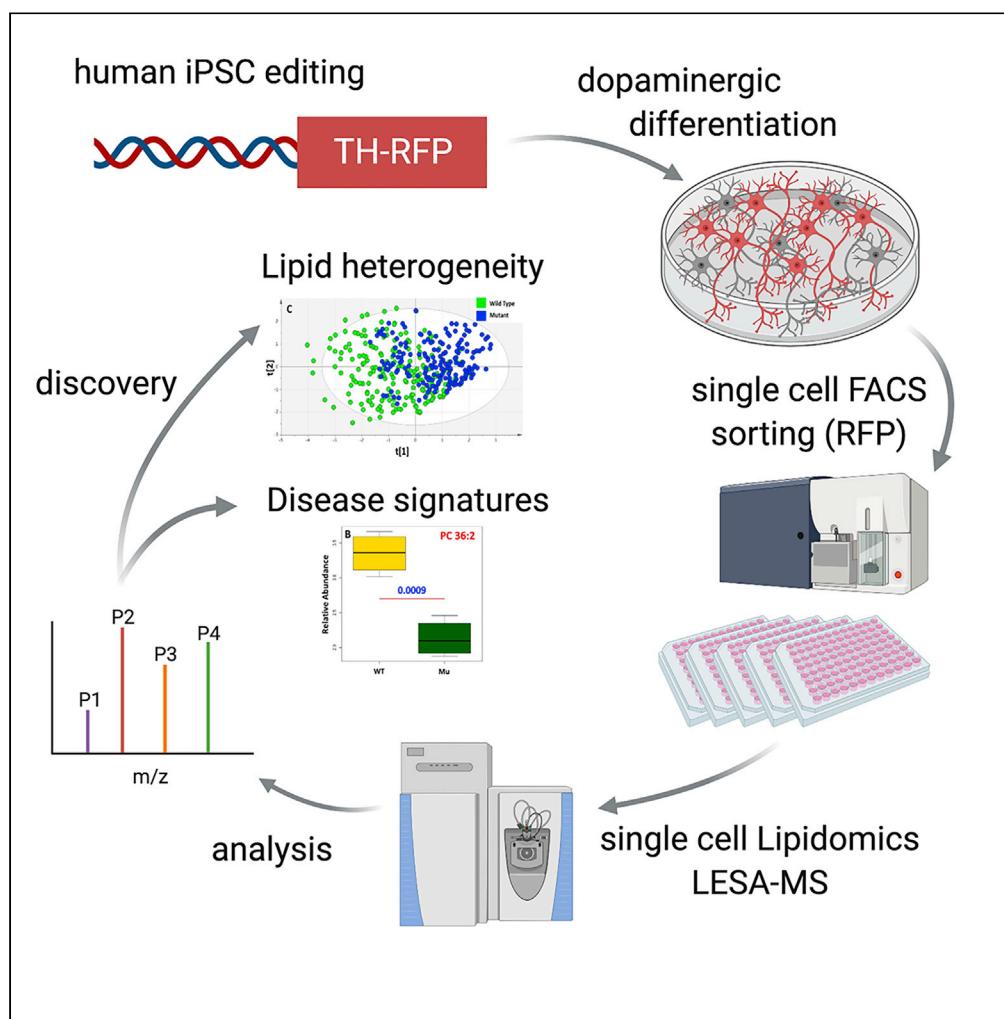


## Article

## Development and Application of High-Throughput Single Cell Lipid Profiling: A Study of SNCA-A53T Human Dopamine Neurons



Stuart G.  
Snowden, Hugo  
J.R. Fernandes,  
Josh Kent, ...,  
Sarah F. Field,  
Emmanouil  
Metzakopian,  
Albert Koulman

em698@medsch.cam.ac.uk  
(E.M.)  
ak675@medsch.cam.ac.uk  
(A.K.)

**HIGHLIGHTS**  
Combining FACS and  
LESA-MS to establish  
high-throughput single  
cell lipid profiling

Lipid differences found  
within and between  
populations of human  
dopamine neurons

Inter-cell lipid  
heterogeneity is increased  
in SNCA-A53T dopamine  
neurons

Identification and  
isolation of human iPSC-  
dopamine neurons with  
a TH-RFP reporter

Snowden et al., iScience 23,  
101703  
November 20, 2020 © 2020  
The Author(s).  
<https://doi.org/10.1016/j.isci.2020.101703>



## Article

# Development and Application of High-Throughput Single Cell Lipid Profiling: A Study of SNCA-A53T Human Dopamine Neurons

Stuart G. Snowden,<sup>1,2,4</sup> Hugo J.R. Fernandes,<sup>3,4</sup> Josh Kent,<sup>1,3</sup> Stefanie Foskolou,<sup>3</sup> Peri Tate,<sup>2</sup> Sarah F. Field,<sup>3</sup> Emmanouil Metzakopian,<sup>3,5,\*</sup> and Albert Koulman<sup>1,5,6,\*</sup>

## SUMMARY

**Advances in single cell genomics and transcriptomics have shown that at tissue level there is complex cellular heterogeneity. To understand the effect of this inter-cell heterogeneity on metabolism it is essential to develop a single cell lipid profiling approach that allows the measurement of lipids in large numbers of single cells from a population. This will provide a functional readout of cell activity and membrane structure. Using liquid extraction surface analysis coupled with high-resolution mass spectrometry we have developed a high-throughput method for untargeted single cell lipid profiling. This technological advance highlighted the importance of cellular heterogeneity in the functional metabolism of individual human dopamine neurons, suggesting that A53T alpha-synuclein (SNCA) mutant neurons have impaired membrane function. These results demonstrate that this single cell lipid profiling platform can provide robust data that will expand the frontiers in biomedical research.**

## INTRODUCTION

The recent developments of single cell approaches have demonstrated the importance of measuring and understanding how cellular heterogeneity affects tissues and organs to fully understand biological processes (Heintzman et al., 2009; Tiklova et al., 2019; Tirosh et al., 2016; Deng et al., 2014). Single cell genomic and transcriptomic approaches have enabled previously unanswered questions to be addressed. For example, Tiklova et al. performed single cell RNA sequencing of *pitx3* mouse midbrain dopamine neurons and identified seven distinct neuronal subtypes, five of which expressed dopaminergic markers and five expressed glutamatergic and GABAergic markers (Tiklova et al., 2019). However, to date there has been limited success in the development of practical metabolic phenotyping tools that can be applied to capture the metabolic heterogeneity at a cellular level.

Lipids are the most abundant class of metabolites in the cell, and the measurement of lipids by mass spectrometry in bulk samples is well described. A handful of studies have previously described proof of principle for single cell lipid profiling (Evers et al., 2019); however, these are not platforms capable or suitable for robust high-throughput readouts of cell activity. Ellis et al. used a low-throughput approach where cell droplets were printed onto a glass slide, which were imaged and analyzed using liquid extraction surface analysis coupled with mass spectrometry (LESA-MS) (Ellis et al., 2012). Neumann et al. used MALDI to measure lipids from a large number of putative single cells from a section of rat cerebellum. However, without imaging all samples lack cell-type specificity and could not guarantee that each sample contained a single cell and not clusters of cells, leading to wide divergence in the number of lipids measured per sample (Neumann et al., 2019). Most single cell mass spectrometry platforms have focused on analyzing immobilized cells; however, Standke et al. (2019) developed an integrated cell manipulation platform that enables single cells to be analyzed from solutions, such as bodily fluids, with minimal sample preparation. More complex derivatization approaches have also been described. Thiele et al. (2019) reported a method for tracing lipid metabolism in cell culture dilutions using click chemistry. This provided detailed coverage but was also unable to give certainty that bona fide single cells were analyzed. Together, these studies are convincing proofs of principle, but they do not represent mature platforms. To achieve this, it is necessary to standardize sample handling, single cell isolation and to establish robust strategies for quality control to ensure that the generated data can be meaningfully compared.

<sup>1</sup>Core Metabolomics and Lipidomics Laboratory, Metabolic Research Laboratories, Institute of Metabolic Science, University of Cambridge, Level 4 Pathology, Cambridge Biomedical Campus, Cambridge CB2 0QQ, UK

<sup>2</sup>Department of Biological Sciences, Royal Holloway University of London, Egham, Surrey TW20 0EX, UK

<sup>3</sup>UK Dementia Research Institute, University of Cambridge, Department of Clinical Neurosciences, Cambridge Biomedical Campus, Cambridge CB2 0AH, UK

<sup>4</sup>These authors contributed equally

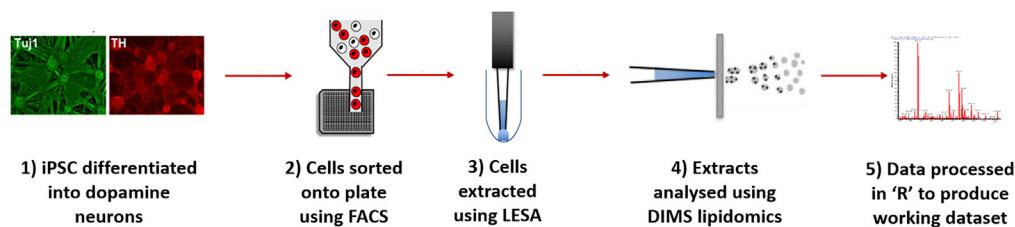
<sup>5</sup>These authors contributed equally

<sup>6</sup>Lead Contact

\*Correspondence:  
[\(E.M.\),](mailto:em698@medsch.cam.ac.uk)  
[\(A.K.\)](mailto:ak675@medsch.cam.ac.uk)

<https://doi.org/10.1016/j.isci.2020.101703>





**Figure 1. Graphical Representation of the Analytical Single Cell Pipeline Used in this Study**

The brain is a lipid-rich organ, and neuronal lipid metabolism regulates a range of biological processes including cell signaling and structural integrity (Tracey et al., 2018). Perturbations of lipid metabolism have been associated with the pathogenesis of Parkinson disease (PD), from genetic risk factors to altered brain lipid profiles (Do et al., 2011; Fabelo et al., 2011). PD is a common neurodegenerative disorder characterized by the loss of dopamine neurons and the accumulation of Lewy bodies, which are composed primarily of alpha-synuclein protein (Spillantini et al., 1997). The relevance of SNCA for the pathology of PD is further highlighted by the fact that mutations in this gene, such as the A53T mutation, lead to the development of PD (Polymeropoulos et al., 1997). Although the underlying mechanisms of this association are unclear, alpha-synuclein has been shown to modulate lipid metabolism in PD models (Golovko et al., 2007; Sharon et al., 2003; Zambon et al., 2019).

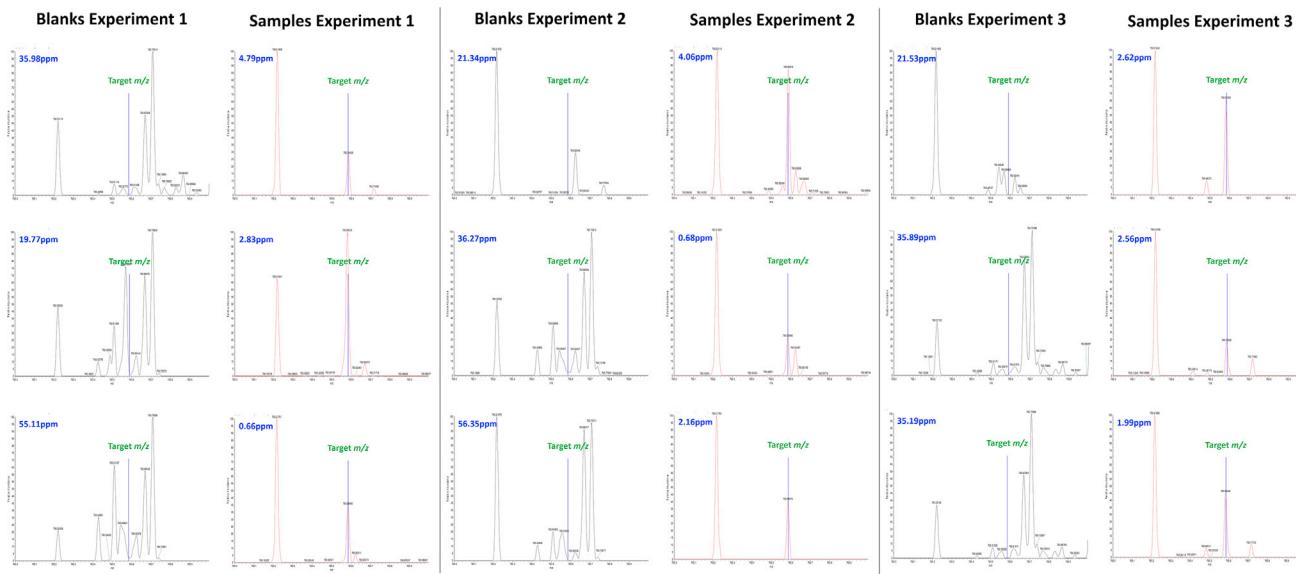
Here we describe the development of a high-throughput (280 single cells/day) untargeted single cell lipid profiling platform, detailing strategies for single cell isolation, data generation, automated signal identification, and quality control (Figure 1). We further demonstrate that this is a mature platform by validating the measured lipids in three independent cohorts and through its application to the analysis of human dopamine neurons derived from induced pluripotent stem cells (iPSCs). We believe this important technological advance for single cell lipid analysis will provide a unique opportunity to address important biomedical questions across various fields of research.

## RESULTS

### Measuring Lipids in Single Cells

Protocols for the differentiation of *in vitro* human dopamine neurons are known to generate heterogeneous neuronal populations (La Manno et al., 2016). To overcome this, we engineered human wild-type (WT) iPSC with a tyrosine hydroxylase (TH) red fluorescence protein (RFP) as previously described (Xia et al., 2017), with modifications. To express tagRFP-T under the influence of the endogenous TH promoter we introduced a P2A-tagRFP-T donor plasmid with homologous arms flanking the TH stop codon and an EF1A-Puromycin selection flanked by *LoxP* sites for drug selection. The targeting plasmid was introduced into iPSCs together with a single guide RNA targeting the 3' end of the TH gene and Cas9 RNP by nucleofection to cause a double-strand break near the stop codon. After Puromycin selection, transient expression of Cre was used to excise the Puromycin cassette. TH catalyzes the hydroxylation of L-DOPA (the rate-limiting step in the synthesis of dopamine) and is considered a bona fide marker for dopamine neurons. After differentiation into dopamine neurons (Siddiqi et al., 2019; Kriks et al., 2011), fluorescence-activated cell sorting (FACS) was used to sort cells based on RFP expression. RFP-positive dopamine neurons were individually sorted into multi-well plates, followed by LES-MS.

