on the raw gene expression counts. We computed the first 5 principal components of gene expression separately in iPSC and IPSDSNs, and used corrplot to compute pairwise correlations among these PCs and sample metadata: gender, iPSC passage number, iPSC culture conditions (wasFeeder), iPSC PluriTest score, IPSDSN fibroblast content, and IPSDSN processing date.

We determined differentially expressed genes between feeder-iPSCs and E8-iPSCs using DESeq2 (Love, Huber, and Anders 2014), using expression counts for genes with median FPKM > 0.1 across iPSC samples (Supplementary Table 5). We removed associations driven by outliers, defined as a maximum Cook's distance >= 5. Similarly, we determined differentially expressed genes in IPSDSNs derived from either feeder-iPSCs or E8-iPSCs (Supplementary Table 8), again for genes with median FPKM > 0.1. We used GeneTrail2 (https://genetrail2.bioinf.uni-sb.de) to do a gene set over-representation analysis for the 717 genes with expression at least 2-fold higher in feeder-iPSCs than E8-iPSCs, and similarly for the 631 genes at least 2-fold higher in E8-iPSCs (Supplementary Tables 6, 7). We did an equivalent gene set over-representation analysis for the 1,159 genes with expression at least 2-fold higher in IPSDSNs differentiated from feeder-iPSCs, and also for the 958 genes at least 2-fold higher in IPSDSNs from E8-iPSCs (Supplementary Tables 9, 10).

To determine genes upregulated on differentiation from iPSCs to IPSDSNs, we first selected the 19,658 genes with expression FPKM > 1 in at least two samples (iPSC or IPSDSN). We used DESeq2 as before, removing genes with maximum Cook's distance > 5, identifying 4,246 differentially expressed genes at FDR 1%.

## QTL calling

### Expression QTLs

To call cis-eQTLs we used RASQUAL (Kumasaka, Knights, and Gaffney 2016), which leverages allele-specific reads in heterozygous individuals to improve power for QTL discovery, while accounting for reference mapping bias and a number of other potential artifacts. With RASQUAL a feature is defined by a set of start and end coordinates; for calling a gene eQTL these are the start and end coordinates for exons, whereas for an ATAC-seq peak these are the peak coordinates. RASQUAL requires as input the allele-specific read counts at each SNP within a feature. We used the Genome Analysis Toolkit (GATK) program ASEReadCounter (Castel et al. 2015) with options '-U ALLOW_N_CIGAR_READS -dt NONE --minMappingQuality 10 -rf MateSameStrand' to count allele-specific reads at SNPs (and not indels). We then annotated the AS read counts in the INFO field of the VCF used as input for RASQUAL.

We used RASQUAL's makeCovariates.R script to determine principal components (PCs) to use as covariates, which determined 12 PCs as appropriate from the expression count data.

We ran RASQUAL separately for each of 35,033 genes (19,796 protein-coding genes and 15,237 noncoding RNAs), passing in VCF lines for all SNPs and indels (MAF > 0.05, INFO > 0.8) within 500 kb of the gene transcription start site. We used the --no-posterior-update option in RASQUAL, as we found that not doing so led to some genes having miniscule p values in permuted data. To correct for multiple testing we used permutations; however, because RASQUAL is computationally intensive, it would not be possible to run a thousand or more permutations for every gene. Therefore we used an approach to balance power and computational time. To correct for the number of SNPs tested per gene, we used EigenMT (Davis et al. 2016) to estimate the number of independent tests per gene, and then performed Bonferroni correction on a gene-by-gene basis. To estimate the false discovery rate (FDR) across genes, we used the --random-permutation option of RASQUAL and re-ran it once for every gene, saving the minimum p value (after eigenMT correction) of the SNPs tested for each gene. This gave a distribution of minimum p values across genes for the permuted data. To determine the FDR for eQTL discovery at a given gene, we used R to compute (#permuted data min p values < p) / (#real data min p values < p), where p is the minimum p value among SNPs for the gene in question. With this procedure we obtained 3,778 genes with a cis-eQTL at FDR 10% (2,628 at FDR 5%).

For QTL calling with FastQTL, we first computed principal components from the CQN-transformed gene expression matrix (cqn v5.0.2 (Hansen, Irizarry, and Wu 2012)). We ran FastQTL with permutations 31 separate times, in each run including the first N principal components (N=0...30) as covariates. For each run we used a cis-window of 500 kb, and included SNPs and indels with MAF > 0.05, INFO > 0.8, as we did for RASQUAL. We plotted the number of eGenes found in each of these runs, which plateaued and remained relatively stable at ~1,400 eGenes (FDR 10%) when anywhere from 16 to 30 PCs were used. We arbitrarily chose to use the FastQTL run with 20 PCs in downstream analyses.

