

# Non-invasive monitoring of midbrain dopaminergic progenitor cell production from human pluripotent stem cells

**Mio Iwasaki**

Kyoto University - Yoshida Campus: Kyoto Daigaku

**Nicola J. Drummond**

The University of Edinburgh School of Biological Sciences

**David E. McNay**

The University of Edinburgh School of Biological Sciences

**Maurice A. Canham**

The University of Edinburgh School of Biological Sciences

**Daniel Tikhomirov**

The University of Edinburgh School of Biological Sciences

**Donia Arafa**

The University of Edinburgh School of Biological Sciences

**Mariah Lelos**

Cardiff University

**Asuka Morizane**

Kyoto University - Yoshida Campus: Kyoto Daigaku

**Tilo Kunath**

[Tilo.Kunath@ed.ac.uk](mailto:Tilo.Kunath@ed.ac.uk)

The University of Edinburgh <https://orcid.org/0000-0002-8805-7356>

---

## Research Article

**Keywords:** pluripotent stem cells, secreted biomarkers, dopaminergic differentiation, non-invasive monitoring, ELISA, Parkinson's disease

**Posted Date:** September 2nd, 2023

**DOI:** <https://doi.org/10.21203/rs.3.rs-2996413/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

# Abstract

**Background** An emerging regenerative medicine for Parkinson's is cell replacement therapy with midbrain dopaminergic (mDA) progenitor cells produced from human embryonic stem cells (hESCs) or induced pluripotent stem cells (iPSCs). Although protocols are well described, the process of hESC/iPSC differentiation remains challenging. Furthermore, different clonal iPSC lines can have variable responses to the same mDA protocol conducted in the same laboratory. **Methods** One solution to address this issue is to identify secreted biomarkers that are predictive of mDA differentiation efficiency. We used candidate approaches and unbiased proteomic methods to interrogate the secretome of cells as they are converting from pluripotent cells to committed mDA progenitor cells. We used conditioned media from two laboratories with different mDA protocols (Edinburgh, Kyoto) and multiple pluripotent stem cell lines (RC17, MasterShef7, 1231A3, 404C2). **Results** We have identified a collection of secreted molecules, including TFF3, CORIN, PDGFC, SERPINF1, and NRP1, that increase in conditioned media during the early stages of mDA differentiation. The kinetics of up-regulation and abundance of secreted biomarkers exhibited unique signatures for each cell line. However, in all cases we could identify biomarker signatures as early as day 4 of mDA protocols that were predictive of mDA differentiation efficiency. We further examined the secretome of differentiating striatal and cortical neurons to discover biomarkers of non-mDA cells, which led to the identification of PLA1 and NCAM1 as diagnostic markers that should remain low during mDA differentiation. Finally, we multiplexed selected positive and negative markers to construct custom Luminex assay systems that can non-invasively analyse media samples during the time-course of mDA progenitor cell production for five hESC/iPSC lines. **Conclusion** We have identified a collection of over 10 secreted biomarkers useful for non-invasively monitoring the production of mDA progenitor cells for day 4 of differentiation from multiple clonal pluripotent stem cell lines.

## Background

Current and emerging regenerative medicine therapies rely on differentiation of human pluripotent stem cells – human embryonic stem cells (hESCs) or induced pluripotent stem cells (iPSCs) – into specialized cell types. One of the most advanced clinical efforts in this area is cell replacement therapy (CRT) for Parkinson's, an incurable neurodegenerative disorder that affects more than 8 million people worldwide [1]. Many of the hallmark motor features of Parkinson's are predominantly due to the loss of midbrain dopaminergic (mDA) neurons of the *substantia nigra* that innervate the dorsal striatum. Grafting of human fetal mesencephalic tissue, the embryonic origin of mDA neurons, into the striatum of Parkinson's patients has demonstrated that CRT can, in some cases, restore dopamine transmission to normal levels, reverse motor dysfunction and reduce dependence on dopaminergic medicines for 15 years or more [2, 3]. However, the use of fetal tissue has considerable ethical and logistical challenges precluding its widespread clinical application. hESCs and iPSCs provide an inexhaustible source of specialized cells that have the potential to replace the need for fetal tissue if functional mDA neurons can be reliably produced from them. The ventral midbrain floor plate is the embryological origin of nigral mDA neurons during development [4, 5]. Recapitulation of key signalling pathways important for ventral midbrain development led to hESC/iPSC differentiation protocols for efficient production of functional mDA neurons [6–8]. Importantly, transplantation of hESC/iPSC-derived mDA cells into the striatum can rescue pre-clinical animal models of Parkinson's [6, 9, 10].

One of the major challenges of applying hESCs and iPSCs to regenerative medicine is the intrinsic differences in differentiation potential between cell lines [11–13]. We have demonstrated that different clonal hiPSC lines

established from the same individual can exhibit marked differences in dopaminergic differentiation potential in culture [14]. The underlying reasons for this variability are complex, and can include genetic mutations, epigenetic status, cell line evolution during clonal expansion, and the state of the pluripotent cells just prior to differentiation initiation. The practical implications of these findings mean that *(i) protocols are not easily transferable between different cell lines, (ii) differentiation efficiencies can vary from experiment to experiment, and (iii) trouble-shooting and protocol optimisation can be lengthy and very inefficient*. This is a challenge for all regenerative medicine therapies that rely on differentiation of hESCs or iPSCs. As part of the solution to this problem we propose to non-invasively monitor the state of the cells during the differentiation process with a focus on the earliest stages of lineage commitment. The first few days and weeks of a differentiation protocol is critical for getting the cells on the correct trajectory to produce the desired cell types. As hESCs/iPSCs are transitioning from a pluripotent state to committed progenitors, and further onto differentiated cell types, significant and dynamic changes in the transcriptome and proteome are occurring over time. In addition to changes in expression of well-known transcription factors and cell surface proteins, there are also concomitant changes to the secretome during differentiation. We hypothesise that *variability and heterogeneity of midbrain dopaminergic differentiation can be accurately monitored in real-time by measuring the abundance of informative lineage-specific secreted factors*. Informative biomarkers should be (i) present in the conditioned medium, (ii) easily quantified by ELISA-based methods, and (iii) change dynamically during the time-course of differentiation in a manner that is reflective of the cell types in culture. The most useful markers should also be able to *predict* the cell types that will emerge during differentiation as it progresses. A library of positive secreted biomarkers for cell types of interest, especially for the earliest stages of differentiation, will provide valuable information of cell identity and cell heterogeneity in real-time, and potentially predict the emergence of cell types at later time-points of the differentiation. This knowledge can be used for protocol optimisation for new cell lines, trouble-shooting differentiation procedures, and importantly, provide quality control measures at different stages of a cell manufacturing process. The latter point has significant cost implications when differentiation is occurring under GMP conditions for clinical applications. There are two reports of non-invasive in-line monitoring of hiPSC differentiation using lineage-specific secreted factors. The first used Cerberus1 (CER1) to monitor the differentiation towards definitive endoderm [15]. The authors show that the amount of CER1, as measured by ELISA, in the conditioned medium as early as day 4 of differentiation, is highly correlated to the percentage of SOX17/FOXA2 double-positive endoderm progenitors that emerge during differentiation [15]. The second study identified a specific secreted biomarker of human pluripotent stem cells that decreased rapidly in conditioned medium upon differentiation [16]. The biomarker is a hyperglycosylated podocalyxin that could be detected by a specific lectin probe, rBC2LCN, and was capable of reporting the presence of pluripotent stem cells in mixed cultures in a dose dependent manner [16]. This latter study highlights the need and utility of having secreted biomarkers of undesired cell types, such as undifferentiated cells. Similarly, negative biomarkers for cell types of incorrect developmental identity to the target cells would also be informative during non-invasive real-time monitoring of hESC/iPSC differentiation.