In spectra generated from single cell samples (Figure 2) there was a clear signal for PC 34:1 with an average deviation of 2.6 ppm and with the highest deviation being 4.8 ppm. However, in the blanks there was no signal for PC 34:1 with a mean deviation between the closest signal and the target mass of 30.8 ppm and the lowest deviation of 15.1 ppm. Comparison of the relative abundance and deviation of the single cell samples to the blanks showed that the samples have a significantly higher abundance and lower deviation (Table S1). Having identified PC 34:1 and other lipids in initial analyses we performed two additional independent experiments to validate these findings. Across these experiments the average deviation in the blanks was 17.8 ppm with only 4 of 63 blanks having a signal falling within 10 ppm of the target mass, whereas in single cells the average deviation was 1.8 ppm and the highest deviation was 7.8 ppm with signals detected in over 90% samples.



**Figure 2. Spectra Showing the Presence of a Signal for PC 34:1 in Single Cell Samples and Its Absence from Extraction Blanks in Three Independent Experiments**

Each spectra shows the location of the “target  $m/z$ ” and the deviation from the “target  $m/z$ ” of the closest measured signal.

Next, we determined how many lipids could be detected in the three datasets. In each dataset we applied four QC filters, signal falling within 10 ppm of the expected mass, a signal-to-noise ratio greater than 1.5, a linear relationship between lipid signal and cell number ( $r > 0.5$ ) in a cell “standard” curve, and a signal had to be present in at least 5% of single cell samples. Once we had identified the lipids measured in each of the three datasets we defined three levels of confidence in the putative assignments. Level 1 putative assignments are lipids that were successfully measured in all three datasets, level 2 putative assignments were present in two of the three, and level 3 assigned lipids were only measured in a single experiment. We successfully identified 37 level 1 lipid assignments corresponding to 25 glycerophosphocholines (PCs), 9 glycerophosphoethanolamines (PE), and 3 sphingomyelins (SM) (Table 1). A further 23 level 2 and 21 level 3 putative assignments were also identified including an additional 15 PCs, 22 PEs, and 1 SM as well as 2 ceramides, 1 diglyceride, 3 glycerophosphoserines, and 1 triglyceride (Table S2)

### Quality Control

Having validated the platform, we expanded our method to address a relevant biological question by assessing the impact of the SNCA-A53T mutation in the lipidome of human dopamine neurons at the single cell level resolution. To reduce inherent variability associated with iPSC modeling studies as previously reported (Volpato and Webber, 2020), we generated isogenic SNCA-A53T mutant iPSC lines using CRISPR-Cas9 on a WT background as previously described (Siddiqi et al., 2019; Bruntraeger et al., 2019). Isogenic SNCA-A53T mutant iPSCs were then engineered with a TH-RFP reporter following the same procedure described for WT cells. Differentiation efficiency into dopaminergic neurons was similar for WT and SNCA-A53T iPSCs (Figure 3A). After differentiation, WT and SNCA-A53T individual dopamine neurons were sorted by FACS across 7 plates of 96 wells. For each genotype, a total of 280 single neurons across four biological replicates were sorted with extraction blanks and quality controls added across all plates.

When comparing the abundance of PC 34:1 measured in the QC samples from all 7 plates (Figure 3B) there was a significant ( $p < 0.05$ ) difference in the abundance of this lipid between plates, with the greatest difference seen between plates 3 and 4. However, after the data were normalized to total signal abundance this difference was corrected for with no significant differences ( $p < 0.05$ ) found in the abundance of PC 34:1 between plates (Figure 3C). Principal-component analysis (PCA) of the normalized lipid profiles from QC samples revealed that samples could not be clustered according to different plates, confirming a robust and standardized sample handling and processing (Figure 3D). Analysis of normalized lipid profile data using PCA showed that there was no relationship between injection order and lipid composition (Figure S1).

|         | Experiment 1 | Experiment 2 | Experiment 3 |
|---------|--------------|--------------|--------------|
| PC 30:1 | 33.3         | 57.5         | 24.3         |
| PC 32:0 | 66.7         | 65.8         | 71.6         |
| PC 32:1 | 88.9         | 65.0         | 67.8         |
| PC 32:2 | 11.1         | 65.0         | 45.9         |
| PC 32:3 | 66.7         | 55.0         | 30.0         |
| PC 32:4 | 44.4         | 46.3         | 23.9         |
| PC 34:0 | 66.7         | 15.0         | 30.2         |
| PC 34:1 | 100.0        | 94.9         | 90.2         |
| PC 34:2 | 88.9         | 86.1         | 60.8         |
| PC 34:3 | 66.7         | 88.8         | 68.5         |
| PC 34:4 | 77.8         | 93.7         | 78.2         |
| PC 36:1 | 77.8         | 89.9         | 74.7         |
| PC 36:2 | 88.9         | 88.6         | 88.4         |
| PC 36:3 | 66.7         | 82.2         | 65.3         |
| PC 36:4 | 88.9         | 91.3         | 86.4         |
| PC 36:5 | 66.7         | 78.7         | 37.2         |
| PC 38:3 | 88.9         | 78.8         | 39.2         |
| PC 38:4 | 55.6         | 67.6         | 44.4         |
| PC 38:5 | 88.9         | 77.5         | 65.5         |
| PC 38:8 | 44.4         | 21.3         | 6.3          |
| PC 40:3 | 44.4         | 58.1         | 36.9         |
| PC 40:4 | 22.2         | 33.8         | 27.7         |
| PC 40:5 | 22.2         | 17.5         | 42.6         |
| PC 40:6 | 11.1         | 7.5          | 58.1         |
| PC 40:9 | 11.1         | 5.0          | 30.0         |
| PE 34:0 | 28.6         | 15.0         | 7.7          |
| PE 34:1 | 44.4         | 46.8         | 30.6         |
| PE 34:2 | 22.2         | 16.5         | 13.3         |
| PE 34:3 | 11.1         | 13.9         | 11.3         |
| PE 36:0 | 22.2         | 17.5         | 8.6          |
| PE 36:1 | 44.4         | 32.9         | 30.9         |
| PE 36:3 | 11.1         | 11.4         | 25.7         |
| PE 36:4 | 11.1         | 10.1         | 7.9          |
| PE 38:6 | 22.2         | 15.0         | 18.6         |

**Table 1. Panel of Lipids Successfully Measured from a Single Dopaminergic Neuron**
*(Continued on next page)*

|         | Experiment 1 | Experiment 2 | Experiment 3 |
|---------|--------------|--------------|--------------|
| SM 34:1 | 33.3         | 37.7         | 30.6         |
| SM 36:1 | 22.2         | 29.5         | 18.9         |
| SM 36:2 | 22.2         | 37.5         | 7.9          |

**Table 1. Continued**

Data showing the percentage of cells in which each lipid was measured in each of the three sample sets. PC, glycerophosphocholine; PE, glycerophosphoethanolamine; SM, sphingomyelin.

### Comparing Single Cell and Population Lipid Profiling

Next, we explored the advantages of using single cell lipid profiling over traditional population level approaches in this dopamine neurons dataset. After data processing and normalization, the mean abundance of each lipid was calculated within each biological replicate to obtain a “population” lipid profile. When these “population” profiles were analyzed with partial least squares-discriminant analysis (PLS-DA) ( $R^2X = 0.679$   $R^2Y = 0.752$   $Q^2 = 0.530$  CV-ANOVA =  $4.50 \times 10^{-6}$ ) there was a clear difference between WT and SNCA-A53T populations (Figure 4A). PC 36:2 was the most important lipid driving the PLS-DA model with univariate analysis also identifying a significant difference between genotypes ( $p = 0.0009$ ) (Figure 4B), with PC 36:1 and PC 34:2 also significantly different between genotypes (Figures S2 and S3). This “population” analysis implied significant lipid alterations associated with the A53T mutation and suggested non-overlapping signatures between WT and SNCA-A53T dopamine neurons.

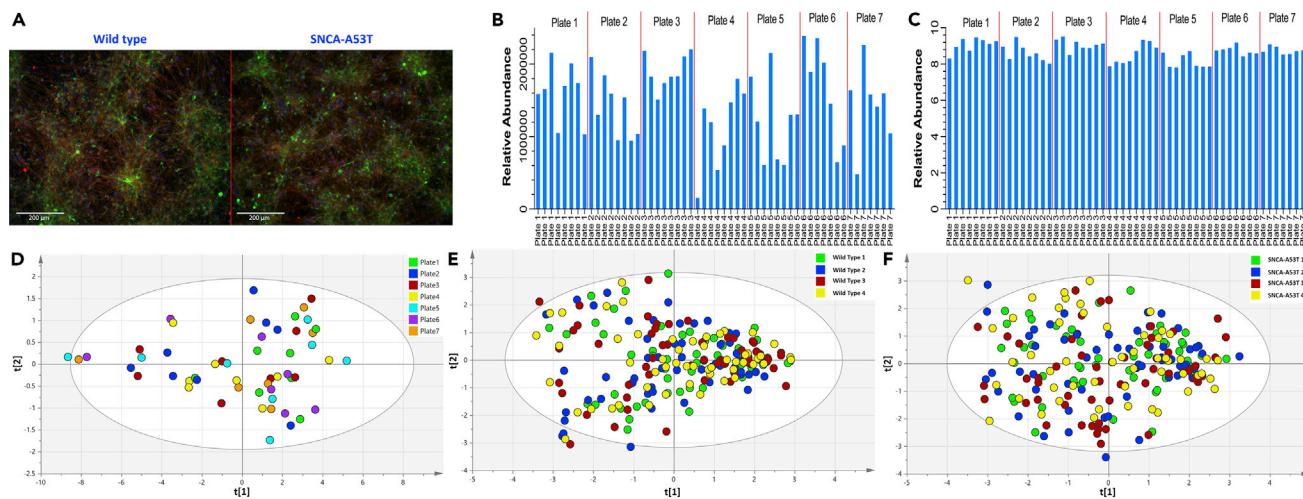
However, when the lipid profiles of individual dopamine neurons were analyzed we found a considerable heterogeneity within both genotypes (Figure 4C). Although the PLS-DA scores plot ( $R^2X = 0.306$   $R^2Y = 0.226$   $Q^2 = 0.195$  CV-ANOVA =  $7.65 \times 10^{-18}$ ) still confirmed differences between WT and SNCA-A53T dopamine neurons, there was some overlap between genotypes. This cellular heterogeneity could be of fundamental importance for a complete understanding of biological function and could only be identified by single cell lipid profiling.

Looking at individual lipids the abundance of PC 36:2 was still significantly different ( $p = 1.4 \times 10^{-7}$ ) between genotypes (Figure 4D), with differences also observed in PC 36:1, PC 34:2, and PC 32:0 (Figures S2–S4). However, there was a high degree of cellular heterogeneity (Figures 4E and 4F) with the A53T mutants showing a high density of neurons with low abundances of PC 36:2 (Figures 4E and 4F).

To further explore this heterogeneity, we stratified samples based on their abundance of PC 36:2 and assessed if the abundance of other lipids mirrored this heterogeneity. This showed that in WT nine lipids were significantly different ( $p < 0.05$ ) between cells with high and low abundances of PC 36:2, whereas eight lipid variables were significant in SNCA-A53T dopamine neurons. Of these, six lipids were common between the genotypes, three were specific to WT, and two were specific to SNCA-A53T including the ratio of total PC to total PE (Table 2). The abundance of PC 36:4 in cells stratified by their PC 36:2 levels revealed similar patterns in both WT and SNCA-A53T populations, with dopamine neurons with high PC 36:2 having high PC 36:4 (Figure 5A). Although there were differences in the abundance and distribution of PC 36:4 in cells with high and low PC 36:2, this was similar between genotypes (Figure 5B). The same was observed when we analyzed the abundance of PE 36:2 with similar patterns observed between genotypes (Figures 5C and 5D); however, in this case neurons with more PC 36:2 have less PE 36:2. Cells with high PC 36:2 had higher levels of PC 34:2 in both genotypes; however, this difference was only significant in WT neurons (Figures 5E and 5F).

### DISCUSSION

This platform is the first single cell approach that enables high-throughput lipid profiling of specific cell types without the need for additional imaging analysis. This coupled with a high degree of standardization and robust quality control has enabled us to apply untargeted lipid profiling to individual cells and explore the functional heterogeneity of cellular populations and the implications for disease pathology. We believe this to be a major technological advance with wide applications for biomedical research.



**Figure 3. Plots Showing the Lipid Abundance in Quality Control Samples from all Seven Plates Both Before and After Batch Correction**

- (A) Immunofluorescence staining of WT iPSC-differentiated dopamine neurons showing expression of pan-neuronal marker beta-3-tubulin (TUJ1) in red, dopamine marker tyrosine hydroxylase (TH) in green, and nuclear DAPI in blue.
- (B) Plot of uncorrected abundance of PC 34:1 in individual QC samples.
- (C) Plot of the corrected abundance of PC 34:1, in individual QC samples.
- (D) PCA scores plot of quality control samples calculated using all corrected lipid data. Mu, SNCA-A53T mutant dopaminergic neurons; PC, glycerophosphocholine; PCA, principal-component analysis.
- (E) PCA scores plot showing the comparability of lipid composition between wild-type cell populations.
- (F) PCA scores plot showing the comparability of lipid composition between SNCA-A53T cell populations.