**ATAC QTLs**

As we did for gene expression, we used featureCounts v1.5.0 to count fragments overlapping consensus ATAC-seq peaks and ASEReadCounter to count allele-specific reads at SNPs (and not indels) within peaks. We ran RASQUAL separately for each of 381,323 peaks, passing in VCF lines for SNPs and indels (MAF > 0.05, INFO > 0.8) within 1 kb of the center of the peak. Since >99.9% of peaks were less than 2 kb in size, this meant that we tested effectively all SNPs within peaks. As we did when calling eQTLs, we ran RASQUAL with the --random-permutation option for every gene, and determined FDR as described above. Note that in this case we used Bonferroni correction based on the number of SNPs tested, without using EigenMT, due to the small size of the windows tested. With this procedure we obtained 6,318 ATAC peaks with a cis-QTL at FDR 10%.

### Splice QTLs

We downloaded LeafCutter from Github (https://github.com/davidaknowles/leafcutter) on April 17, 2016. We used the LeafCutter bam2junc.sh script to determine junction counts for each sample, followed by leafcutter_cluster.py. This resulted in 254,057 junctions in 59,736 clusters. To focus on splicing events likely to be significant, we applied a number of filters, including: (a) removing junctions accounting for less than 2% of the cluster reads, (b) removing introns used (i.e. having at least 1 supporting read) in fewer than 5 samples, (c) retaining only clusters where at least 10 samples had 20 or more reads in the cluster. This yielded a filtered set of 95,786 junctions in 30,591 clusters. We first determined the read proportions for all junctions within alternatively excised clusters. We then Z-score standardised each junction read proportion across samples, and then quantile-normalised across introns. We used this as our phenotype matrix for input to FastQTL to test for associations between intron usage and variants within 15 kb of the center of each intron. We chose a cis-window size of 30 kb (2 x 15 kb) because >91% of introns are < 30 kb in size, and so this tests variants near exon/intron boundaries for the great majority of introns, while maximising power.

We ran FastQTL in nominal pass mode 31 times specifying the first 0 to 30 principal components as covariates, and examined the number of intron QTLs with minimum SNP p value < $10^{-5}$. This showed that the number of QTLs plateaued when 5 PCs were used, and so we used 5 PCs in subsequent runs. We next ran FastQTL with 10,000 permutations to determine empirical p values for each alternatively excised intron. To correct for the number of introns tested per cluster, we used Bonferroni correction on the most significant intron p value per cluster. We then used the Benjamini-Hochberg method to estimate FDR across tested clusters. This yielded 2,079 significant SNP associations for intron usage (sQTLs) at FDR 10%.

For significant sQTLs we used bedtools closest with GRCh38 release 84 to annotate the gene(s) nearest the lead SNP for the association. To ensure we had relevant genes, we filtered the annotation to include only genes where one of the exon boundaries matched the intron boundary for the sQTL.

## Identifying tissue-specific eQTLs

We determined the set of tissue-specific eQTLs using the same procedure and code as in the HIPSCI project (Kilpinen, Helena, et al. 2016). Briefly, we considered the full cis eQTL output of sensory neuron eQTLs and 44 tissues analyzed by the GTEx Project (GTEx Consortium 2015). To enable comparison, lead SNP positions for sensory neuron eQTLs were first lifted back from GRCh38 to GRCh37 using Crossmap (Zhao et al. 2014). For each discovery tissue (including sensory neurons), we tested for the replication of all lead eQTL - target eGene pairs reported at FDR 5%. If the lead eQTL variant was not reported in the comparison tissue, then the best high-LD proxy of the lead variant ($r^2 > 0.8$ in the UK10k European reference panel) was used as the query variant. Replication was defined as the query variant having a nominal eQTL $p < 2.2 \text{x} 10^{-4}$ (corresponding to $p = 0.01 / 45$, where 45 refers to the total number of tissues tested) for the same eGene. We then extracted eGenes for which the lead eQTL did not show evidence of replication in any other tissue ($p > 2.2 \text{x} 10^{-4}$) or could not be tested (i.e. was not measured or reported as expressed in any other tissue).

This analysis gave 954 eGenes where the eQTL is specific to sensory neurons (Supplementary Table 15). We note that some of these "tissue-specific" eGenes could be due to the difference in QTL-calling methods used, notably that we used RASQUAL, a method incorporating both allele-specific and population-level expression variation. Therefore, some of the tissue-specific eGenes we report may actually be present more broadly in GTEx tissues but missed by the linear QTL model used in GTEx. Among the 1,403 eGenes called by FastQTL, 208 were tissue-specific to IPSDSNs.

## Motif enrichment analyses

We used the R Bioconductor package LOLA (Sheffield and Christoph 2015) to identify enrichments in transcription factor binding sites (TFBS) and motifs. We defined three sets of loci to consider for enrichment: 1) tissue-specific eQTL SNPs with a window of 50 bp (+/- 25) around the SNP position, 2) all eQTL SNPs (50 bp window), and 3) all ATAC-seq peaks. For the QTLs we used all GTEx eQTL lead SNPs as the "universe" set against which we tested TFBS for enrichment. For this loaded GTEx eQTLs in R and used the liftOver function from rtracklayer to convert their coordinates to GRCh38. We tested for enrichment against the LOLA core database considering only ENCODE TFBS enrichments (Supplementary Tables 16 and 17). We also tested ATAC-seq peaks for enrichment relative to DNaseHS for many tissues from (Sheffield et al. 2013), which are available in the LOLA catalog. Motif enrichments in ATAC-seq peaks are reported in Supplementary Table 18.