Here we have used candidate and unbiased approaches to identify a library of secreted biomarkers for the production of mDA progenitor cells from hESCs/iPSCs. We took advantage of publicly-available transcriptomic datasets of hESCs undergoing midbrain floor plate differentiation [6], recent single-cell RNAseq data of mDA lineages [17], and the efforts to define and annotate the human secretome [18]. We also adopted an unbiased quantitative mass spectrometry approach to analyse the conditioned media of two hESC lines and two iPSC lines differentiating into mDA progenitor cells in two laboratories using distinct modifications of the midbrain

floor plate protocol. These combined efforts led to the identification and validation of 11 secreted biomarkers (TFF3, SERPINF1, NRP1, CORIN, PDGFC, SHH, IGFBP2, ALCAM, SPP1, PLA1, and NCAM1) that can non-invasively reflect the collective state of a cell population during mDA differentiation. Furthermore, some of the biomarkers can be detected at day 4 of the midbrain floor plate protocol, which provides significant and early predictive information on mDA progenitor cell differentiation trajectories.

## Methods

### Human ESC and iPSC culture

Human iPSCs (1231A3 and 404C2) were maintained on iMatrix-511 silk (0.5  $\mu\text{g}/\text{cm}^2$ , Matrixome)-coated plates in StemFiT AK-02N (Ajinomoto) [19]. When passaging the cells onto iMatrix-511 silk-coated plates, the iPSCs were dissociated into single cells with TrypLE select (Thermo Fisher Scientific) and replated at a density of  $2.0 \times 10^4$  cells per 6-well plate with StemFiT AK-02N supplemented with 10  $\mu\text{M}$  Y-27632 (Fujifilm). Approval for the use of MasterShef7 (MShef7) and RC17 hESCs used in this study was granted by the MRC Steering Committee for the UK Stem Cell Bank and for the Use of Stem Cell Lines (ref. SCSC13-18 for MShef7 and ref. SCSC13-19 for RC17). hESCs and NAS2 iPSCs were maintained in self-renewing conditions on Laminin-521-coated plates (L521, 5  $\mu\text{g}/\text{ml}$ , Biolamina) in StemMACS™ iPS-Brew XF (iPS-B, Miltenyi Biotec). Once cells reached 70–90% confluency, they were passaged as clumps with EDTA (0.5 mM Thermo Fisher Scientific), at a split ratio of 1:3 to 1:6.

### mDA neuronal differentiation with Kyoto protocol

Human iPSCs were differentiated to mDA progenitors with the protocol we previously reported [10, 20]. Briefly the cells were seeded at a density of 500,000 cells/ $\text{cm}^2$  to iMatrix-511 silk-coated plates. Cells were differentiated in the medium supplemented with KnockOut Serum Replacement (Thermo Fisher Scientific) using dual SMAD inhibition strategy. The medium was changed every day from day 1 to day 11 as the cell density was higher in the Kyoto protocol than in Edinburgh's one. The volumes of the medium were increased step-by-step; 1.0 ml/well for day 0–2, 1.5 ml/well for day 3–6, and 2.0 ml/well for day 7–11. At day 11, the cells were lifted with Accumax (Innovative Cell Technologies) and replated to iMatrix-511 silk-coated 24-well plates at a density of 500,000 cells/ $\text{cm}^2$ . In the original protocol, the differentiated cells were sorted with Corin antibodies at differentiation day 12–14 and switched the condition from attachment culture to the suspension culture forming cell aggregates. In this research, we modified the protocol in which the cells were just passaged at day 11 without sorting and kept the attachment culture to be compared with other mDA protocols. From day 11, cells were fed with neural differentiation medium (NDM) with ascorbic acid (AA, 0.2 mM, Sigma A4034), brain-derived neurotrophic factor (BDNF, 20 ng/ml, R&D 248-BD-025) and glial cell line-derived neurotrophic factor (GDNF, 10 ng/ml, R&D 212-GD). Y27632 (20 $\mu\text{M}$ ) was present in medium at day 0 and 11. From day 21 dibutyl cyclic AMP (dcAMP, 0.5 mM, Sigma D0627) and DAPT (1  $\mu\text{M}$ , Tocris) were added.

### mDA neuronal differentiation with Edinburgh protocol

Self-renewing hESCs were lifted using EDTA (0.5 mM) and seeded at a density of 40,000 cells/ $\text{cm}^2$  onto Laminin-111 (L111, 5  $\mu\text{g}/\text{ml}$ , Biolamina) coated plates. Cells were plated for differentiated in medium made up of 50% DMEM/F12 (Thermo Fisher Scientific) and 50% Neurobasal™ medium (Thermo Fisher Scientific), B27 Supplement (without Vitamin A, 1:50, Thermo Fisher Scientific), N2 (1:100, Thermo Fisher Scientific) and L-

glutamine (2mM, Thermo Fisher Scientific) containing Y27632 (10  $\mu$ M, Tocris), SB431542 (10  $\mu$ M, Tocris), LDN193189 (100 nM, Stemgent), Shh-C24II (0-600 ng/ml, R&D) and CHIR99021 (0–2  $\mu$ M Miltenyi Biotec). On day 2 the medium was changed with the above medium without Y27632. On day 4 and day 7 base medium consisted of 50% DMEM/F12, 50% Neurobasal™ medium, B27 Supplement (1:100), N2 (1:200) and L-glutamine (2 mM) supplemented with SB431542 (10  $\mu$ M), LDN193189 (100 nM), Shh-C24II (600 ng/ml, R&D) and CHIR99021 (0–2  $\mu$ M Miltenyi Biotec). On day 9 the above base medium was used and supplemented with FGF8b (100 ng/ml, R&D) and heparin (1  $\mu$ g/ml, Sigma). On day 11, cells were lifted using Accutase (Sigma) and replated at 800,000 cells/cm<sup>2</sup> on L111-coated plates. From day 11 cells were cultured in base medium consisting of Neurobasal™ medium, B27 (1:50) and L-glutamine (2 mM). At day 11 medium was supplemented with ascorbic acid (0.2 mM, Sigma), BDNF (20 ng/ml, Peprotech), GDNF (10ng/ml, Peprotech), FGF8b (100 ng/ml), heparin (1  $\mu$ g/ml) and Y27632 (10  $\mu$ M). On day 14 the medium was changed with the above medium without Y27632. On day 16 cells were lifted with Accutase and replated at a density of 800,000 cells/cm<sup>2</sup> in base medium as on day 11 but supplemented with BDNF (20 ng/ml), GDNF (10 ng/ml), ascorbic acid (0.2 mM), dibutyl cyclic AMP (db-cAMP, 0.5 mM, Sigma), DAPT (1  $\mu$ M) and Y27632 (10  $\mu$ M). From day 18 onwards medium was changed every 2–3 days with day 16 medium but without Y27632. A detailed version of this protocol is available at: [dx.doi.org/10.17504/protocols.io.bddpi25n](https://dx.doi.org/10.17504/protocols.io.bddpi25n).

Minor modifications of the Edinburgh protocol were implemented for mDA induction of 1231A3 and 404C2 iPSC lines. The initial cell density at day 0 was optimized: 1231A3 was 40,000 cells/cm<sup>2</sup>, and 404C2 was 100,000 cells/cm<sup>2</sup>. The matrix coating was changed from LN111 to iMatrix-511silk. For ventralization, purmorphamine (0.25  $\mu$ M, FujiFilm) was supplemented in addition to 600 ng/ml Shh-C24II. As in previous studies, the optimal concentration of the CHIR99021 needed to be optimized for both cell lines as estimated by immunocytochemical analysis with LMX1A, FOXA2, OTX2 and CORIN antibodies.