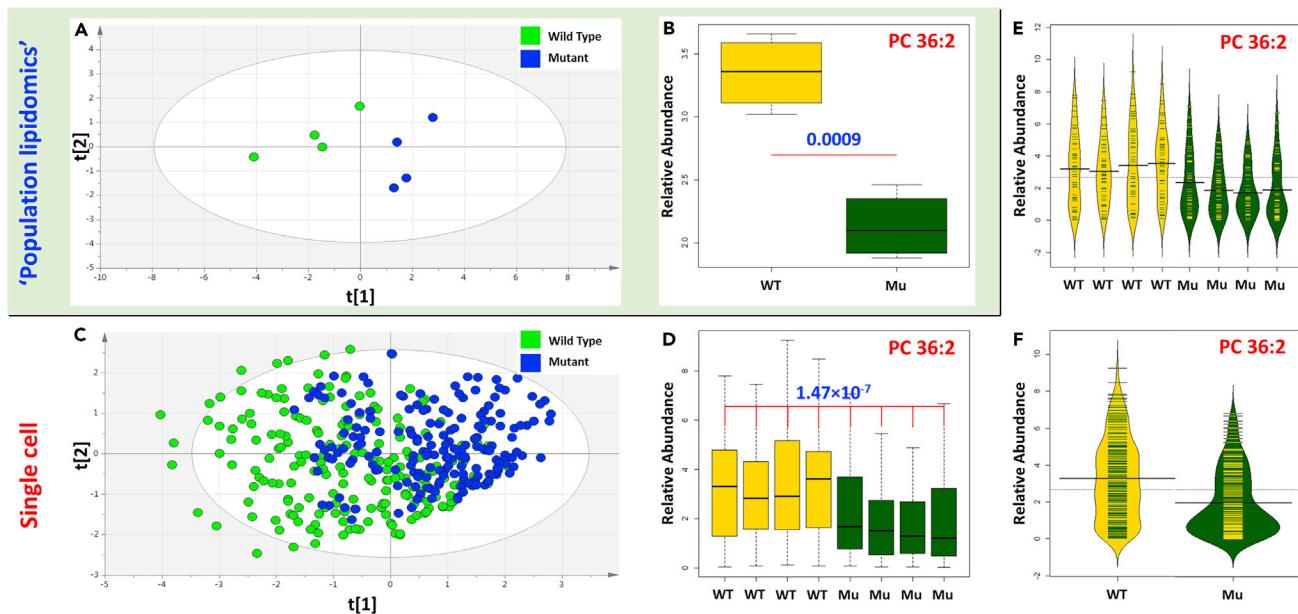
### Cell Location, Throughput, Quality Control, Signal Coverage

The advent of iPSCs offers the opportunity to study otherwise inaccessible cell types such as functional human neurons. However, the limitations of this model need to be acknowledged, in particular the inherent variability associated with iPSC techniques highlighted in previous studies (Volpato and Webber, 2020). To overcome this limitation, we generated isogenic SNCA-A53T iPSCs, a well-established approach in the field of iPSC modeling to reduce cellular variability.

As with single cell genomics and transcriptomics the requirement to analyze large numbers of cells means that these approaches need to be high throughput. In this study, we analyzed four independent replicates of two genotypes of human dopamine neurons. To capture the cellular heterogeneity of these samples we analyzed 70 cells per biological replicate leading to a total of 560 analyzed cells. Our platform enables these 560 samples to be analyzed in 2 days (including the automated data processing), underlining the high-throughput nature of this method.

We used FACS to sort TH-positive dopamine neurons into individual wells using an RFP reporter to improve the throughput and specificity of the platform. The use of the TH-RFP coupled with an appropriate gating strategy provides specificity and ensures that all cells analyzed are individual dopamine neurons. Using FACS also allowed for quick sample isolation or large number of cells without the requirement for additional imaging analysis to confirm cell numbers and identity as previously required.

This need to analyze large numbers of samples requires a high-throughput platform with robust quality control strategies to ensure the quality of the data generated. In the development of this platform we automated as many processes as possible, to both maximize the number of samples that can be analyzed in a given time and to standardize analysis to minimize sources of analytical variation. With cells being sorted onto all seven analytical plates at the same time and stored at  $-80^{\circ}\text{C}$  before analysis it was important to ensure that storage was not effecting the lipid composition of the cells. Although in the raw data batches effects were observed between plates, normalization of the data removed these effects with no plate to plate differences seen in either QC or single cell samples at either the univariate or multivariate level demonstrating that sample storage is not introducing any bias into the lipid composition of the cells. In the normalized data we also showed that there was no relationship between injection order and lipid



**Figure 4. Comparison of the Information Provided by Population and Single Cell Lipid Profiling Approaches**

- (A) PLS-DA scores plot of population level lipid profiling comparing wild-type and A53T mutants (Mu).  
 (B) Boxplot of the abundance of PC 36:2 in wild-type and A53T mutants (Mu); p value calculated using generalized linear model.  
 (C) PLS-DA scores plot of single cell lipid profiling comparing wild-type and A53T mutants (Mu).  
 (D) Boxplots of the abundance of PC 36:2 in each biological replicate, p value calculated using generalized linear model.  
 (E) Beanplot showing the relative distribution of the abundance of PC 36:2 in each biological replicate cells where each line represents the abundance measured in a cell and the width of the outline represents the density of samples.  
 (F) Beanplot showing the relative distribution of the abundance of PC 36:2 in combining all wild-type and all mutants where each line represents the abundance measured in a cell and the width of the outline represents the density of samples. All p values were calculated using generalized linear models.  
 Mu, SNCA-A53T mutant dopaminergic neurons; PC, glycerophosphocholine; PLS-DA, partial least squares-discriminant analysis.

profile composition (Figure S1) demonstrating that the data generated from the first and last samples are comparable.

It is crucial to control for cell size as this is the most important factor controlling the measured lipid abundance, as a larger cell will have greater total signal abundance because it contains more lipids. However, as cell size will also determine the total lipid signal observed, normalizing lipid abundances to the total signal controls for the size of the cell analyzed.

For this method we used LESA as it requires lower solvent volume than traditional extraction protocols producing more concentrated extracts. However, when using LESA-MS-based approaches for single cell lipid profiling locating the cell for analysis is a significant challenge. Although other methods have used imaging approaches to locate the cell for analysis (Ellis et al. 2012; Neumann et al., 2019), this is time consuming and would significantly reduce the throughput of the method. Instead we calibrated the FACS to maximize the chance of a cell being dispensed in the center of the well, with the LESA analysis subsequently performed on that spot; however, in some cases the LESA missed the cell. We identified these failed analysis by searching the generated spectra for signals from the two most abundant lipids (PC 34:1 and PC 36:2), and if neither of these were present then the sample was considered a failure and excluded from subsequent analysis. It is also important to identify and exclude samples where two cells had been erroneously dispensed into the same well. We did this by looking for samples that had significantly higher abundance of all the lipids measured (as an analysis of two cells will increase the amount of all lipids in the well relative to a single cell), and we excluded samples from further analysis if 80% of lipids had an abundance 2.5 standard deviations above the mean.

As in traditional lipidomic approaches we used signal to noise, mass accuracy (ppm), and missing value thresholds to identify lipid signals, although we applied more relaxed cutoffs for each of these parameters than are used in normal lipidomic methods (Harshfield et al., 2019). The lower signal abundance achieved

|         | Mutant                |       | Wild-Type             |       |
|---------|-----------------------|-------|-----------------------|-------|
|         | p Value               | FC    | p Value               | FC    |
| PC 34:1 | 0.0001                | 0.679 | $1.28 \times 10^{-7}$ | 0.628 |
| PC 34:2 |                       |       | $1.13 \times 10^{-7}$ | 0.578 |
| PC 36:1 |                       |       | 0.0002                | 0.591 |
| PC 36:4 | $4.76 \times 10^{-6}$ | 0.610 | $2.16 \times 10^{-7}$ | 0.641 |
| PC 38:4 | 0.0009                | 2.32  | 0.038                 | 1.61  |
| PC 40:3 |                       |       | 0.0041                | 3.29  |
| PE 36:1 | 0.0061                | 4.31  |                       |       |
| PE 36:2 | 0.047                 | 1.79  | 0.037                 | 1.81  |
| PE 36:3 | 0.0023                | 3.96  | 0.013                 | 3.17  |
| SM 36:1 | 0.0047                | 6.29  | 0.0005                | 3.64  |
| PC/PE   | 0.011                 | 0.36  |                       |       |

**Table 2. Lipids with Significantly Different Abundances in Cells with High and Low Levels of PC(36:2)**

FC, fold change relative to cells with low PC(36:2); PC, glycerophosphocholine; PE, glycerophosphoethanolamine; SM, sphingomyelin.

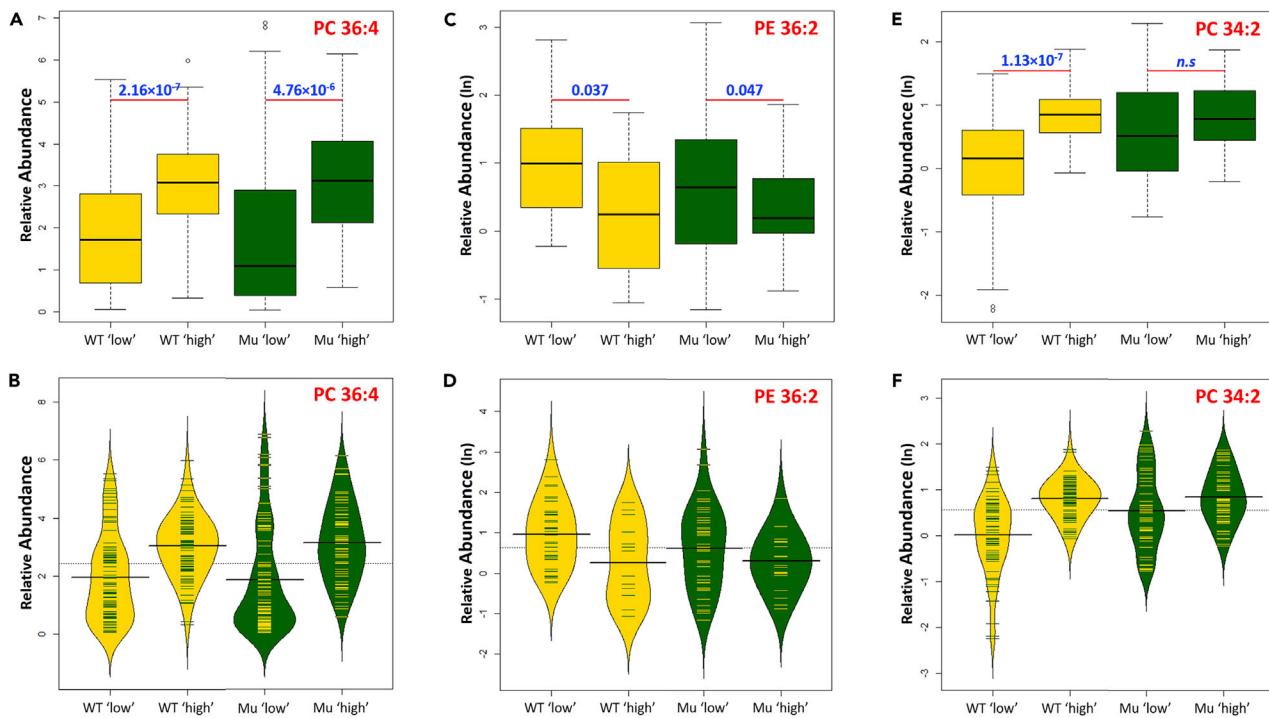
when analyzing a single cell means that signal-to-noise ratios will be lower than when analyzing bulk cell pellets, tissue, or bio-fluid samples. This low intensity does also leads to slightly lower mass accuracy and higher percentage of samples in which a lipid will fall between the lower limit of detection, which means that if we applied the standard values for these threshold real lipid signals would have been missed.

For signal identification, although the QC cutoffs we applied are less stringent than in traditional lipidomics approaches they are more stringent than those used in other single cell lipid profiling methods. For example, whereas we required a signal-to-noise threshold of  $>1.5$ , mass accuracy of  $<10$  ppm,  $r > 0.5$  in a cell “standard” curve, and a signal to be present in at least 10% of analyzed cells with many of the identifications validated across our multiple experiments, Neumann et al. (2019) did not apply any signal to noise, mass accuracy, or cell “standard” curve cutoffs and included lipids measured in less than 1% of cells.

After optimization and quality control we measured 37 lipids (Table 1), which is lower than reported in this previous study; however, this is a result of our quality controls being more stringent to allow meaningful comparison of the generated data. Although the number of lipids reported here is lower than for a traditional lipidomics study (using tissue or plasma, Ellis et al., 2012; Neumann et al., 2019; Thiele et al., 2019), it provides sufficient coverage to interrogate the lipid metabolism of a single cell. When we look at the abundances of lipids measured from a bulk cell sample (Figure S5) it can be seen that the distribution is not linear, and unsurprisingly it can also be seen that the lipids measured in a single cell are generally the lipids that are the most abundant in the bulk samples (Figure S5). Potentially of more interest is the fact that just below the lower limit of detection (LLOD) in the single cell method there are a lot of lipids within a narrow range of abundances, suggesting that a small increase in the sensitivity of the method will likely lead to a large increase in the number of lipids measured. Despite current limitations, developments in mass spectrometric instruments’ sensitivity should soon provide technical improvements to expand this platform to a larger number of detectable lipids.

### Advantages of Single Cell over “Population” Lipid Profiling Approaches

When we compared the “population” and single cell lipid profiles we observed differences between WT and SNCA-A53T mutant dopamine neurons driven predominantly by PC 36:2. However, only by single cell analysis we were able to identify an overlap between the two genotypes both in the multivariate and univariate analyses (Figure 4). This shows that the population level analysis failed to capture the heterogeneity of the biological processes at work in human dopamine neurons. There was also significant



**Figure 5. Boxplots and Beanplots Showing Lipid Abundances in Wild-Type and A53T Mutant Neurons after Stratifying Samples by Their Abundance of PC 36:2**

(A) Boxplot showing the abundance of PC 36:4 in cells with high and low levels of PC 36:2; p values calculated using generalized linear model.  
 (B) Beanplot showing the relative distribution of the abundance of PC 36:4 in cells with high and low levels of PC 36:2.