## Power simulations

Gene expression values were normalized to counts per million. We selected the 544 eGenes discovered by RASQUAL at FDR 1% which met the following criteria:

- at least 10 P2-protocol samples homozygous for each allele of the lead eQTL variant,
- mean expression among homozygous carriers was consistent with RASQUAL's reported direction of effect, and
- CV < 2 (this filter removed only 8 eGenes)

For each gene we resampled the normalized expression values, with replacement, from IPSDSN samples to achieve a specified number of samples ($N \in \{4,6,10,20,40\}$) with each homozygous genotype. From 100 such resamplings, we defined the power to discover a given variant's effect as the fraction of cases with $p < 0.05$ from a Wilcoxon rank sum test comparing expression in each genotype category. We determined the allelic fold change between genotypes using RASQUAL's effect size (pi), as:

fold change = max( pi / (1-pi), (1-pi) / pi)

We used ggplot2 with geom_smooth to display the 95% confidence interval around the fitted mean TPR at each parameter combination. As can be seen on the plots, the deviation about this mean for individual genes is larger than the standard error of the mean.

## QTL overlap with GWAS catalog

The GWAS catalog was downloaded from https://www.ebi.ac.uk/gwas/ on 2016-5-08. To determine overlap between variants in the GWAS catalog and our lead QTLs, we first extracted all lead variants (both QTLs and GWAS catalog variants) from the full VCF file. We used vcftools v0.1.14 (Danecek et al. 2011) to compute the correlation $R^2$ between all lead variants within 500 kb of each other among our samples. We determined overlap separately for eQTLs, sQTLs, and ATAC QTLs, and retained only overlaps with $R^2 > 0.8$ between lead variants. Note that a given GWAS variant may be in LD with an eQTL for more than one gene, and vice versa, an eQTL for a single gene may be in LD with more than one GWAS catalog entry.

To determine whether our QTL overlaps were enriched in any specific GWAS catalog traits relative to other traits, we computed overlap with all GWAS catalog SNPs ($p < 5x10^{-8}$) but sought to eliminate redundant overlaps. For traits that were reported with differing names (e.g. "Alzheimer's disease (cognitive decline)" and "Alzheimer's disease in APOE e4-carriers"), we grouped these into a single trait name (e.g. "Alzheimer's disease"). We then sorted overlaps by decreasing LD $R^2$, and kept the single overlapping QTL with the highest $R^2$ for each GWAS catalog entry. Similarly, we removed duplicates with the same reported

GWAS catalog SNP and trait, such as when successive GWAS of the same trait report the same SNP association. We counted the number of such unique GWAS-QTL overlaps separately for eQTLs, sQTLs, and caQTLs, and we report these in Table 1. To avoid bias due to correlation between GWAS power and LD patterns, we restricted our analysis to the 41 traits with at least 40 GWAS catalog associations. We then considered the binomial probability of the observed overlap with each trait, with the expected overlap frequency being the proportion of QTL overlaps among all trait associations (6.2%). After correcting for multiple testing, no traits showed significantly greater overlap with our QTL catalog than other traits.

To test for overall enrichment of QTLs overlapping with GWAS catalog SNPs, we used vcftools to identify 1000 Genomes SNPs in LD $R^2 > 0.8$ with a GWAS catalog SNP, and removed duplicate SNPs. We used our IPSDSN eQTL lead SNPs as input to SNPsnap (https://data.broadinstitute.org/mpg/snpsnap/), and computed 1000 random sets of SNPs matched for LD partners, MAF, gene density, and distance to nearest gene. IPSDSN eQTL lead SNPs had more overlaps (92) with GWAS catalog + $R^2$>0.8 SNPs than did any of the matched sets (median: 58, range 37-87).

## Acknowledgments

The iPSC lines were generated under the Human Induced Pluripotent Stem Cell Initiative (HIPSCI) funded by a grant from the Wellcome Trust and Medical Research Council, supported by the Wellcome Trust (WT098051) and the NIHR/Wellcome Trust Clinical Research Facility. HIPSCI funding was used for sensory neuron RNA-sequencing. We acknowledge Life Science Technologies Corporation as the provider of Cytotune. Pfizer Neuroscience (Pfizer Ltd.) funded neuronal differentiation, functional assays, single-cell RNA-sequencing, and collection and sequencing of dorsal root ganglion samples. We thank Natsuhiko Kumasaka for help with RASQUAL, and Florian Merkle for comments on the manuscript.

## Conflicts of Interest

SF, RF, CB, AW, MB, EI, LC, SL, AJL, PJW and AGu were all employees of Pfizer at the time the experiments were performed.