## Cortical neuron induction of MShef7 and RC17 hESCs

Self-renewing hESCs were lifted using EDTA (0.5 mM) and seeded at a density of 80,000 cells/cm<sup>2</sup> onto Laminin-111 (L111, 5  $\mu$ g/ml, Biolamina) coated plates. Cells were plated for differentiated in medium made up of 50% DMEM/F12 (Thermo Fisher Scientific) and 50% Neurobasal™ medium (Thermo Fisher Scientific), B27 Supplement (without Vitamin A, 1:50, Thermo Fisher Scientific), N2 (1:100, Thermo Fisher Scientific) and L-glutamine (2mM, Thermo Fisher Scientific) containing Y27632 (10  $\mu$ M, Tocris), SB431542 (10  $\mu$ M, Tocris), LDN193189 (100 nM, Stemgent). On day 2 the medium was changed with the above medium without Y27632. On day 4, day 7 and day 9 medium was changed consisting of 50% DMEM/F12, 50% Neurobasal™ medium, B27 Supplement (1:100), N2 (1:200) and L-glutamine (2 mM) base medium supplemented with SB431542 (10  $\mu$ M), LDN193189 (100 nM). On day 11 cells were lifted with Collagenase IV (200 U/ml, Thermo Fisher Scientific) and replated at a ratio of 1:1.5 in Laminin-111 coated plates. Cells were plated in 50% DMEM/F12, 50% Neurobasal™ medium, B27 Supplement (1:100), N2 (1:200) and L-glutamine (2 mM) base medium supplemented with Y27632 (10  $\mu$ M). On day 14 medium was changed to remove Y27632 (10  $\mu$ M). On day 16 cells were lifted as clumps as on day 11 using Collagenase IV. Medium was changed on day 18 to remove Y27632. Day 19, day 21, day 23 medium was changed. On day 25 cells were lifted using Accutase, passed through a cell strainer (40  $\mu$ m) and plated at 70,000 cells/cm<sup>2</sup> in base medium supplemented with BDNF (20 ng/ml) and GDNF (10 ng/ml) and Y27632 (10  $\mu$ M). On day 28 the medium was changed with the above medium without Y27632. 50% of the medium was replaced every 2–3 days.

# Lateral ganglionic eminence (LGE) induction protocol

Human iPSCs were differentiated into LGE progenitors with the previously reported protocol using Activin A [21]. In brief, at day 0, the differentiation was started as suspension culture using the SFEBq method [22], combined with the dual SMAD inhibition strategy [23]. At day 9 the cell aggregations were treated with Accumax, dissociated into single cells, and then plated to iMatrix-511 silk-coated plates to mono-layer culture with the addition of Activin A (day 9–21, 25 ng/mL, R&D).

## Gene expression analysis by RT-qPCR

RNA extraction was performed using either the MasterPure™ Complete DNA and RNA Purification Kit (Epicenter, MC85200), according to manufacturer's instructions or RNA Plus Mini Kit (Qiagen), according to manufacturer's instructions. RNA concentration was quantified using NanoDrop spectrophotometer. RNA (1 µg in 10 µl) was used for cDNA synthesis, using M-MLV reverse transcriptase (Thermo Fisher Scientific). Samples were incubated with 1 µl dNTP mix (10 mM, Thermo Fisher Scientific) and 1 µl Random Primer Mix (60 µM NEB) at 65°C for 5 minutes and chilled on ice. Samples were centrifuged and 4 µl 5x First strand buffer (Thermo Fisher Scientific) 2 µl DTT (0.1 M, Thermo Fisher Scientific) and 1 µl RNaseOUT (40 units/µl, Thermo Fisher Scientific) were added. Samples were incubated at 37°C for 2 minutes then 1 µl M-MLV reverse transcriptase (200 units/µl, Thermo Fisher Scientific) was added. This was incubated at room temperature for 10 minutes before incubating at 37°C for 1 hour. Inactivation of the enzyme by incubation at 90°C for 10 minutes was performed. qPCR was performed using the Roche LightCycler® 480 System with either the Universal Probe library (UPL) (Roche) system or IDT predesigned qPCR Assays. The Roche UPL Assay design centre was used to design intron-spanning Forward (Primer 1) and Reverse (Primer 2) primers with a specific UPL Probe for each gene (Table 1). Reactions (10 µl) were performed containing cDNA primers, UPL Probe, LightCycler® 480 Probes Master mix (Roche) and PCR water in 384-well plates according to manufacturer's instructions. Alternatively, reactions (10 µl) were performed containing IDT Predesigned qPCR primers (Table 2) and probe mix, IDT PrimeTime gene expression master mix and PCR water in 384-well plates according to manufacturer's instructions. The data was normalized to the level of *TATA-binding protein (TBP)* and data of technical replicates or replica experiments were plotted.

Table 1  
Primers for UPL system qPCR

| Gene Name       | Primer 1                | Primer 2                    | UPL Probe |
|-----------------|-------------------------|-----------------------------|-----------|
| <i>TBP</i>      | GAACATCATGGATCAGAACAACA | ATAGGGATTCCGGGAGTCAT        | 87        |
| <i>LMX1A</i>    | TGGAGGAGAACTTCCAAAGC    | CAGACAGACTTGGGGCTCAC        | 3         |
| <i>FOXA2</i>    | GGGTGATTGCTGGTCGTTT     | ATACTGGAAGCCGAGTGCAT        | 5         |
| <i>CORIN</i>    | AAGAGCGCTACCGCAGAG      | ACCGGAGGAGGTTAGCAGTC        | 46        |
| <i>SHH</i>      | AGCAAAGCCTGATGTGTAAGC   | GTCCTAAGAGAAGGTTGTTGGTG     | 10        |
| <i>SPON1</i>    | ACTTCTACAAGCCGGAACC     | GGGCAATTAATGTGAATCCTCT      | 52        |
| <i>TFF3</i>     | GCTGCTGCTTTGACTCCAG     | TGGAGGTGCCTCAGAAGGT         | 4         |
| <i>DKK1</i>     | CAGGCGTGCAAATCTGTCT     | AATGATTTTGATCAGAAGACACACATA | 4         |
| <i>SERPINF1</i> | GTGTGGAGCTGCAGCGTAT     | TCCAATGCAGAGGAGTAGCA        | 57        |
| <i>PKDCC</i>    | AACGAGAAGCGGAACCTCTA    | AGCAGAGGACGCAGTGAAG         | 51        |
| <i>NPY</i>      | CTCGCCCGACAGCATAGTA     | CCCTTCTTGCCCAAACTCTT        | 83        |
| <i>THBS4</i>    | ATTGGGAAATTTTGGGCAGT    | ATGGGGATGGAAGAGGAGAT        | 58        |
| <i>BMP7</i>     | AAGAAGCACGAGCTGTATGTCA  | ACAGTAGTAGGCGGCGTAGC        | 26        |
| <i>NTNG1</i>    | ACATAAAGGTGCGAGGAAGG    | TCGCATGTCAATTTGCTGTT        | 59        |
| <i>FBLN1</i>    | CGAATGCAAGACGGGTTACT    | GCTGGCACTCGTTGACATC         | 18        |
| <i>NRCAM</i>    | AGGGACCAGAGCATGTGAAC    | CAAAGAAGCTCCGAGAACCA        | 27        |
| <i>FLRT2</i>    | GACCCCTAAAACAATACCATGC  | TGGTTAAGC GJUTCAATCTGCAATC  | 10        |