(C) Boxplot showing the abundance of PE 36:2 in cells with high and low levels of PC 36:2; p values calculated using generalized linear model.  
 (D) Beanplot showing the relative distribution of the abundance of PE 36:2 in cells with high and low levels of PC 36:2.

(E) Boxplot showing the abundance of PC 34:2 in cells with high and low levels of PC 36:2; p values calculated using generalized linear model.  
 (F) Beanplot showing the relative distribution of the abundance of PC 34:2 in cells with high and low levels of PC 36:2.

PC, glycerophosphocholine; Mu, SNCA-A53T mutant dopaminergic neurons.

overlap in the abundance of PC 36:2 (Figure 4D) with the WT cells showing a much more even distribution abundance across their concentration range (Figures 4E and 4F) with a high number of SNCA-A53T neurons presenting with low PC 36:2. These results highlight the need to develop and apply methods for single cell analysis to capture both the heterogeneity of the model in study and important cell-type-specific differences. This is further supported by our recent single cell transcriptomic analysis of this model that leads to the identification and characterization of cellular heterogeneity, which would not be possible using traditional bulk approaches (Fernandes et al., 2020).

### Impaired Membrane Function in SNCA-A53T Dopamine Neurons

PCs are key components of the cell membrane (van Meer et al., 2008), and changes in the balance of species of different chain lengths as well as the PC to PE ratio (van der Veen et al., 2017) affect membrane fluidity, structure, and function (van der Veen et al., 2017; Paris et al., 2010; Sciacca et al., 2012). Several studies have suggested that SNCA plays a role in lipid transport and synaptic membrane biogenesis (Jo et al., 2000), so the SNCA-A53T mutation could result in changes in the composition of the cell membrane and thus potential impairment of synaptic transmission. The results of the “population” lipid profiling show higher levels of PC 36:2 in all WT replicates with no overlap between the genotypes (Figure 4B). This suggests that the A53T mutation is causing membrane remodeling to occur evenly across all cells, potentially resulting in a distinct cellular morphology not present in the WT neurons.

After stratifying lipid abundances based on PC 36:2 11 lipid variables were significantly different between cells with high and low abundance of this lipid. Of these, six were common to both genotypes, and with four

lipids differing significantly between WT and A53T mutant dopamine neurons multiple factors appear to be controlling the lipid composition of a cell. Therefore, single cell approaches provide an opportunity to explore how the A53T mutation interacts with these factors to modulate lipid metabolism. The ratio of total PC to total PE in the cell membrane has been shown to have major effects on the fluidity of the bilayer and vesicle formation as well as the composition and prevalence of transmembrane proteins (Li et al., 2006; Brown and Seelig 1978). Thus it is interesting that in SNCA-A53T dopamine neurons we see a larger difference in the PC:PE ratio between cells with high and low PC 36:2 levels than is observed in the WT (Table 2). The majority of the mutant cells have low PC 36:2, resulting in low PC:PE ratios for a higher proportion of these cells.

Healthy cells have a range of mechanisms to respond to normal physiological stimuli as well as to adapt to stress. In this study we have shown a greater heterogeneity in lipid composition in WT neurons, and given that the cell membrane is the largest source of lipids in the cell it is likely that this reflects heterogeneity in the makeup of the cell membrane. It is plausible that this heterogeneity in membrane composition is integral to the mechanisms by which cell populations respond to membrane stress and the skewed distribution in the mutant neurons could impair their ability to respond to stimuli (Fu et al., 2011; Hannun and Luberto, 2000).

### Strengths of Study

#### High Throughput

Single cell approaches require the analysis of large numbers of cells to capture inter-cell heterogeneity. This results in the requirement of a large number of cells from each biological replicate, significantly increasing the number of analysis required per experiment. Our method is capable of measuring the lipid profile of 280 samples in 24 h meaning that in just 2 days we were able to analyze sufficient samples to identify different distributions of lipid abundance between WT and SNCA-A53T dopamine neurons.

#### Cell-Type Specificity

Tissue and cell culture models are heterogeneous and made up of diverse cell types in different proportions, which can make single cell analysis difficult. Previous lipidomics methods have coupled with imaging approaches to identify the type of cell being analyzed. Although this approach is useful for some applications, it is not translated into a high-throughput platform capable of analyzing large number of samples required to address questions of cellular heterogeneity. Our approach of sorting cells using FACS before analysis ensures that we are specifically analyzing a single and specific cell type of interest. This means that differences detected between generated lipid profiles are the result of heterogeneity within a specific cell population rather than differences between the compositions of different cell types.

#### Quality Control

The need to analyze large numbers of samples in single cell approaches means that robust quality control procedures need to be implemented to ensure that all the generated data is comparable. Although some studies have measured lipids in single cells (Ellis et al., 2012; Neumann et al., 2019; Thiele et al., 2019; Standke et al., 2019) and some have measured in large numbers of cells (Neumann et al., 2019), these approaches have incorporated limited quality control procedures to ensure data consistency across sample sets. In this study we also included a pooled quality control corresponding to unsorted cells, which was added to each plate and enabled us to correct for inter-plate variability (Figure 3). This enabled us to perform large-scale experiments while accounting for possible sources of analytical variance that could confound our results.

#### Validation of Signal Identification

Signal identification in single cell lipid profile data is challenging because the majority of signals are low abundance and are often not measured in a large proportion of samples making it difficult to distinguish real signals from analytical noise. To avoid reporting false-positives in this study we analyzed TH-positive dopamine neurons from three independent experiments to identify lipids that were measured consistently. To provide clarity we have defined three levels of confidence in our putative assignments with species identified in all three experiments given our highest confidence (Level 1) and signals only putatively identified in one experiment given our lowest confidence (Level 3).

In conclusion, this work provides a method for high-throughput lipid profiling at the single cell level. As a proof of principle, we used this platform to profile individual human dopamine neurons resulting in the identification of lipid alterations that will further our understanding of PD pathology. This technological advance potentially has major applications across a wide range of biomedical research fields, allowing the opportunity to explore and overcome metabolic heterogeneity in complex cellular populations.

### Limitations of Study

#### Putative Lipid Assignment

Ideally glycerophospholipids would be identified as their lipid head group and the specific pair of fatty acid side chains rather than by the head group and the total number of carbons and double bonds, e.g., PC 16:1/18:0 rather than PC 34:1. This additional structural detail can improve biological interpretation as the different fatty acid configurations can have differing biological functions. This additional structural detail is usually obtained by fragmenting the parent molecule. However, it is extremely difficult to obtain clean fragmentation spectra from low-abundance signals, making it impossible to obtain this detailed structural information in current single cell approaches. It is for this reason that other single cell approaches have also only been able to annotate lipids to the level of the total number of carbons and double bonds. However, in the future we hope that the availability of more sensitive mass spectrometers will allow us to generate fragmentation spectra and obtain these more detailed annotations.

#### Coverage of the Lipidome

Lipids represent the largest and most diverse group of metabolites in the human body consisting of dozens of structural classes and potentially thousands of structurally distinct molecules. In this study we have identified 81 lipid species from six lipid classes, which is only a small proportion of the total lipids present. However, if we look at the distribution of the abundance of the identified lipids in a lipid dataset generated from a bulk cell pellet it can be seen ([Figure S5](#)) that the assigned lipids are the most abundant lipids. It can clearly be seen that the distribution of lipid abundances is not linear, so even a small improvement in the lower limit of detection can potentially greatly improve the coverage of the lipidome. This improvement can potentially be achieved by using more sensitive mass spectrometry equipment in the future.

#### Exclusion of Internal Standards

Traditional lipidomics methods include internal standards to correct for analytical variation within the data. However, internal standards only work when the amount of lipid in a sample is known, as analytical variation will then be the major source of stochastic variance. If we consider single cell approaches, each sample will consist of one cell, yet the amount of total lipid will differ between samples due to differences in cell size. This means that the intensity of the lipid signals will be independent of the intensity of the internal standards. To support this, we showed that lipid and internal standard signal correlated in plasma samples but did not in single cell samples ([Figure S6](#)).

#### Missing Data

In an ideal scenario all the lipids detected would be measured in all the samples analyzed; however, this is never the case as across samples some lipids will fall below the lower limit of detection. Large numbers of missing values can be problematic in both multivariate and univariate statistical analysis by potentially obscuring the relationship of lipid abundances. This is particularly an issue in single cell approaches where low and medium abundance lipids will only be measured in larger cells. Improving the sensitivity of the platform and lowering the LLOD will increase the proportion of samples that we can measure of a given lipid in helping to alleviate this problem.

### Resource Availability

#### Lead Contact

Further information and requests for resources and reagents should be directed to the lead contact Albert Koulman ([ak675@medsch.cam.ac.uk](mailto:ak675@medsch.cam.ac.uk)).

#### Material Availability

This study did not generate any unique reagents.

### Data and Code Availability

All the code used in this study is shown in the [Supplemental Information](#) of the article.

### METHODS

All methods can be found in the accompanying [Transparent Methods supplemental file](#).

### SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at <https://doi.org/10.1016/j.isci.2020.101703>.

### ACKNOWLEDGMENTS

This work was supported by the Michael J. Fox Foundation (USA) grant ID 16486. S.G.S. was also supported by the Biotechnology and Biological Sciences Research Council (BB/P028195/1) and A.K. by the National Institute of Health Research Cambridge Biomedical Research Centre (146281). E.M. is funded by UK Dementia Research Institute grant ID RRZA/175. H.J.R.F. is funded by Open Targets (OTAR035). We would also like to thank Chiara Cossetti and Dr. Reiner Schulte at the Cambridge Institute for Medical Research Flow Cytometry core facility for their help with cell sorting.

### AUTHORS CONTRIBUTIONS

Conceptualization, H.J.R.F., S.G.S., E.M., and A.K.; Methodology, S.G.S., H.F., and P.T.; Investigation, S.G.S., H.J.R.F., J.K., S.F., and S.F.F.; Writing – Original Draft, S.G.S. and H.J.R.F.; Writing – Review and Editing – S.G.S., H.J.R.F., J.K., E.M., and A.K.; Funding Acquisition, H.J.R.F. and E.M.; Supervision, E.M. and A.K.

### DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: July 1, 2020

Revised: September 10, 2020

Accepted: October 15, 2020

Published: November 20, 2020

### REFERENCES

- Brown, M.F., and Seelig, J. (1978). Influence of cholesterol on the polar region of phosphatidylcholine and phosphatidylethanolamine bilayers. *Biochem* 17, 381–384.
- Bruntraeger, Michaela, Byrne, Meg, Long, Kathleen, and Bassett, Andrew (2019). Editing the Genome of Human Induced Pluripotent Stem Cells Using CRISPR/Cas9 Ribonucleoprotein Complexes. *Methods in molecular biology* 1961, 153–183, [https://doi.org/10.1007/978-1-4939-9170-9\\_11](https://doi.org/10.1007/978-1-4939-9170-9_11).
- Deng, Q., Ramsköld, D., Reinarius, B., and Sandberg, R. (2014). Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells. *Science* 343, 193–196.
- Do, C.B., Tung, J.Y., Dorfman, E., Kiefer, A.K., Drabant, E.M., Francke, U., Mountain, J.I., Goldman, S.M., Tanner, C.M., Langston, W., et al. (2011). Web-based genome-wide association study identifies two novel loci and a substantial genetic component for Parkinson's disease. *PLoS Genet.* 7, e1002141.
- Ellis, S.R., Ferris, C.J., Gilmore, K.J., Mitchell, T.W., Blanksby, S.J., and Panhuis, M. (2012). Direct lipid profiling of single cells from inkjet printed microarrays. *Anal. Chem.* 84, 9679–9683.
- Evers, T.M.J., Hochane, M., Tans, S.J., Heeren, R.M.A., Semrau, S., Nemes, P., and Mashaghi, A. (2019). Deciphering metabolic heterogeneity by single-cell analysis. *Anal. Chem.* 91, 13314–13323.
- Fabelo, N., Martin, V., Santpere, G., Marin, R., Torrent, L., Ferrer, I., and Diaz, M. (2011). Severe alterations in lipid composition of frontal cortex lipid rafts from Parkinson's disease and incidental Parkinson's disease. *Mol. Med.* 17, 1107–1118.
- Fernandes, H.J.R., Patikas, N., et al. (2020). Single cell transcriptomics of Parkinson's Disease human *in vitro* models reveals dopamine neuron specific stress responses. *Cell Rep.* 33, 108263.
- Fu, S., Yang, L., Hofmann, O., Dicker, L., Hide, W., Lin, X., Watkins, S.M., Ivanov, A.R., and Hotamisligil, G.S. (2011). Aberrant lipid metabolism disrupts calcium homeostasis causing liver endoplasmic reticulum stress in obesity. *Nature* 473, 528–531.
- Golovko, M.Y., Rosenberger, T.A., Feddersen, S., Faergeman, N.J., and Murphy, E.J. (2007).  $\alpha$ -synuclein gene ablation increases docosahexanoic acid incorporation and turnover in brain phospholipids. *J. Neurochem.* 101, 201–211.
- Hannun, Y.A., and Luberto, C. (2000). Ceramide in the eukaryotic stress response. *Trends Cell Biol.* 10, 73–80.
- Harshfield, E.L., Koulman, A., Ziemek, D., Marney, L., Fauman, E.B., Paul, D.S., Stacey, D., Rasheed, A., Lee, J., Shah, N., et al. (2019). An unbiased lipid phenotyping approach to study the genetic determinants of lipids and their association with coronary heart disease risk factors. *J. Proteome Res.* 18, 2397–2410.
- Heintzman, N.D., Hon, G.C., Hawkins, D., Kheradpour, P., Stark, A., Harp, L.F., Ye, Z., Lee, L.K., Stuart, R.H., Ching, C.W., et al. (2009). Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* 459, 108–112.
- Jo, E., McLaurin, J., Yip, C.M., George-Hyslop, P., and Fraser, P.E. (2000).  $\alpha$ -synuclein membrane interactions and lipid specificity. *J. Biol. Chem.* 275, 34328–34334.
- Kriks, S., Shim, J., Piao, J., Ganat, Y.M., Wakeman, D.R., Xie, Z., Carillo-Reid, L., Auyeung, G., Antonacci, C., Buch, A., et al. (2011). Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature* 480, 547–551.

Li, Z., Agellon, L.B., Allen, T.M., Umeda, M., Jewell, L., Mason, A., and Vance, D.E. (2006). The ration of phosphatidylcholine to phosphatidylethanolamine influences membrane integrity and statoregulation. *Cell Metab.* 3, 321–331.

La Manno, G., Gyllborg, D., Codeluppi, S., Nishimura, K., Salto, C., Zeisel, A., Borm, L.E., Stott, S.R.W., Toledo, E.M., Villaescusa, C., et al. (2016). Molecular diversity of midbrain development in mouse, human, and stem cells. *Cell* 167, 566–580.

van Meer, G., Voelker, D.R., and Feigenson, G.W. (2008). Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* 9, 112–124.

Neumann, E.K., Ellis, J.F., Triplett, A.E., Rubakhin, S.S., and Sweedler, J.V. (2019). Lipid analysis of 30000 individual Rodent cerebellar cells using high-resolution mass spectrometry. *Anal. Chem.* 91, 7871–7878.

Paris, L., Cecchetti, S., Spadaro, F., Abalsamo, L., Lugini, L., Pisani, M.E., Iorio, E., Natali, P.G., Ramoni, C., and Podo, F. (2010). Inhibition of phosphatidylcholine-specific phospholipase C downregulates HER2 overexpression on plasma membrane of Breast cancer cells. *Breast Cancer Res.* 12, R27.

Polymeropoulos, M.H., Lavedan, C., Leroy, E., Ide, S.E., Dehejia, A., Dutra, A., Pike, B., Root, H., Rubenstein, J., Boyer, R., et al. (1997). Mutation in the alpha-synuclein gene identified in families with Parkinson's disease. *Science* 276, 2045–2047.

Sciacca, M.F.M., Breider, J.R., Lee, D.K., and Ramamoorthy, A. (2012). Phosphatidylethanolamine enhances amyloid

fiber-dependent membrane fragmentation. *Biochem* 51, 7676–7684.

Sharon, R., Bar-Joseph, I., Mirick, G.E., Serhan, C.N., and Selkoe, D.J. (2003). Altered fatty acid composition of dopaminergic neurons expressing  $\alpha$ -synuclein and human brains with  $\alpha$ -synucleinopathies. *J. Biol. Chem.* 278, 49874–49881.

Siddiqi, F.H., Menzies, F.M., Lopez, A., Stamatakou, E., Karabiyik, C., Ureshino, R., Ricketts, T., Jimenez-Sánchez, M., Esteban, M.A., Lai, L., et al. (2019). Felodipine induces autophagy in mouse brains with pharmacokinetics amenable to repurposing. *Nat. Commun.* 10, 1817.

Spillantini, M.G., Schmidt, M.L., Lee, V.M., Trojanowski, J.Q., Jakes, R., and Goedert, M. (1997). Alpha-synuclein in Lewy bodies. *Nature* 388, 839–840.

Standke, S.J., Colby, D.H., Bensen, R.C., Burgett, A.W.G., and Yang, Zhibo. (2019). Mass spectrometry measurement of single suspended cells using combined cell manipulation system and a single-probe device. *Anal. Chem.* 91, 1738–1742.

Thiele, C., Wunderling, K., and Leyendecker, P. (2019). Multiplexed and single cell tracing of lipid metabolism. *Nat. Meth.* 16, 112–1130.

Tiklova, K., Bjorklund, A.K., Lahti, L., Fiorenzano, A., Nolbrant, S., Gillberg, L., Vloakakis, N., Yokota, C., Hilscher, M.M., Hauling, T., et al. (2019). Single -cell RNA sequencing reveals midbrain dopamine neuron diversity emerging during mouse brain development. *Nat. Commun.* 10, 581.

Tirosh, I., Izar, B., Prakadan, S.M., and Wadsworth, M.H. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science* 352, 189–196.

Tracey, T.J., Steyn, F.J., Wolvetang, E.J., and Ngo, S.T. (2018). Neuronal lipid metabolism: multiple pathways driving functional outcomes in health and disease. *Front. Mol. Neurosci.* 11, 1–10.

van der Veen, J., Kennelly, J.P., Wan, S., Vance, J.E., Vance, D.E., and Jacobs, R.L. (2017). The critical role of phosphatidylcholine and phosphatidylethanolamine metabolism in health and disease. *Biochim. Biophys. Acta* 1859, 1558–1572.

Volpatto, Viola, and Webber, Caleb (2020). Addressing variability in iPSC-derived models of human disease: guidelines to promote reproducibility. *Disease Models & Mechanisms* 13, 1–9, <https://doi.org/10.1242/dmm.042317>.

Xia, N., Fang, F., Zhang, P., Cui, J., Tep-Cullison, C., Hamerley, T., Lee, H.J., Palmer, T., Bothner, B., Lee, J.H., and Pera, R.P. (2017). A knockin reporter allows purification and characterization of mDA neurons from heterogeneous populations. *Cell Rep.* 18, 2533–2546.

Zambon, F., Cherubini, M., Fernandes, H.J.R., Lang, C., Ryan, B.J., Volpatto, V., Bengoa-Vergniory, N., Vingil, S., Attar, M., Booth, H.D.E., et al. (2019). Cellular alpha-synuclein pathology is associated with bioenergetic dysfunction in Parkinson's iPSC-derived dopamine neurons. *Hum. Mol. Genet.* 28, 2001–2013.

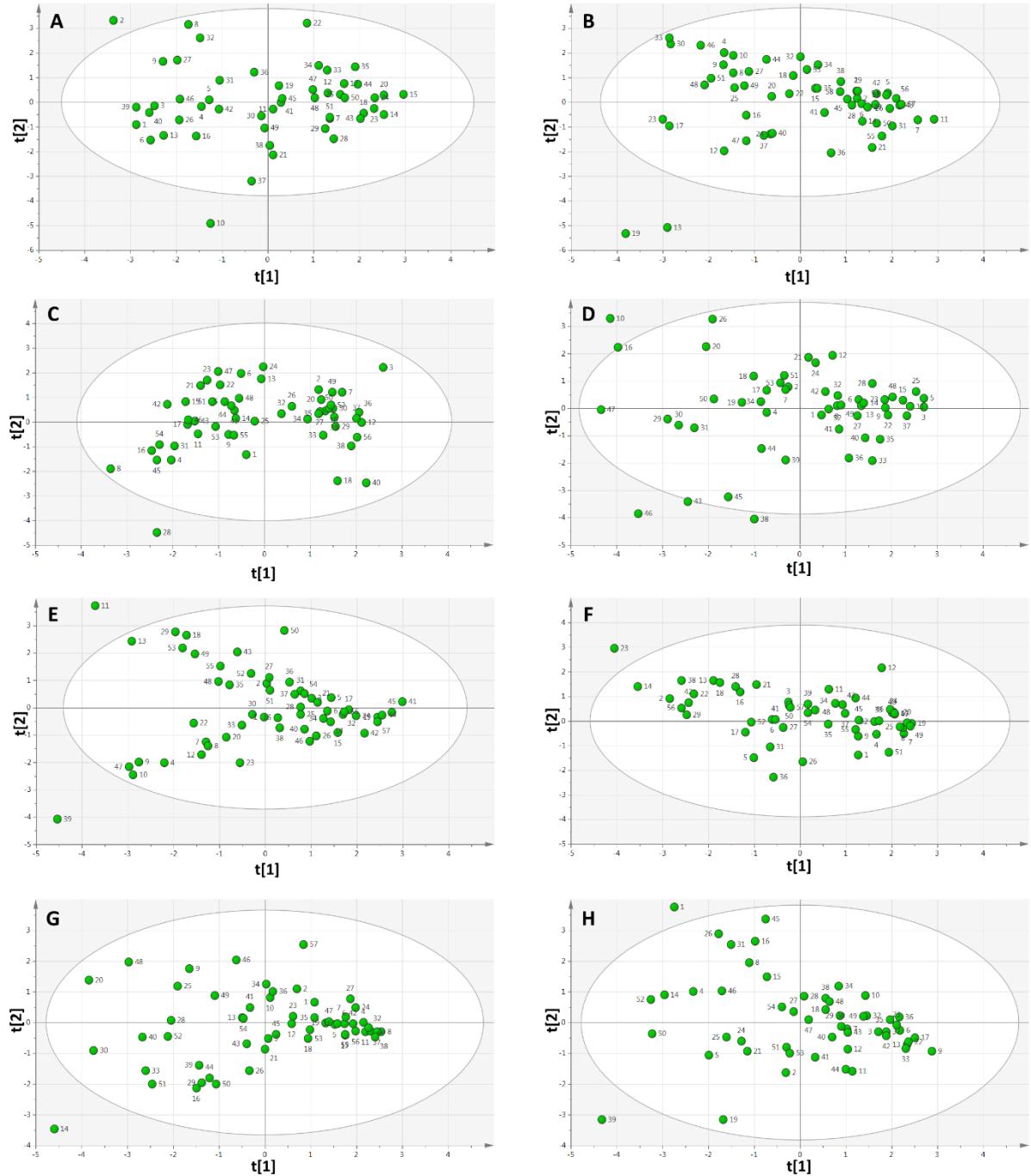
## **Supplemental Information**

### **Development and Application of High-Throughput Single Cell Lipid Profiling: A Study of SNCA-A53T Human Dopamine Neurons**

**Stuart G. Snowden, Hugo J.R. Fernandes, Josh Kent, Stefanie Foskolou, Peri Tate, Sarah F. Field, Emmanouil Metzakopian, and Albert Koulman**

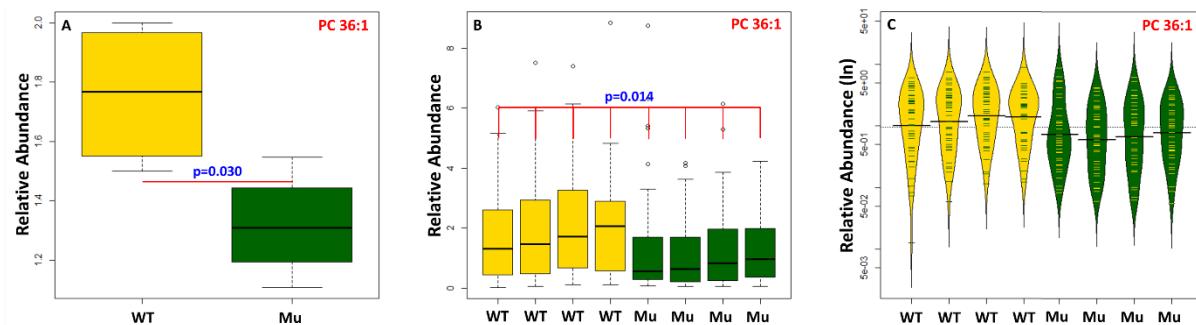
## Supplemental Figures

**Supplemental Figure 1 PCA Scores plots showing the compositional relationship between cells from the same population and their relative injection order. Related to Figure 3.**



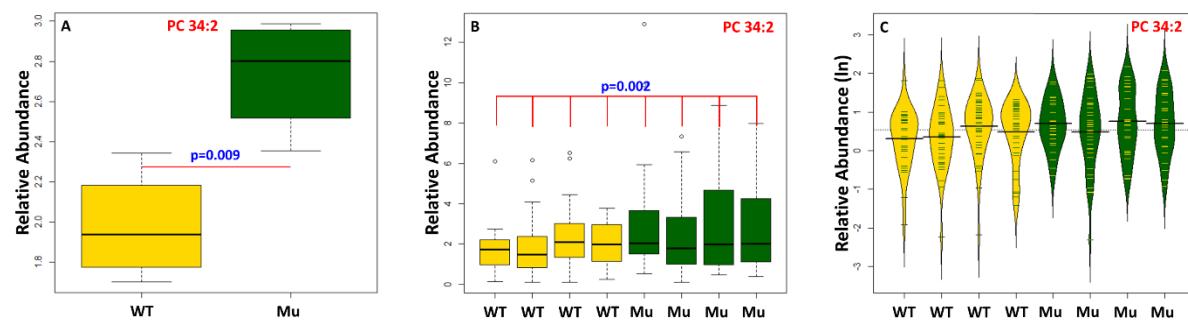
A) Wild type 1 B) Wild type 2 C) Wild type 3 D) Wild type 4 E) SNCA-A53T 1 F) SNCA-A53T 2 G) SNCA-A53T 3 H) SNCA-A53T 4

**Supplemental Figure 2 Plots comparing the abundance of PC 36:1 in wild type and A53T mutant cells. Related to Figure 4.**



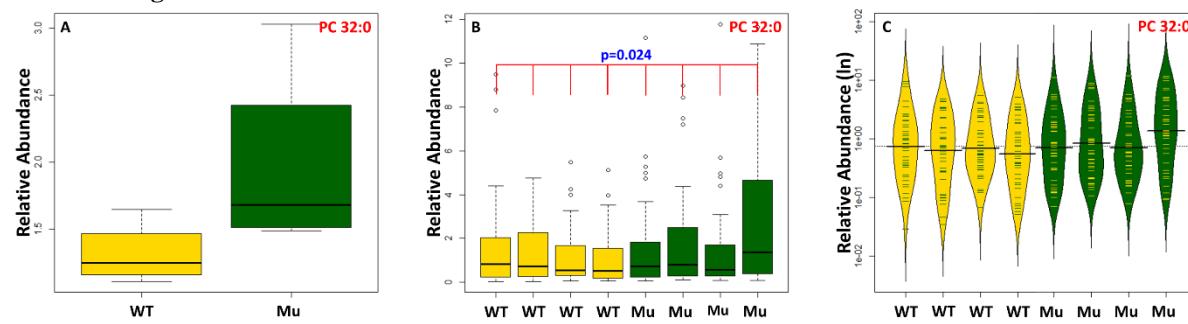
A) Boxplot comparing the abundance of wild type and mutant cells (Mu) in ‘population’ level data B) Boxplot comparing the abundance in individual wild type and mutant cells (Mu) populations using single cell data, p-value calculated using generalised linear models C) Beanplot comparing the distribution of abundances in individual wild type and mutant cell (Mu) populations using single cell data. Mu; SNCA-A53T mutant dopaminergic neurons.