Table 2  
Primers for IDT predesigned qPCR Assays

| Gene Name       | Primer 1                | Primer 2                 | IDT Assay Name       |
|-----------------|-------------------------|--------------------------|----------------------|
| <i>TBP</i>      | GCTGTTTAACTTCGCTTCCG    | CAGCAACTTCCTCAATTCCTTG   | Hs.PT.58v.39858774   |
| <i>LMX1A</i>    | AGAACTTCCAAAGCGCGAT     | AGAAACCTGTCCAAGATGACC    | Hs.PT.58.27592353    |
| <i>FOXA2</i>    | GGAGCGGTGAAGATGGAAG     | TGTTTCATGCCGTTTCATCCC    | Hs.PT.58.26032236    |
| <i>EN1</i>      | GTACTGCACACGTTATTCGGA   | TTGTCCTCCTTCTCGTTCTTC    | Hs.PT.58.14559889    |
| <i>CORIN</i>    | GCTATGAGTTGGCACAGTT     | GAGCCCCATGAAGTTAATCCAA   | Hs.PT.58.26629660    |
| <i>TFF3</i>     | TGTCCTCCAGCTCTGCT       | CAGTCCACCCTGTCCTTG       | Hs.PT.58.1814807     |
| <i>SERPINF1</i> | ACCTGCAAGAGATCAACAACT   | GAGTCAAACCTTTGTTACCCACTG | Hs.PT.58.20553847    |
| <i>NRP1</i>     | GCAGTATTCCCCAAACCACT    | GACACCATACCCAACATTCCA    | Hs.PT.58.38768935    |
| <i>PDGFC</i>    | CTTGAAGACCCAGAAGATGACA  | CAGTACCAGAACCACACCAG     | Hs.PT.58.3039105     |
| <i>SPP1</i>     | CCCCACAGTAGACACATATGATG | TTCAACTCCTCGCTTTCCAT     | Hs.PT.58.19252426    |
| <i>ALCAM</i>    | CAAGCAACCATCTAAACCTGAA  | TCCTGTACCATTGATATTGCC    | Hs.PT.56a.20098970.g |
| <i>IGFBP2</i>   | ACCTCTACTCCCTGCACAT     | TTCACACACCAGCACTCC       | Hs.PT.58.2693427     |

## Immunostaining

Cells were fixed with 4% formaldehyde for 20 minutes and washed three times in PBS. Fixed cells were blocked for 30 minutes with 0.1% Triton X-100, 2% donkey serum in PBS. Primary antibodies tyrosine hydroxylase (TH, 1:1000, rabbit, Millipore) and  $\beta$ -III tubulin (TuJ1, 1:1000, mouse IgG2a, R&D) were incubated overnight at 4°C and then washed three times in PBS with 0.1% Triton X-100. Secondary antibodies donkey anti rabbit 594 (Thermo Fisher Scientific) and donkey anti mouse 488 (Thermo Fisher Scientific) were incubated overnight at 4°C and then washed three times in PBS with 0.1% Triton X-100. Cells were incubated with DAPI (10  $\mu$ g/ml, Thermo Fisher Scientific) and washed before imaging on the Opera Phenix Plus High-Content Screening System (PerkinElmer).

For iPSC differentiation, mDA cells at day 11 were dissociated, replated at a density  $3.4 \times 10^5$  cells/well of 96-well microplate with polymer coverslip bottom (ibidi) coated with iMatrix-511 silk. Three days later (total day 14) cells were fixed with 4% paraformaldehyde. The cells were permeabilized and blocked with 0.3% TritonX and 2% normal donkey serum, and stained immunocytochemically with primary and secondary antibodies (Table 3). All secondary antibodies raised in donkey were from Thermo Fisher Scientific. The images were obtained by fluorescence microscopy (BZ-X810, Keyence).

Table 3  
Primary and secondary antibodies

| Antibodies                           | Company   | Cat#     | dilution |
|--------------------------------------|-----------|----------|----------|
| LMX1A/B (rabbit)                     | Millipore | AB10533  | 1000     |
| Foxa2 (goat)                         | R&D       | AF2400   | 500      |
| Corin (rat)                          | R&D       | MAB2209  | 500      |
| OTX2 (goat)                          | R&D       | AF1979   | 1000     |
| TH (rabbit)                          | Chemicon  | AB152    | 400      |
| $\beta$ -III tubulin (TuJ1) (rabbit) | Biologend | PRB-435P | 500      |
| CTIP2 (25B6) (rat)                   | abcam     | ab18465  | 500      |
| DARPP32 (rabbit)                     | abcam     | ab40801  | 400      |

| Secondary antibodies | Alexa Fluor™ 488 | Alexa Fluor™ 594 | dilution |
|----------------------|------------------|------------------|----------|
| Anti-mouse IgG       | A21202           | A21203           | 400      |
| Anti-rabbit IgG      | A21206           | A21207           | 400      |
| Anti-goat IgG        | A11055           | A11058           | 400      |
| Anti-rat IgG         | A21208           | A21209           | 400      |

## Proteomic analysis

Conditioned media samples were mixed with equal volume of ice-cold 2X lysis buffer consists of 200 mM Tris-HCl (pH9.0) (Nippon Gene), 24 mM sodium deoxycholate (WAKO) and 24 mM sodium lauroyl sarcosinate (WAKO) supplemented with 1% phosphatase inhibitors (Sigma-Aldrich) and 1% protease inhibitor (Sigma-Aldrich). The lysates were subjected to reduction (final 1 mM dithiothreitol, FujiFilm), alkylation (final 10 mM iodoacetamide, FujiFilm), Lys-C/trypsin digestion (enzyme ratio: 1/100, Promega) and desalting by in house SDB-XC StageTip (PMID: 12585499), as previously described [24]. Peptides (200 ng) were loaded and separated on an Aurora column (25 cm length, 75  $\mu$ m i.d., IonOpticks) using a nanoElute (Bruker) for subsequent analysis by timsTOF Pro 2 system (Bruker). The mobile phases were composed of 0.1% formic acid (solution A) and 0.1% formic acid in acetonitrile (solution B). A flow rate of 400 nl/min of 2–17% solution B for 60 min, 17–25% solution B for 30 min, 25–37% solution B for 10 min, 37–80% solution B for 10 min, and 80% solution B for 10 min was used (120 min in total). The applied spray voltage was 1400 V, and the interface heater temperature was 180°C. To obtain MS and MS/MS spectra, the Parallel Accumulation Serial Fragmentation (PASEF) acquisition method with data-independent acquisition (DIA) mode was used (diaPASEF) [25]. For diaPASEF settings, 1.7 s per one cycle with precursor ion scan and 16 times diaPASEF scans were conducted with the MS/MS isolation width of 25 m/z, precursor ion ranges of 400 – 1200 m/z, ion mobility ranges of 0.57 – 1.47 V s cm<sup>-2</sup>. The obtained DIA data were searched by DIA-NN (v1.8.1) [26] against selected human entries of

UniProt/Swiss-Prot release 2020\_03 with the carbamidomethylation of cysteine as the fixed modification and protein N-terminal acetylation and methionine oxidation as the variable modification. For the other DIA-NN parameters, Trypsin/P protease, one missed cleavage, peptide length range of 7–30, precursor m/z range of 300–1800, precursor charge range of 1–4, fragment ion m/z range of 200–1800, and 1% precursor FDR were used. The values “Precursor.Quantity” from the results were used as peptide area values for the comparison. To detect mDA-specific proteins, we used the following strategy with in-house Perl scripts.

1. The peptide area values of technical duplicate were averaged.
2. The peptides and proteins were grouped into the five sample types of LGE, Cortex, iPSC, medium without (w/o) cells, and mDA.
3. Proteins specifically identified for each cell type were extracted.
4. Clustering analysis was performed using daily expression profiles of cell type specific identified proteins by R script with degPatterns module.

## ELISA

Conditioned media samples were collected at day 0, 2, 4, 7, 9, 11, 14 and 16 of the mDA differentiation. Samples were centrifuged at 10,000g for 5 minutes to remove any cell debris before being frozen. ELISAs were performed for target proteins TFF3, CORIN, PDGFC, NRP1, SERPINF1, SHH-N, NCAM1, PLA1 and IGFBP2 (R & D Systems, Table 4) according to manufacturer’s instructions. Optimization of appropriate dilution factors for each time point was required initially. The optical density of samples was measured on the FLUOstar Omega plate reader (BMG Labtech) using 450 nm and wavelength correction at 570 nm. Four (4) parameter logistic curves were created using the AAT Bioquest tool and the concentrations of target proteins within the samples determined.

Table 4  
ELISA kits

| Protein  | ELISA Kit                                                | Catalogue number |
|----------|----------------------------------------------------------|------------------|
| TFF3     | Human TFF3 Quantikine ELISA Kit                          | DTFF30           |
| TFF3     | Human TFF3 DuoSet ELISA                                  | DY4407           |
| CORIN    | Human CORIN Quantikine ELISA                             | DCRN00           |
| PDGFC    | Human PDGF-CC Quantikine ELISA Kit                       | DCC00            |
| PDGFC    | Human PDGF-CC DuoSet ELISA                               | DY1687-05        |
| NRP1     | Human Neuropilin-1 DuoSet ELISA                          | DY3870-05        |
| SERPINF1 | Human SerpinF1/PEDF DuoSet ELISA                         | DY1177-05        |
| SHH-N    | Human Sonic Hedgehog/Shh N-Terminus Quantikine ELISA Kit | DSHH00           |
| NCAM1    | Human NCAM1/CD56 DuoSet ELISA                            | DY2408           |
| PLAU     | Human u-plasminogen activator/Urokinase DuoSet ELISA     | DY1310           |
| IGFBP2   | Human IGFBP-2 DuoSet ELISA                               | DY674            |

## Multiplex Luminex Assay

Conditioned media samples were collected at day 0, 4, 7, 11, 14 and 16 of the Edinburgh mDA differentiation protocol or day 0, 4, 7, 11 and 21 of the Kyoto differentiation protocol. Samples were centrifuged at 10,000g for 5 minutes to remove any cell debris before being frozen. Multiple analytes were measured simultaneously using a custom designed Luminex discovery 6-plex assay (R & D Systems) containing TFF3, SERPINF1, NRP1, PDGFC, PLAU and NCAM1 or a custom designed 8-plex assay (R & D Systems) containing TFF3, SERPINF1, NRP1, IGFBP2, ALCAM, SPP1, PLAU and NCAM1. The Luminex assay was performed according to the manufacturer's instructions and the analyte concentrations were measured using the BioRad 200 HTF Luminex (BioRad). Several dilutions were included for each time point for all analytes to be within the standard curve measurements. The Bio-Plex Manager™ software (BioRad) was used to determine the analyte concentrations within each sample.

## Results

Differentiation of human pluripotent stem cells into midbrain floor plate cells requires an orchestrated manipulation of diverse signalling pathways to achieve the desired cell types. The manipulations include (i) inhibition of TGFβ/activin/nodal and BMP signalling (dual-Smad inhibition), (ii) strong activation of the SHH pathway, (iii) activation of FGF signalling, and (iv) moderate activation of WNT signalling [6, 7, 23, 28–30]. We re-analysed transcriptomic data that compared dual-Smad inhibition with LDN193189 and SB431542 (LSB) to the complete mDA protocol at day 11 and day 25 of differentiation, as well as LSB/Shh/FGF8 to the full mDA protocol that includes the indirect WNT activator, CHIR99021, for the same time-points [6]. The significantly up-regulated genes from these transcriptomic experiments were cross-referenced with the published human secretome list of 2640 proteins [18]. At day 11 of differentiation, genes for 21 secreted proteins were up-

regulated when compared to the LSB conditions, and 22 secreted proteins were significantly increased in mDA cells when compared the LSB/Shh/FGF8 condition (Supplementary Table 1). At day 25 of differentiation, 21 and 26 secreted proteins were up-regulated in mDA cells when compared to LSB and LSB/Shh/FGF8 conditions, respectively (Supplementary Table 2). The overlap of genes for secreted proteins within these datasets was significant with 11 genes common at day 11 of differentiation, and 18 genes in common at day 25 (Fig. 1A). Four genes, *TFF3*, *SPON1*, *SERPINF1*, and *SULF2* were common to all four datasets. Total RNA from hESC-derived neural cells differentiated in 0 ng/ml, 200 ng/ml, or 600 ng/ml SHH-C24II for 11 days was subjected to RT-qPCR for selected up-regulated genes. Gene expression analysis for known ventral midbrain markers, *FOXA2*, *LMX1A*, *CORIN*, and *SHH*, exhibited a dose-dependent increase in expression with increasing SHH-C24II concentration (Fig. 1B). Candidate secreted biomarkers, *TFF3*, *SERPINF1*, *PKDCC*, *NPY*, *BMP7*, *NTNG1*, *NRCAM*, and *FLRT2* also exhibited a SHH dose-dependent up-regulation of expression (Fig. 1B). *SPON1* and *FBLN1* increased in expression at 200 ng/ml, but not 600 ng/ml SHH-C24II, while *DKK1* and *THBS4* did not exhibit SHH-dependent expression (Fig. 1B).

We next examined the WNT-responsiveness of selected biomarker genes by performing qPCR on hESCs differentiated in 0.5  $\mu$ M, 1.0  $\mu$ M, and 1.5  $\mu$ M CHIR99021 (CHIR), with 1.0  $\mu$ M being the ideal concentration of CHIR to obtain correctly patterned ventral midbrain floor plate cells. At day 11 of differentiation well-characterized ventral midbrain markers *LMX1A*, *FOXA2*, *EN1*, and *CORIN* exhibited highest expression at 1.0  $\mu$ M CHIR (Fig. 2A). *TFF3* and *SERPINF1* exhibited a similar peak of expression at 1.0  $\mu$ M CHIR (Fig. 2A). Additional candidate genes were investigated based on the literature and whether they were compatible to be multiplexed on the Luminex assay platform [31]. ALCAM is a cell surface and secreted protein that marks ventral brain regions [32, 33], and mRNA for this gene also peaked at 1.0  $\mu$ M CHIR (Fig. 2A). SPP1, a secreted protein also known as osteopontin, was investigated since it is significantly up-regulated in A9 nigral neurons when compared to A10 ventral tegmental area neurons of the mouse midbrain [34]. *SPP1* exhibited highest expression at day 11 in cells differentiated in 1.0  $\mu$ M CHIR and was almost undetectable at 0.5  $\mu$ M and 1.5  $\mu$ M (Fig. 2A). Two additional Luminex-compatible genes, *Neuropilin-1* (*NRP1*) and *PDGFC*, also exhibited SHH-dependent gene expression, and maximal expression at 1.0  $\mu$ M CHIR (Fig. 2B). NRP1 is a secreted protein that is highly expressed in a sub-cluster of dopaminergic neurons that are triple-positive for *DAT*, *TH*, and *ALDH1A1* [17]. *PDGFC* is expressed in a radial glia sub-type of the ventral midbrain [35], and it is co-expressed with *TFF3*, *SERPINF1*, *WNT5A*, and *EN1* in “stage II” mDA-fated neural progenitor cells [36]. Finally, IGFBP2, a secreted protein with reagents available on the Luminex platform and specifically expressed in DAT-high cells [17], exhibited highest expression in day 11 mDA progenitors cells differentiated with 1.0  $\mu$ M CHIR (Fig. 2C). Neural cells induced in the three concentrations of CHIR were differentiated up to day 39 and immunostained for the pan-neuronal marker,  $\beta$ -III tubulin, and the dopaminergic marker, tyrosine hydroxylase (TH). Only cells cultured in 1.0  $\mu$ M CHIR produced neuronal networks that were double-positive for TH: $\beta$ -III tubulin (Fig. 2D).