**Supplemental Figure 3 Plots comparing the abundance of PC 34:2 in wild type and A53T mutant cells. Related to Figure 4.**



A) Boxplot comparing the abundance of wild type and mutant cells (Mu) in ‘population’ level data B) Boxplot comparing the abundance in individual wild type and mutant cell (Mu) populations using single cell data, p-value calculated using generalised linear models C) Beanplot comparing the distribution of abundances in individual wild type and mutant cell (Mu) populations using single cell data. Mu; SNCA-A53T mutant dopaminergic neurons.

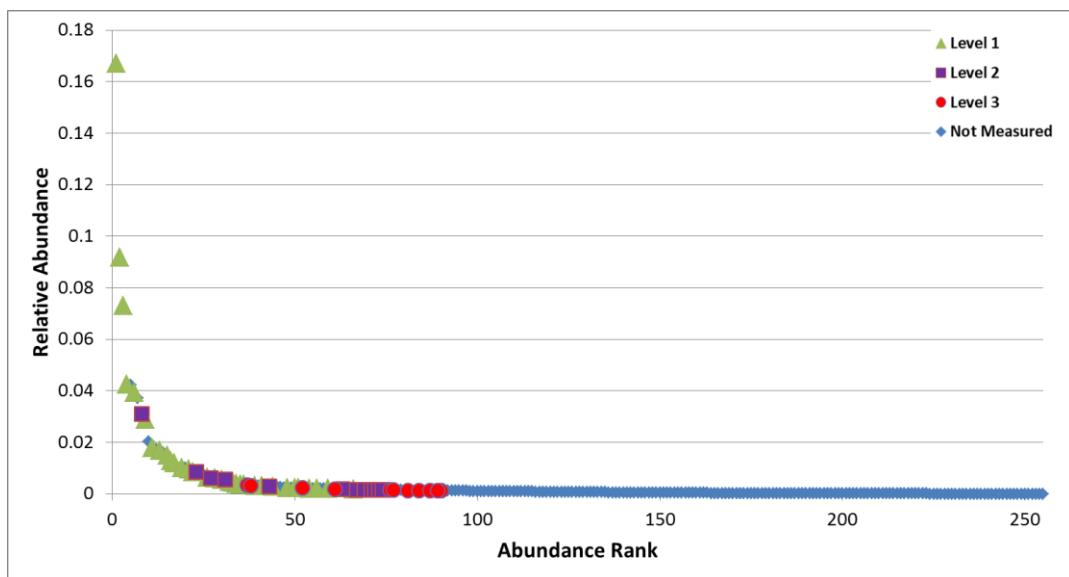
**Supplemental Figure 4 Plots comparing the abundance of PC 32:0 in wild type and A53T mutant cells. Related to Figure 4.**



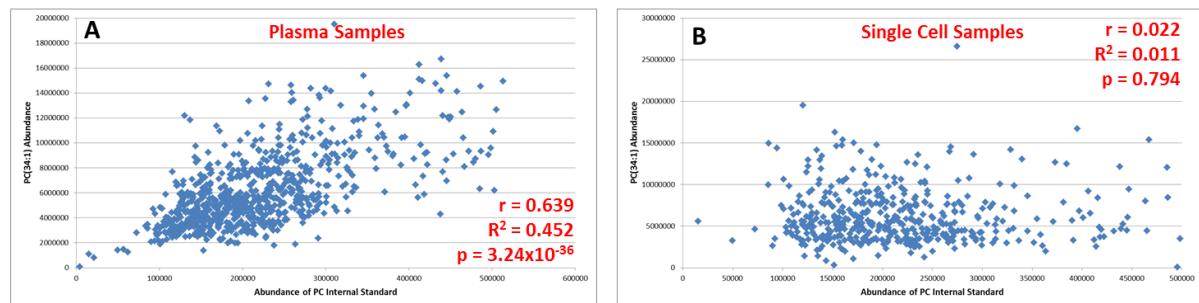
A) Boxplot comparing the abundance of wild type and mutant cells (Mu) in ‘population’ level data B) Boxplot comparing the abundance in individual wild type and mutant cell (Mu) populations using single cell data, p-value calculated using generalised linear models C) Beanplot comparing the distribution of abundances in

individual wild type and mutant cell (Mu) populations using single cell data. Mu; *SNCA-A53T* mutant dopaminergic neurons.

**Supplemental Figure 5** Scatter plot showing the relative abundance of level 1,2 and 3 single cell annotations in a bulk lipidomics dataset generated from the same cell lines. Related to Figure 3.

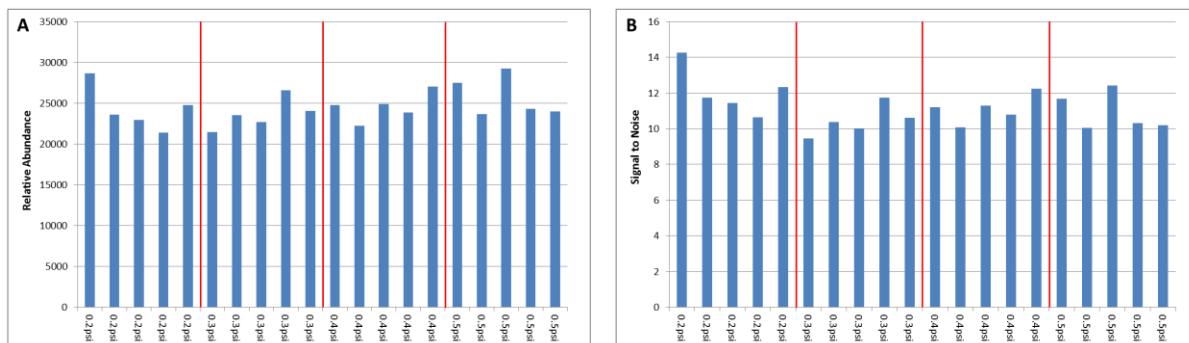


**Supplemental Figure 6** Scatter plots showing the relationship between the abundance of PC internal standard and the abundance of PC 34:1 in both plasma and single cell samples. Related to Figure 3.



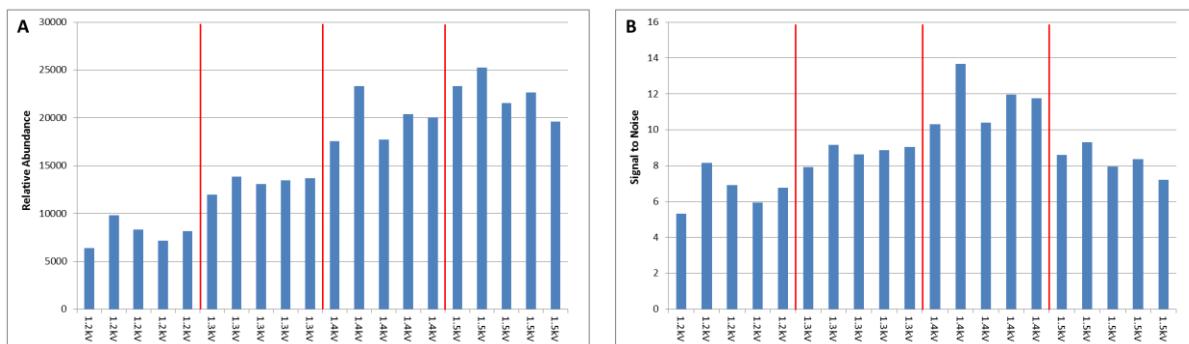
p-values calculated using generalised linear models. PC; glycerophosphocholine.

**Supplemental Figure 7 Plots of the signal abundance and signal to noise ratio of PC C16:0-d31/C18:1 when optimising infusion back pressure. Related to Figure 3.**



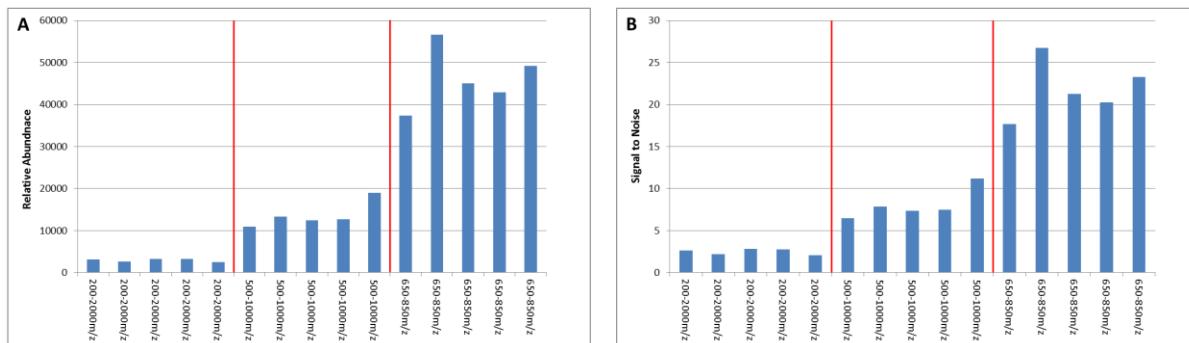
A) signal abundance in individual test samples B) signal to noise in individual test samples

**Supplemental Figure 8** Plots of the signal abundance and signal to noise ratio of PC C16:0-d31/C18:1 when optimising electrospray ionisation current. Related to Figure 3.



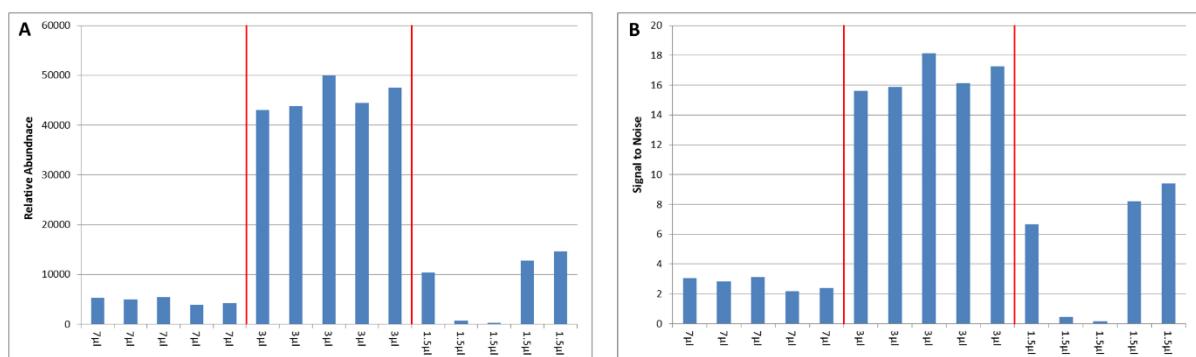
A) signal abundance in individual test samples B) signal to noise in individual test samples

**Supplemental Figure 9 Plots of the signal abundance and signal to noise ratio of PC C16:0-d31/C18:1 when optimising the mass window analysed. Related to Figure 3.**



A) signal abundance in individual test samples B) signal to noise in individual test samples

**Supplemental Figure 10** Plots of the signal abundance and signal to noise ratio of PC C16:0-d31/C18:1 when optimising the solvent volume used for LESA. Related to Figure 3.



A) signal abundance in individual test samples B) signal to noise in individual test samples

## Supplemental scheme

**Supplemental scheme 1 R script used to run our in-house pipeline for comparing generated spectra to list of known lipids. Related to Table 1.**

```
source("hrms.R")
files = list.files(".", pattern=".mzXML")
system.time(
  for (i in 1:length(files)) {
    main(files[i],rtwin=c(20,78),mzwin=c(650,850))
  }
)
results <- signals_deviations()
```

**Supplemental Scheme 2 R script contained in “hrms.R” referred to in supplemental scheme 1 that compares generated spectra to list of known lipids listed in Supplemental table 3. Related to Table 1.**

```
main <- function(filename,rtwin,mzwin) {
  require(xcms)
  require(data.table)
  targets <- read.table("./LipidList.csv", header=T, sep=',')
  targets <- data.table(targets)
  options("nwarnings" = (length(targets$mz)+50)) # we need to get at least as many warnings as targets
  spectrum <- getspectra(filename=filename, rt=rtwin, mz=mzwin)
  tgts <- peaktable(targets,spectrum)
  write.csv(tgts, file = gsub(pattern=".mzXML", x=filename, replacement=".csv"), row.names=F)
}

signals_deviations <- function() { # the csv files must be in the active directory
  x <- read.csv(gsub(pattern=".mzXML", x=files[1], replacement=".csv"),header=T)
  targets <- read.table("./LipidList.csv", header=T, sep=',')
  targets <- data.table(targets)
  signals <- data.frame(x$targets.name,targets$mz)
  for (i in 1:length(files)) {
    x <- read.csv(gsub(pattern=".mzXML", x=files[i], replacement=".csv"),header=T)
    signals[files[i]] <- data.frame(x$signal)
  }

  x <- read.csv(gsub(pattern=".mzXML", x=files[1], replacement=".csv"),header=T)
  deviations <- data.frame(x$targets.name,targets$mz)
  for (i in 1:length(files)) {
    x <- read.csv(gsub(pattern=".mzXML", x=files[i], replacement=".csv"),header=T)
    deviations[files[i]] <- data.frame(x$mz_deviations)
  }

  results <- list(signals,deviations)
  write.csv(results[[1]],file="signals.csv", row.names=F)
  write.csv(results[[2]],file="deviations.csv", row.names=F)
  return(results)
}

getspectra <- function(filename,rt,mz) {
  spectra <- list()
```

```

spectrum <- getSpec(xcmsRaw(filename), rrange=rt, mzrange=mz)
spectrum[,"mz"] <- round(spectrum[,"mz"], digits=4)
spectrum <- as.data.table(spectrum)
spectrum <- spectrum[,mean(intensity), by=mz]
spectrum <- na.omit(spectrum)
setkey(spectrum,mz)
return(spectrum)
}

peakable <- function(targets,spectra) {
  nearest_mz <- vector(length=length(targets$mz)) #predefine length later
  signal <- vector(length=length(targets$mz)) #predefine length later
  for (i in 1:length(targets$mz)) {
    target <- targets[i,mz]
    peak <- peakfind_midpoint(target,spectra,0.01,warnings)
    nearest_mz[i]<-peak[1,mz]
    signal[i]<-peak[1,V1]
  }

  mz_deviation <- targets[,mz] - nearest_mz
  peak_id <- data.frame(targets$name,targets$mz,nearest_mz,mz_deviation,signal)
  return(peak_id)
}

peakfind_midpoint <- function(target,spectra,hwidth,warnings) {
  window <- subset(spectra, spectra$mz > target-hwidth & spectra$mz < target+hwidth)
  if (nrow(window)==0) { # no data for target?
    peak = data.table('mz'=target,'V1'=0) # enter zero intensity for target mass
  } else if (sum(window$V1) < 5000) { # very low s/n?
    peak <- peakfind_max(target,spectra,hwidth) # uses older peakmax finder for low s/n peaks, while
    less accurate this helps with exception handling dramatically
    warning(paste("low signal/noise found for target mass-", target, "-using older peakmax finder.
Identification may not be accurate", sep=" "))
  } else { # now we run the peak width peak finder
    peak <- peakfind_max(target,spectra,hwidth)
    hh_close=peak[1,V1]/2
    window <- subset(spectra, spectra$mz > peak$mz-hwidth & spectra$mz < peak$mz+hwidth)
    if (window$V1[length(window$mz)] > hh_close | window$V1[1] > hh_close) { # window doesn't
sample the width of the peak?
      setkey(window,V1) #this will sort table by intensity, thus finding peak maximum as last entry in
      table
      peak <- window[length(window$mz)] #get last entry of table for the peak maximum
      window <- subset(spectra, spectra$mz > peak$mz-hwidth & spectra$mz < peak$mz+hwidth)
      if (window$V1[length(window$mz)] > hh_close | window$V1[1] > hh_close) { # is the bad
sampling of peak due to interference?
        setkey(window,V1) #this will sort table by intensity, thus finding peak maximum as last entry in
        table
        peak <- window[length(window$mz)] #get last entry of table for the peak maximum
        warning((paste("interfered peak detected for target mass-", target, "-older peak_max() function
used", sep=" ")))
      }
    } else {
      ##for resolved peaks (at half height) the follow code is run ##
      setkey(window,V1) #this will sort table by intensity, thus finding peak maximum as last entry in
      table
      peak <- window[length(window$mz)] #get last entry of table for the peak maximum
    }
  }
}