In a parallel and unbiased approach we performed quantitative mass spectrometry on conditioned media (CM) from a time-course of midbrain floor plate differentiation of four human pluripotent stem cell lines using two mDA protocols (Edinburgh (EDI) and Kyoto). We also investigated CM from cortical and lateral ganglionic eminence (LGE) neuronal differentiation protocols, as well as media conditioned by self-renewing iPSCs. In an effort to increase the robustness of the dataset, the Kyoto iPSC lines (404C2, 1231A3) were differentiated in a modified Edinburgh (mEDI) protocol, as well as the Kyoto mDA protocol. Immunostaining for markers of mDA identity (*FOXA2*, *LMX1A/B*, *CORIN*, *OTX2*) at day 14 of the mEDI protocol showed good agreement with the same

iPSCs differentiated in the Kyoto protocol (Fig. 3A). Differentiation to day 35 in the mEDI produced neurons positive for both TH and  $\beta$ -III tubulin markers (Fig. 3B). The LGE protocol efficiently produced neurons double-positive for the striatal markers CTIP2 and DARPP32 (Fig. 3C).

CM was collected from hESCs and iPSCs during induction into (i) cortical, (ii) LGE, and (iii) mDA progenitor cells. The latter mDA samples were collected from 4 cell lines (RC17, MShef7, 404C2, 1231A3) in three mDA protocols (EDI, Kyoto, mEDI). CM from self-renewing iPSCs was also collected and analyzed. Finally, media without cells (w/o cells) was also analyzed to determine the protein content in the neural induction medium. In total, 113 samples, most in technical duplicates, were processed, digested, and subjected to unbiased quantitative mass spectrometry using the Parallel Accumulation Serial Fragmentation (PASEF) acquisition method with data-independent acquisition (DIA) mode [25] (Fig. 4A). Over 2500 proteins were identified across all samples and almost 2000 proteins were uniquely present in at least one mDA sample (Fig. 4B). However, the list of proteins that were present in mDA samples from all four cell lines and in the three mDA protocols was only 20, including TFF3, CORIN, SLIT2, and LGI1 (Fig. 4C). Only a fraction (< 15%) of the proteins identified in CM were annotated as “secreted” by UniProt or the published human secretome list [18] (Fig. 4B; Supplementary Table 3).

We next investigated individual secreted biomarkers by ELISA during a time-course of mDA differentiation of RC17 hESCs, and over a concentration range of SHH (0–600 ng/ml) and CHIR99021 (0–2.0  $\mu$ M). TFF3 was undetectable in self-renewing hESCs, but was observed in CM from day 4 of mDA progenitor cell induction (Fig. 5A). It exhibited a SHH-dose-dependent expression both at day 4 and day 7 of differentiation, and had a peak of expression at a CHIR concentration between 0.9–1.1  $\mu$ M (Fig. 5B,C). SERPINF1 was also detectable from day 4 of mDA induction, but peaked at day 7 before decreasing as differentiation progressed (Fig. 5D). Its sensitivity to SHH and CHIR at day 4 was not as marked as TFF3, but it showed dose-dependent expression that was similar to TFF3 at day 7, that is highest expression at 600 ng/ml SHH and a peak of expression between 0.9–1.1  $\mu$ M CHIR (Fig. 5E,F). Neuropilin-1 (NRP1) was low, but detectable in self-renewing hESCs and the early stages of differentiation, and began to increase from day 7 (Fig. 5G). NRP1 exhibited SHH-dependent expression at day 7, and peaks of expression in the middle range of CHIR concentrations, especially at day 14 (Fig. 5H,I). CORIN, a cell-surface marker that can be used by flow cytometry to enrich for mDA progenitor cells [10], is also cleaved from the cell surface [37]. This protein was one of 20 observed by mass spectrometry in all mDA samples for all cell lines. CORIN was first detectable in CM from day 9 and peaked at day 14 and 16 of mDA induction (Fig. 6A). It was sensitive to SHH concentration at day 14, and had a peak of expression between 0.9–1.1  $\mu$ M CHIR at the same time-point (Fig. 6B,C). PDGFC was undetectable in the early stages of mDA differentiation and became detectable at day 14 (Fig. 6D). It was not expressed in the absence of SHH, but did not show a dose-dependent expression at day 14 (Fig. 6E). However, PDGFC was particularly sensitive to CHIR concentration, exhibited a peak at 1.0  $\mu$ M CHIR, the ideal concentration for the RC17 hESC line in the EDI protocol. SHH-N was only measured in self-renewing hESC cultures from day 11 onwards since it is added to the mDA induction medium until day 9. SHH-N is absent in hESC CM, but readily detectable from day 11 (Fig. 6G). It was not sensitive to SHH concentration when measured at day 14 (Fig. 6H), but it did show differences of expression at differing CHIR concentrations (Fig. 6I). IGFBP2 was detectable in hESC CM, and did not significantly increase until day 14 of mDA induction (Fig. 6J). However, IGFB2 showed a CHIR-dose-dependent increase in expression as early as day 4 and day 7 of differentiation (Fig. 6K,L). PLAU was identified as a non-mDA marker from the proteomic analysis. It was detectable in hESC CM, and rapidly decreased as differentiation proceeded (Fig. 6M). At day 4 of mDA induction its expression was lowest in the optimal CHIR concentration (1.0

$\mu\text{M}$ ) for the RC17 hESCs (Fig. 6N). At day 7, when PLAU expression is 10-fold lower than at day 4, its expression in 1.0  $\mu\text{M}$  CHIR cultures was significantly lower than 1.5  $\mu\text{M}$  cultures and similar to neural progenitors differentiated in 0.5  $\mu\text{M}$  CHIR (Fig. 6O). NCAM1 was identified as a marker that increased in all neuronal induction protocols, but its increase was smallest for progenitors committed to the mDA lineage. Although it is not a true non-mDA marker, its kinetics of expression during the course of differentiation provides a useful indicator of lineage progression. For example, NCAM1 should not exceed 2 ng/ml in CM during the first 14 days of mDA induction of RC17 hESCs (Fig. 6P). NCAM1 is also sensitive to CHIR concentration at day 4 and day 7 of mDA induction with its highest expression in 1.5  $\mu\text{M}$  CHIR conditions (Fig. 6Q,R).

Our final aim was to multiplex several informative biomarkers on the Luminex platform. This would allow the quantification of numerous analytes from a single sample of conditioned medium in one assay [31]. There are only 300 Luminex-compatible antibody pairs, so not all useful biomarkers could be multiplexed. For example, high-quality ELISA assays are available for CORIN, but the antibody pairs for this analyte are not available on the Luminex platform. We custom-designed 6-plex and 8-plex Luminex assays and tested CM from a time-course of mDA induction for the 4 cell lines used for the mass spectrometry studies, and a fifth human iPSC line, NAS2, to examine the usefulness of the biomarkers for an independent cell line. The 6-plex Luminex assay combined TFF3, SERPINF1, NRP1, PDGFC, PLAU, and NCAM1. NAS2 was differentiated using the EDI mDA induction protocol and CM was collected up to day 14 for two separate differentiation attempts. TFF3 and SERPINF1 increased significantly, similar to RC17 differentiation, with the exception that TFF3 could not be detected at day 4 (Fig. 7A). NRP1 and PDGFC showed increased expression at day 14 agreeing with RC17 data that these are late markers (Fig. 7A). PLAU increased from day 0 to day 4 and then declined, which differs from the RC17 pattern where PLAU is highest at day 0 and then immediately declines (Fig. 7A). NCAM1 was undetectable in iPSC CM and increased to almost 15 ng/ml at day 7 and was highly variable (Fig. 7A). This was in contrast to RC17 mDA induction where NCAM1 did not go above 2 ng/ml. For 404C2 and 1231A3 Kyoto iPSC lines, CM was collected up to day 21 of mDA progenitor cell induction using the Kyoto protocol. TFF3 increased in a time-dependent manner as it did for NAS2 and RC17, but the levels from the 1231A3 were 10-fold lower (Fig. 7B,C). The SERPINF1 levels increased for both cell lines in a similar manner exceeding 100 ng/ml by day 11 (Fig. 7B,C). The low TFF3 levels appeared to be a unique signature of the 1231A3 cell line. NRP1 increased at day 7, decreased at day 11, and then increased again at day 21 for both iPSC lines (Fig. 7B,C). PDGFC was up-regulated at day 11 and 21, but was only detectable at day 7 for the 1231A3 cell line, a unique characteristic for this cell line since PDGFC was not detected at day 7 for RC17 or NAS2 cells (Fig. 7A,B,C; Fig. 6D). PLAU showed similar kinetics during the 404C2 and 1231A3 time-courses to NAS2 iPSCs, increasing during differentiation peaking at over 15 ng/ml at day 7 and then decreasing at later stages (Fig. 7B,C). NCAM1 increased at day 7 for both iPSC lines and maintained a similar level of expression at day 21. However, the levels were 10-fold lower in the 1231A3 mDA cultures and it was again variable (Fig. 7B,C).

For the 8-plex Luminex assay we replaced PDGFC with three positive mDA markers, IGFBP2, ALCAM, and SPP1, that all exhibited highest expression in 1.0  $\mu\text{M}$  CHIR conditions (Fig. 2A,C). We used this Luminex assay to simultaneously quantify the 8 analytes for two hESC lines (RC17, MShef7) over a time-course of mDA induction (day 0 to 14) for two independent differentiation runs per cell line. TFF3 increased significantly during mDA induction, and was detectable by day 4 for the RC17 differentiations, and by day 7 for the MShef7 attempts (Fig. 7D,E). SERPINF1 kinetics of expression was similar to the Kyoto iPSC lines with an initial increase followed by a decrease at later stages of mDA induction. The overall levels of SERPINF1 were 3-fold lower in the MShef7

cultures compared to RC17 and the Kyoto iPSC lines (Fig. 7B,C,D,E). NRP1 did not increase until day 14 for both hESC lines, similar to the results from the NAS2 iPSCs (Fig. 7A,D,E). IGFBP2 was detectable in self-renewing hESCs and increased at each time-point of the differentiation for all four differentiation runs, and the levels were similar for both RC17 and MShef7 (Fig. 7D,E). ALCAM was only detectable at day 14 of mDA induction for both hESC lines and was 4-fold lower in MShef7 cultures (Fig. 7D,E). SPP1 (also known as osteopontin) was highly expressed in hESC cultures, decreased significantly at the start of mDA induction, and then increased at later time-points (Fig. 7D,E). Although the initial SPP1 concentration was 2.5-fold higher in CM from RC17 cells, it reached a similar level of about 12–15 ng/ml at day 14 for both cell lines (Fig. 7D,E). PLAU was highest at the start of differentiation, but 20-fold lower in MShef7 hESCs, and decreased as mDA induction proceeded (Fig. 7D,E). This was in contrast to PLAU kinetics in the NAS2, 404C2, and 1231A3 iPSC lines where PLAU exhibited an initial increase in expression before decreasing at the later stages (Fig. 7). Finally, NCAM1 increased during mDA induction, but with different kinetics for RC17 and MShef7. NCAM1 increased at day 7 and 14 for MShef7 mDA cultures, but undetectable at day 4 (Fig. 7E). In contrast, NCAM1 was detected at day 4 in one of the two RC17 mDA differentiations, undetectable at day 7 for both mDA inductions, and then present at day 14, but at 5-fold lower concentrations than MShef7 CM at the same time-point (Fig. 7D,E).

## Discussion

The process of pluripotent stem cell differentiation is highly dynamic and complex, and the outcome of a protocol is usually not known until the end of a procedure and often involves invasive sampling of cells. Here we have used candidate and unbiased approaches to discover secreted biomarkers that can be quantified during mDA differentiation to provide valuable near real-time information on cell identity in a non-invasive manner.

Since different clonal hESC and iPSC lines can exhibit minor, and sometimes major, differences in propensity to response to the same lineage induction protocol, we used four genetically distinct hESC/iPSC lines subjected to mDA floor plate induction protocols in two laboratories (Edinburgh, Kyoto). We aimed to identify informative positive and negative secreted biomarkers that could be broadly applicable to the global cell replacement therapy (CRT) efforts for Parkinson's. By investigating candidate factors based on the literature, and our large-scale unbiased proteomic screen we determined that some secreted biomarkers were broadly useful for all cell lines (eg. TFF3, Corin), and that each cell line also had a unique mDA-specific secretome signature (Fig. 4).

For the common mDA biomarkers we determined the earliest time-points possible for their detection. Early quality-control indicators will be highly valuable for clinical applications, since each day of manufacturing in GMP facilities is costly. A number of the biomarkers we discovered could provide discriminating information as early as day 4 of mDA induction. TFF3, SERPINF1, and IGFBP2 increased in mDA induction cultures and SPP1 decreased at this time-point (Figs. 5,7). PLAU levels also provided useful information at day 4, but it increased for some cell lines, and decreased for others, demonstrating the unique secretome signature of different cell lines. For all cell lines PLAU was reduced to low levels at the late stages of mDA floor plate induction. The absolute amount of a given biomarker in the media is also a critical and informative indicator of differentiation progression and efficiency. For example, for RC17 hESC mDA induction using the EDI protocol, TFF3 levels should be above 2 ng/ml by day 4 and above 15 ng/ml by day 7 if the differentiation is proceeding normally. TFF3 levels below these values would indicate problems with the differentiation, and action could be taken to adjust the protocol or abort the protocol at an early time-point. Although, TFF3 is a useful biomarker for other cell lines, the absolute values at particular time-points and kinetics of up-regulation displayed unique signatures. For

example, for 404C2 iPSCs TFF3 in CM should reach 1 ng/ml by day 7 and above 5 ng/ml by day 11 using the Kyoto mDA protocol (Fig. 7).

Although single secreted biomarkers can provide early, non-invasive information on differentiation progress via commercial ELISAs, we sought to design and construct multiplexed ELISA systems to simultaneously quantify as many markers as possible from each sample of CM collected. Towards this end, we constructed 6-plex and 8-plex Luminex assays that provided multi-dimensional secreted biomarker data over a time-course of mDA neural induction for 5 pluripotent stem cell lines (Fig. 7). Further refinements of mDA-specific multiplexed assays will be possible when more positive and negative markers are validated.

The associated study by Rifes *et al.* used an unbiased mass spectrometry approach to identify secreted biomarkers specific to the caudal ventral midbrain (cVM) identity, the cells which produce mDA neurons required for Parkinson's CRT [38]. They specifically identified markers that can distinguish cVM from the rostral VM (rVM) – a region that does not produce correctly specified mDA neurons needed for CRT. Two of their three cVM-specific biomarkers, LGI1 and PDGFC, were also identified in our study. They also determined that “common” mDA markers in our study, TFF3 and Corin, had differing levels of expression in cVM versus rVM, with lower levels in the cVM [38]. This further highlights the concept that the absolute levels of the secreted biomarkers provide the most useful information. For example, TFF3 levels should increase during mDA floor plate induction, but if it exceeds a certain absolute value it might indicate the cultures are becoming more rVM in character and less cVM. Ratios of biomarkers is also a useful concept proposed by Rifes *et al.* Multiplexed assays for secreted mDA biomarkers during the time-course of differentiation will provide valuable (i) single analyte data and (ii) analyte ratio data to provide as rich information as possible during manufacturing of mDA progenitor cells for Parkinson's clinical applications.

## Conclusions

Here we have characterized the secretome of four human pluripotent stem cell lines during a time-course of differentiation into mDA progenitor cells. We have identified and validated 11 secreted biomarkers that can provide non-invasive quality-control data of mDA progenitor cell differentiation from as early as day 4 of manufacturing.