```

```

hh=peak[1,V1]/2
nearmz=peak[1,mz]
left_mzs <- c(max(subset(window, window$mz < nearmz & window$V1 <
hh)$mz),min(subset(window, window$mz < nearmz & window$V1 > hh)$mz))
left_int <- c(max(subset(window, window$mz < nearmz & window$V1 <
hh)$V1),min(subset(window, window$mz < nearmz & window$V1 > hh)$V1))
right_mzs <- c(max(subset(window, window$mz > nearmz & window$V1 <
hh)$mz),min(subset(window, window$mz > nearmz & window$V1 > hh)$mz))
right_int <- c(max(subset(window, window$mz > nearmz & window$V1 <
hh)$V1),min(subset(window, window$mz > nearmz & window$V1 > hh)$V1))
midpoints <- data.frame(left_mzs, left_int, right_mzs, right_int)
### there has got to be a way to combine the last five rows into one row
coordinates <- list()
left <- coefficients(lm(left_int ~ left_mzs, data=midpoints))
right <- coefficients(lm(right_int ~ right_mzs, data=midpoints))
midpoint <- ((hh-left[1])/left[2]+(hh-right[1])/right[2])/2 # midpoint between the intersection
points of both lines from a y=hh flat line
peak$mz <- round(midpoint[1], digits=4) # modify the peaks variable with the new more accurate
m/z value
}
}
return(peak)
}

peakfind_max <- function(target,spectra,hwidth) {
window <- subset(spectra, spectra$mz > target-hwidth & spectra$mz < target+hwidth)
setkey(window,V1) #this will sort table by intensity, thus finding peak maximum as last entry in table
peak <- window[length(window$mz)] #get last entry of table for the peak maximum
#plot(window, type='h', lwd=1)
return(peak)
}

if(!interactive()){
args <- commandArgs(trailingOnly = TRUE)
f <- args[1]
main(f,rtwin=c(85,130),mzwin=c(190,1800))
}

```

## Supplemental Tables

**Supplemental Table 1** comparison of signal abundance and mass deviation of individual lipids in both extraction blanks and single cell samples. Related to Figure 1.

|                | Abundance |         |                       | Deviation |         |                       |
|----------------|-----------|---------|-----------------------|-----------|---------|-----------------------|
|                | Blanks    | Samples | p-value               | Blanks    | Samples | p-value               |
| <b>PC 30:1</b> | 42        | 803     | $3.1 \times 10^{-4}$  |           |         |                       |
| <b>PC 32:0</b> | 70        | 3117    | $5.1 \times 10^{-10}$ | 8.6       | 2.2     | $3.0 \times 10^{-10}$ |
| <b>PC 32:1</b> | 633       | 7122    | 0.003                 |           |         |                       |
| <b>PC 32:2</b> | 215       | 651     | 0.001                 |           |         |                       |
| <b>PC 32:4</b> | 80        | 232     | 0.01                  |           |         |                       |
| <b>PC 34:0</b> | 5         | 625     | $4.1 \times 10^{-6}$  |           |         |                       |
| <b>PC 34:1</b> | 614       | 10330   | $4.6 \times 10^{-13}$ | 14.7      | 1.8     | $4.6 \times 10^{-13}$ |
| <b>PC 34:2</b> | 97        | 2977    | $7.1 \times 10^{-32}$ | 8.3       | 2.1     | $2.1 \times 10^{-9}$  |
| <b>PC 34:3</b> | 309       | 1445    | $8.7 \times 10^{-12}$ |           |         |                       |
| <b>PC 34:4</b> | 71        | 1575    | $4.5 \times 10^{-34}$ | 6.3       | 8.4     | 0.03                  |
| <b>PC 36:1</b> | 160       | 2062    | $3.8 \times 10^{-7}$  | 11.5      | 2.6     | $9.7 \times 10^{-11}$ |
| <b>PC 36:2</b> | 75        | 2619    | $1.2 \times 10^{-46}$ | 10.8      | 3.0     | $9.9 \times 10^{-15}$ |
| <b>PC 36:3</b> | 124       | 1945    | $9.1 \times 10^{-38}$ | 9.3       | 8.1     | 0.02                  |
| <b>PC 36:4</b> | 400       | 3541    | $1.2 \times 10^{-8}$  | 16.2      | 5.1     | $3.8 \times 10^{-11}$ |
| <b>PC 36:5</b> | 96        | 2295    | 0.07                  | 13.3      | 6.2     | 0.05                  |
| <b>PC 38:3</b> | 41        | 294     | $1.9 \times 10^{-9}$  |           |         |                       |
| <b>PC 38:4</b> | 56        | 179     | $1.7 \times 10^{-6}$  | 7.4       | 7.1     | 0.50                  |
| <b>PC 38:5</b> | 33        | 792     | $2.2 \times 10^{-20}$ |           |         |                       |
| <b>PC 38:8</b> | 3         | 357     | 0.03                  |           |         |                       |
| <b>PC 40:3</b> | 26        | 371     | $4.8 \times 10^{-6}$  | 6.7       | 4.4     | 0.01                  |
| <b>PC 40:4</b> | 328       | 686     | 0.03                  |           |         |                       |
| <b>PC 40:5</b> | 11        | 263     | $4.5 \times 10^{-6}$  |           |         |                       |
| <b>PC 40:6</b> | 826       | 1249    | 0.04                  |           |         |                       |
| <b>PC 40:9</b> | 6         | 322     | 0.02                  |           |         |                       |
| <b>PE 34:0</b> | 64        | 229     | 0.02                  |           |         |                       |
| <b>PE 34:1</b> | 8         | 553     | $4.3 \times 10^{-5}$  | 11.3      | 4.6     | $2.1 \times 10^{-7}$  |
| <b>PE 34:2</b> | 53        | 510     | 0.01                  | 7.2       | 1.9     | $8.4 \times 10^{-7}$  |
| <b>PE 34:3</b> | 67        | 235     | 0.04                  | 8.1       | 5.7     | 0.04                  |
| <b>PE 36:0</b> | 42        | 347     | 0.06                  |           |         |                       |
| <b>PE 36:1</b> | 1030      | 2365    | 0.03                  | 7.4       | 3.4     | $1.2 \times 10^{-5}$  |
| <b>PE 36:3</b> | 39        | 104     | 0.11                  | 8.6       | 6.4     | 0.01                  |
| <b>PE 36:4</b> | 39        | 129     | 0.0002                | 8.5       | 5.4     | 0.003                 |
| <b>PE 38:6</b> | 61        | 154     | 0.006                 |           |         |                       |
| <b>SM 34:1</b> | 66        | 754     | 0.0009                | 7.8       | 1.3     | $3.2 \times 10^{-11}$ |
| <b>SM 36:1</b> | 48        | 323     | $5.6 \times 10^{-6}$  | 7.8       | 1.4     | $1.1 \times 10^{-8}$  |
| <b>SM 36:2</b> | 135       | 294     | 0.04                  |           |         |                       |

PC; glycerophosphocholine, PE; glycerophosphoethanolamine, SM; sphingomyelin.

**Supplemental Table 2 comparison of signal abundance and mass deviation of individual lipids in both extraction blanks and single cell samples. Related to Table 1.**

|                          | <b>Experiment 1</b> | <b>Experiment 2</b> | <b>Experiment 3</b> |
|--------------------------|---------------------|---------------------|---------------------|
| Cer 42:1 ( <b>Lv2</b> )  | 33.3                | 18.8                |                     |
| Cer 44:2 ( <b>Lv2</b> )  |                     | 23.7                | 37.8                |
| DG 42:2 ( <b>Lv2</b> )   | 11.1                | 12.5                |                     |
| PC 30:0 ( <b>Lv2</b> )   |                     | 60.0                | 40.5                |
| PC 34:5 ( <b>Lv3</b> )   |                     | 92.5                |                     |
| PC 36:0 ( <b>Lv2</b> )   | 55.5                |                     | 17.3                |
| PC 38:2 ( <b>Lv2</b> )   |                     | 38.8                | 32.9                |
| PC 38:6 ( <b>Lv2</b> )   | 22.2                |                     | 47.3                |
| PC 38:7 ( <b>Lv3</b> )   |                     |                     | 36.5                |
| PC 40:7 ( <b>Lv3</b> )   | 33.3                |                     |                     |
| PC 40:8 ( <b>Lv3</b> )   | 22.2                |                     |                     |
| PC-O 32:0 ( <b>Lv3</b> ) |                     |                     | 10.8                |
| PC-O 34:1 ( <b>Lv2</b> ) |                     | 30.0                | 17.6                |
| PC-O 34:2 ( <b>Lv2</b> ) | 55.5                |                     | 19.4                |
| PC-O 36:3 ( <b>Lv3</b> ) |                     | 7.3                 |                     |
| PC-O 36:4 ( <b>Lv2</b> ) |                     | 13.8                | 19.4                |
| PC-O 36:5 ( <b>Lv3</b> ) | 22.2                |                     |                     |
| PC-P 34:4 ( <b>Lv3</b> ) | 33.3                |                     |                     |
| PE 32:2 ( <b>Lv2</b> )   |                     | 6.3                 | 11.9                |
| PE 38:2 ( <b>Lv3</b> )   |                     |                     | 26.1                |
| PE 38:3 ( <b>Lv2</b> )   |                     | 11.7                | 20.4                |
| PE 38:4 ( <b>Lv3</b> )   |                     |                     | 4.3                 |
| PE 38:5 ( <b>Lv3</b> )   |                     |                     | 49.8                |
| PE 40:4 ( <b>Lv2</b> )   | 22.2                | 6.3                 |                     |
| PE 40:5 ( <b>Lv3</b> )   |                     |                     | 23.4                |
| PE 40:6 ( <b>Lv3</b> )   |                     |                     | 14.6                |
| PE 40:7 ( <b>Lv2</b> )   | 33.3                |                     | 9.9                 |
| PE-O 34:2 ( <b>Lv3</b> ) |                     |                     | 22.1                |
| PE-O 36:1 ( <b>Lv3</b> ) |                     |                     | 6.8                 |
| PE-O 36:2 ( <b>Lv2</b> ) | 22.2                |                     | 43.2                |
| PE-O 36:3 ( <b>Lv3</b> ) |                     |                     | 13.3                |
| PE-O 36:4 ( <b>Lv2</b> ) |                     | 28.8                | 15.8                |
| PE-O 36:5 ( <b>Lv2</b> ) |                     | 1.3                 | 7.9                 |
| PE-O 38:2 ( <b>Lv2</b> ) | 11.1                | 10.0                |                     |
| PE-O 38:4 ( <b>Lv3</b> ) |                     |                     | 10.6                |
| PE-O 38:5 ( <b>Lv3</b> ) | 55.5                |                     |                     |
| PE-O 38:6 ( <b>Lv3</b> ) |                     |                     | 1.8                 |
| PE-O 40:5 ( <b>Lv3</b> ) |                     |                     | 12.4                |
| PE-O 40:6 ( <b>Lv2</b> ) |                     | 41.3                | 27.3                |
| PE-P 40:7 ( <b>Lv2</b> ) | 11.1                |                     | 49.3                |
| PS 38:5 ( <b>Lv2</b> )   |                     | 23.7                | 51.4                |
| PS 40:4 ( <b>Lv3</b> )   |                     |                     | 35.6                |
| PS 40:6 ( <b>Lv2</b> )   |                     | 20.0                | 27.0                |
| SM 34:0 ( <b>Lv2</b> )   | 11.1                |                     | 11.5                |
| TG 42:4 ( <b>Lv2</b> )   |                     | 18.8                | 13.3                |

Cer; ceramide, DG; diglyceride, PC; glycerophosphocholine, PE; glycerophosphoethanolamine, PS; glycerophosphoserine SM; sphingomyelin, TG; triglyceride.

## Transparent Methods

### Key resources Table

| Reagent or Resource                    | Source                   | Identifier       |
|--|--------------------------|------------------|
| <b>Biological Samples</b>              |                          |                  |
| KOLF-2 iPSC (WT)                       | HipSci                   | RRID:CVCL_9S58   |
| KOLF-2 iPSC ( <i>SNCA-A53T</i> )       | Fernandes et al, 2020    |                  |
|  |                          |                  |
| <b>Chemicals and consumables</b>       |                          |                  |
| Isopropanol                            | Sigma Aldrich            | Cat# I9516-1L    |
| Methanol                               | Thermo Fisher            | Cat# 10675112    |
| Chloroform                             | Sigma Aldrich            | Cat# 366927-2.5L |
| Methyl-tertiary butyl ether            | Sigma Aldrich            | Cat# 34875-2.5L  |
| Ammonium acetate                       | Fluka                    | Cat# 17836-250G  |
| Phosphatidylcholine C16:0-d31:C18:1    | Avanti                   |                  |
| Glass coated 96 well plate             |                          |                  |
| SAG                                    | Enzo Life Sciences       | ALX-270-426-m001 |
| LDN-193189                             | Cambridge Bioscience     | 2092-5           |
| Recombinant Human FGF-8a Protein       | R&D Systems              | 4745-F8-050      |
| StemMACS Purmorphamine                 | Miltenyi Biotec          | 130-104-465      |
| Bovine Serum Albumin 7.5%              | Sigma-Aldrich            | A8412            |
| 3',5'-Dibutyryl Cyclic AMP             | Sigma-Aldrich            | D0627-1G         |
| L-ascorbic acid                        | Sigma-Aldrich            | A4544-25G        |
| DAPT                                   | Tocris                   | 2634/50          |
| SB-431542                              | Tocris                   | 1614/10          |
| Y-27632                                | Tocris                   | 1254/10          |
| CHIR99021                              | Tocris                   | 4423/10          |
| 2-Mercaptoethanol                      | Thermo Fisher Scientific | 31350-010        |
| Antibiotic-Antimycotic (100X)          | Thermo Fisher Scientific | 15240062         |
| L-Glutamine 200mM                      | Thermo Fisher Scientific | 25030081         |
| N-2 Supplement (100X)                  | Thermo Fisher Scientific | 17502048         |
| B-27 Supplement (50X), minus Vitamin A | Thermo Fisher Scientific | 12587-010        |
| Geltrex LDEV-Free hESC-qualified       | Life Technologies        | A1413302         |
| Knockout Replacement Serum             | Life Technologies        | 10828-028        |
| Neurobasal Media                       | Life Technologies        | 21103-049        |
| StemPro Accutase                       | Life Technologies        | A1110501         |
| Dimethyl Sulfoxide                     | Merck                    | D2650-100ML      |
| Recombinant Human BDNF                 | Peprotech                | 450-02           |
| Recombinant Human GDNF                 | Peprotech                | 450-10           |
| Recombinant Human TGF-b3               | Peprotech                | PHG9305          |
| MEM Non-essential amino acids (100X)   | Life Technologies        | 11140-035        |
| Knockout DMEM                          | Life Technologies        | 10829-018        |
| TeSR-E8                                | STEMCELL Technologies    | 05990            |
|  |                          |                  |
| <b>Deposited Data</b>                  |                          |                  |
| Raw and analysed data                  | This paper               |                  |
| Analysis scripts used                  | This paper               |                  |
|  |                          |                  |
| <b>Software</b>                        |                          |                  |
| R                                      | Open source              | version 3.4.2    |

## **Resource availability**

### **Lead contact**

Further information and requests for resources and reagents should be directed to the lead contact Albert Koulman ([ak675@medschl.cam.ac.uk](mailto:ak675@medschl.cam.ac.uk)).

### **Material availability**

This study did not generate any unique reagents.

### **Data and code availability**

All of the code used in this study is shown in the supplemental information of the paper

## **Experimental model and subject details**

### **Cell culture and dopaminergic differentiation**

Human iPSCs were cultured in TeSR-E8 medium on Vitronectin coated plates. Cells were passaged with 0.5mM EDTA when reaching 70% confluence at a ratio of 1:6. Differentiation into dopaminergic neurons was performed according modify version of existing experimental protocols (Siddiqi et al., 2019, Kriks et al., 2011). In brief, iPSCs were first dissociated into single cells, plated at 150,000 cells/cm<sup>2</sup> on Geltrex coated plates and grown for 11 days in Knockout Serum Replacement media (KSR) containing KO DMEM media, KSR (15%), Non Essential Amino Acids (1:100), 2-Mercaptoethanol (10 µM) and 2 mM L-glutamine. KSR medium was gradually changed to NNB medium containing Neurobasal medium, N2 (0.5X) and B27 (0.5X) and 2 mM L-glutamine from day 6. Media was changed to NB medium on day 12 containing Neurobasal medium, B27 (1X) and 2 mM L-glutamine. Medias were supplemented with LDN-193189 (100nM) from days 0-10; SB431542 (10 µM) from days 0-4; SAG (100 nM) from days 1-6; Purmorphamine (2 µM) from days 1-6; FGF8a (100 ng/ml) from days 1-6; and CHIR99021 (3 µM) from days 3-12. From Day 12 onwards, the following supplements were added: BDNF (20ng/ml), GDNF (20ng/ml), Ascorbic Acid (200 µM), TGFβ3 (1 ng/ml), dibutyryl cAMP (500 µM), and DAPT (10 µM). At day 21 cells were dissociated with StemPro Accutase and replated at 300,000 cells/cm<sup>2</sup> in dishes pre-coated with Geltrex and fed every second day for 2 weeks before analysis.

Protocols for the differentiation of *in vitro* human dopamine neurons are known to generate heterogenous neuronal populations (Le Manno et al., 2016). To overcome this, we engineered human wild-type (WT) iPSC with a tyrosine hydroxylase (TH) red fluorescence protein (RFP) (Xia et al., 2017), with modifications. In order to express tagRFP-T under the influence of the endogenous TH promoter we introduced a P2A-tagRFP-T donor plasmid with homologous arms flanking the TH stop codon and an EF1A-Puromycin selection flanked by LoxP sites for drug selection. The targeting plasmid was introduced into iPS cells together with an sgRNA targeting the 3' end of the TH gene and Cas9 RNP by nucleofection to cause a double strand break near the stop codon. After Puromycin selection, transient expression of Cre was used to excise the Puromycin cassette. TH catalyses the hydroxylation

of L-DOPA (the rate limiting step in the synthesis of dopamine) and is considered a *bona fide* marker for dopamine neurons. After differentiation into dopamine neurons (Siddiqi et al., 2019; Kriks et al., 2011), fluorescence-activated cell sorting (FACS) was used to sort cells based on RFP expression. RFP positive dopamine neurons were individually sorted into multi-well plates, followed by LESA mass spectrometry.

### **TH-RFP engineering**

iPSCs were engineered with a TH red fluorescence protein (RFP) reporter following previous methods (Xia et al., 2017) with modifications. We introduced a P2A-tagRFP-T donor plasmid with TH homologous arms and an EF1A-Puromycin selection flanked by LoxP sites for drug selection. The targeting plasmid was introduced into iPS cells together with a TH sgRNA (GACGCCGTGCACCTAGCCAA TGG) against the 3' end of the TH gene and Cas9 (HiFi) RNP by nucleofection (AMAXA 4D nucleofector, “CA 137” pulse code). Nucleofection was performed using the P3 Primary Cell 4D-Nucleofector X Kit L (cat. V4XP-3024) according to the manufacturer’s instructions. Briefly,  $1.5 \times 10^6$  cells were used for each nucleofection reaction together with 2 $\mu$ g of donor plasmid, 16 $\mu$ g of TH sgRNA and 20 $\mu$ g of HiFi Cas9. After Puromycin selection (0.25 $\mu$ g/ml for 10 days), transient expression of Cre (Tat-Cre, 4 $\mu$ M, 3 hours) was used to excise the Puromycin cassette. Clones were picked and genotyped for correct insertion using the following primers: GCCCTCGACCACCTTGATTCTCATGG and CCAAGACCAGACGTACCACTCAGTC; GCCAGGGGCATCTTACAGAGTCTG and GGCGACAAAGAGACCTACG; CAGGAGCTATGCCTCACGCATCCAG and GTGACGGTATTGGGGCAGCAGAC.

### **Flow cytometry**

Following differentiation until day 35, neuronal cultures were washed with PBS and incubated with StemPro Accutase (1X) for 20 min at 37°C. An equal volume of Tripsin-EDTA (0.25%) was added followed by additional incubation at 37°C for 10 min. Neurons were then gently dissociated into a single cell suspension with a P1000 pipette and reaction was neutralized by the addition of an equal volume of Neurobasal media containing Y-27632 (10 $\mu$ M). Cell suspension was then filtered through a 40 $\mu$ m FALCON Cell Strainer (Corning) to remove any undissociated cell clumps, and spun at 300xg for 5 minutes. Cells were resuspended in PBS and sorted based on their expression of RFP using a BD-Influx cell sorter (Becton-Dickinson, San Jose, CA, USA). Buffer used during sorting was BD FACSFlow (cat. 342003, BD Biosciences).

### **Method details**

#### **Liquid extraction surface analysis – mass spectrometry (LESA-MS)**

Liquid extraction surface analysis (LESA) on TH positive dopamine neurons from the bottom of a well of a glass coated 96 well plate using a Triversa Nanomate (Advion, Ithaca, USA). Lipid extraction was achieved using a solvent of 20 mM ammonium acetate in isopropanol, methanol and chloroform (4:2:1, v:v:v), with 3  $\mu$ l aspirated from the solvent reservoir with 1.0  $\mu$ l subsequently dispensed onto the sample area and held there for 7 seconds before being re-aspirated and infused into the mass spectrometer for 90 seconds.

Data was acquired on an Exactive Orbitrap (Thermo, Hemel Hampstead, UK) with data acquired between 650-850  $m/z$  in full scan in the positive ionisation mode. Parameters were optimised to maximise sensitivity (Supplemental Information 1) with the optimal settings being an infusion back pressure of 0.4 psi, 1.4 Kv electrospray current, mass accuracy of 100,000, a balanced AGC target, and an injection time of 100 m/s.

### **Quality control and extraction blanks**

To help ensure that there was no bias in our experimental design an equal number of cells from each replicate was sorted onto each plate, with the samples from each replicate evenly distributed across the run order to help reduce within batch effects. Extraction blanks were created by diverting a droplet of FACS sheath fluid into 8 wells on each plate, with these wells subsequently treated identically to single cell samples. Quality control samples were generated by pooling additional neurons by FACS and spinning them down to produce a bulk pellet. This pellet was then extracted by adding 1.8 ml of MTBE:MeOH (5:1, v:v) and was then vortexed for 60 seconds after which 600  $\mu$ l of water was added. Subsequently the sample was spun at 2000 rpm ( $845 \times g$ ) for 2 minutes to produce a phase separation with the upper organic phase removed. Prior to adding QC's to the plate, the material was analysed by LESA to assess concentration with MTBE added to reduce the concentration to give a signal within 2 orders of magnitude of the single cell signal. Once the QC sample was diluted 25  $\mu$ l was added to 8 wells per plate and was subsequently dried under a continuous flow of nitrogen.

### **Optimisation of analytical method**

Method optimisation was performed using a 1 nM solution of glycerophosphocholine (C16:0-d31/C18:1) dried down in a glass coated 96 well plate under a continuous stream of nitrogen. Liquid extraction surface analysis (LESA) was performed on these samples with a single parameter at a time with data collected from 5 replicates of each setting.

We demonstrated that at this low analyte concentration increasing infusion back pressure had no effect on signal abundance or signal to noise ratio (Supplemental Figure 6). The data also showed that increasing the electrospray current increased both signal abundance and signal to noise ratio up to a current of 1.4 kV, however when it was increased to 1.5kV we observe a slight increase in signal abundance but a slight reduction in signal to noise ratio (Supplemental Figure 7). When looking at the analysed mass window it can be seen that narrowing the mass range analysed increased both signal abundance and signal to noise ratio (Supplemental Figure 8). Reducing the solvent volume used to extract the sample from 7  $\mu$ l to 3  $\mu$ l increased both signal abundance and signal to noise ratio (Supplemental Figure 9). However, reducing this further to 1.5  $\mu$ l actually reduced the measured signal abundance and signal to noise ratio (Supplemental Figure 9).

### **Quantification and statistical analysis**

#### **Data processing**

Data processing was performed in R using an in-house pipeline (for code see Supplemental Information 2). Scans collected between 20 and 78 seconds were integrated and compared to

a list of 578 known lipid *m/z*'s (Supplemental Table 3) identifying the signal abundance and its deviation from the expected mass. Signal identification was done by comparing the abundance and mass deviation of the nearest signal to the target *m/z* in both FACS blanks and single cell samples. A signal was defined as present if it had a signal to noise ratio of  $> 1.5$  and an average deviation of less than 10 ppm, if no signal was observed within 10 ppm the signal was defined as missing, with a positive signal identified in at least 5% of cells in two of the three analysed sample sets. To be defined as a measured signal a linear response between signal abundance and cell number ( $r>0.5$ ) had to be observed (Table 1). Data processing for 560 samples and the accompanying 56 blanks and 56 QC samples took approximately 3 hours with signal identification and data clean up taking about 2 hours, which is comparable with other high throughput lipidomics methods (Harshfield et al., 2019).

### **Statistical Analysis**

Prior to statistical analysis data was normalised to the total signal, this was done by calculating the abundance of all signals within each sample with all the measured signals then divided by the mean abundance. The mean abundance of each lipid was calculated for each biological replicate to give a ‘population’ lipid profile, which could be compared to single cell profiles. Multivariate analysis using principal component analysis (PCA) and partial least square discriminant analysis (PLS-DA) were performed with SIMCA v13.0.4 (Umetrics, Umeå, Sweden) all data was logarithmically transformed (base10) and scaled to unit variance (UV). The relationship between individual lipid species and given Y-variables were determined using generalised linear models (GLM) applied to the whole of the dataset, all models were calculated in ‘R’ (version 3.4.2).

### **Data and code availability**

All R code used in this study is included in the supplemental information. All data will be upload to a publically available repository on publication of the manuscript.