## Abbreviations

CHIR: CHIR99021; CM: conditioned media; CRT: cell replacement therapy; cVM: caudal ventral midbrain; DA: dopaminergic; DIA: data-independent acquisition; DTT: dithiothreitol; ELISA: Enzyme-Linked Immunosorbent Assay; GMP: Good Manufacturing Practice; hESC: human embryonic stem cell; IAA: Iodoacetamide; iPSC: induced pluripotent stem cell; LGE: lateral ganglionic eminence; mDA: midbrain dopaminergic; MS: Mass spectrometry; PASEF; Parallel Accumulation Serial Fragmentation; qRT-PCR: quantitative Reverse Transcription Polymerase Chain Reaction; rVM: rostral ventral midbrain; SHH: Sonic Hedgehog; TFF3: trefoil factor 3; TH: tyrosine hydroxylase; UPL: Universal probe library

## Declarations

## Ethics approval and consent to participate

Not applicable

## Consent for publication

Not applicable

## Availability of data and materials

The mass spectrometry data have been deposited to the ProteomeXchange Consortium via jPOSTrepo [27] (<https://repository.jpostdb.org/>) with the data set identifier JPST002168 (PXD042509).

URL

<https://repository.jpostdb.org/preview/2086704224648be7d7d2fb7>

Access key

7614

## Competing interests

The authors declare that they have no competing interests.

## Funding

This study was supported by AMED under Grant Number JP20bm0704054, MRC Regenerative Medicine Grant Mr/V00560x/1, MRC Confidence-in-Concept award, Wellcome Trust Institutional Translational Partnership Award (iTPA), and Cure Parkinson's project grant. The funding bodies played no role in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

## Authors' Contributions

MI, NJD, AM, and TK designed the study. MAC, NJD, DA, AM and ML conducted differentiations and provided conditioned media for ELISAs and mass spectrometry. MI conducted all mass spectrometry experiments. NJD, DT, and MAC conducted ELISAs and RT-qPCRs. DEM conducted RT-qPCRs immunostaining, and Opera imaging. NJD and DT performed Luminex assays. MI, NJD, and AM prepared figures and contributed to the writing. TK wrote the manuscript. All authors read and approved the manuscript.

## Acknowledgements

We thank K. Ogata for the technical assistance of the cell culture experiments in Japan. We thank Y. Miyamoto for the technical assistance for the proteome analysis. We thank Helen Henderson, Marilyn Thomson, and Theresa O'Connor for cell culture support in Edinburgh, Linda Ferguson for Luminex assay assistance, and Justyna Cholewa-Waclaw for imaging with the Opera Phenix Plus.

## References

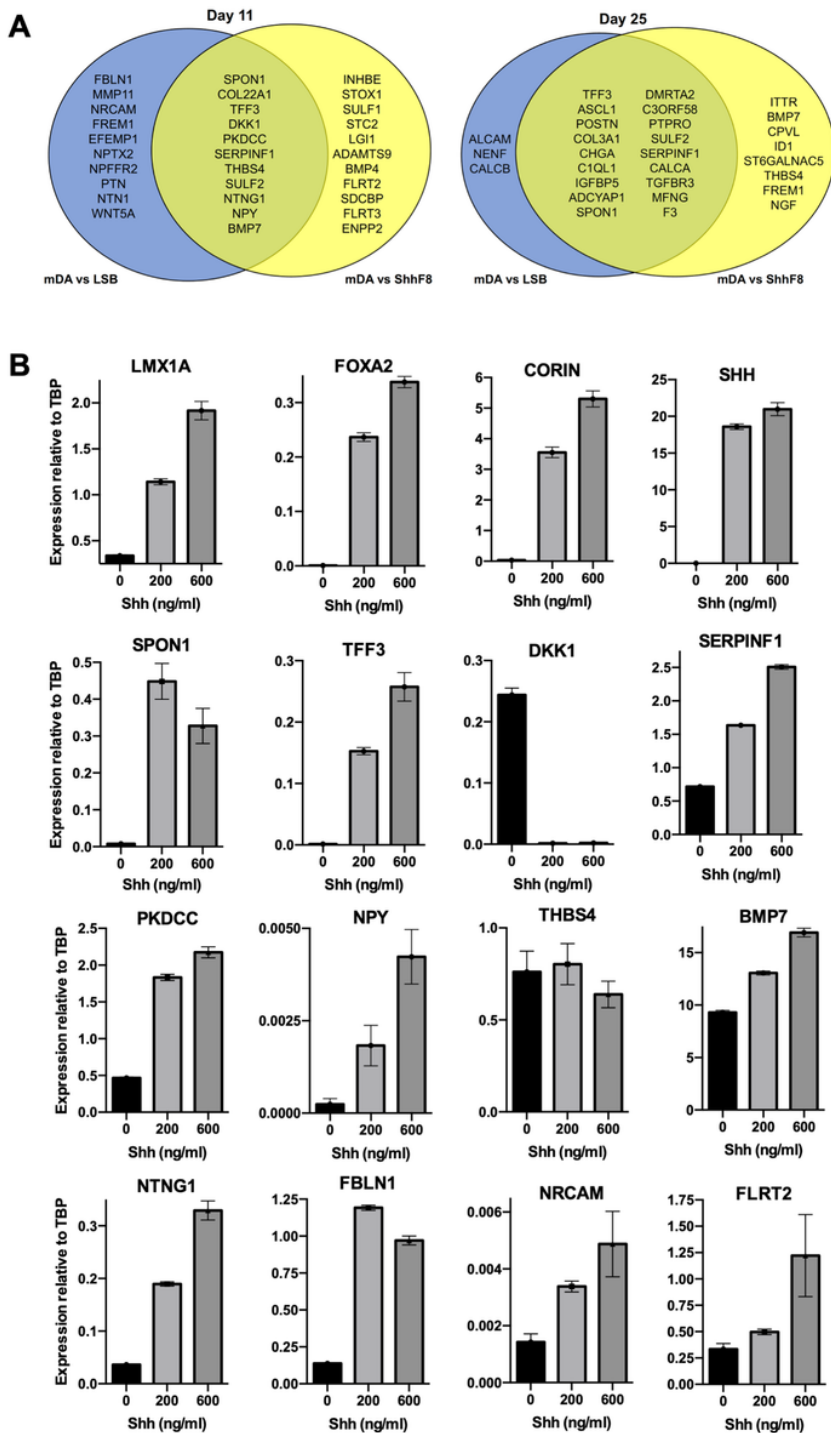
1. Barker RA, Parmar M, Studer L, Takahashi J. Human Trials of Stem Cell-Derived Dopamine Neurons for Parkinson's Disease: Dawn of a New Era. *Cell Stem Cell*. 2017;21:569–73.

2. Barker RA, Barrett J, Mason SL, Björklund A. Fetal dopaminergic transplantation trials and the future of neural grafting in Parkinson's disease. *Lancet Neurol.* 2013;12:84–91.
3. Kefalopoulou Z, Politis M, Piccini P, Mencacci N, Bhatia K, Jahanshahi M, et al. Long-term clinical outcome of fetal cell transplantation for Parkinson disease: two case reports. *JAMA Neurol.* 2014;71:83–7.
4. Ono Y, Nakatani T, Sakamoto Y, Mizuhara E, Minaki Y, Kumai M, et al. Differences in neurogenic potential in floor plate cells along an anteroposterior location: midbrain dopaminergic neurons originate from mesencephalic floor plate cells. *Development.* 2007;134:3213–25.
5. Bonilla S, Hall AC, Pinto L, Attardo A, Götz M, Huttner WB, et al. Identification of midbrain floor plate radial glia-like cells as dopaminergic progenitors. *Glia.* 2008;56:809–20.
6. Kriks S, Shim J-W, Piao J, Ganat YM, Wakeman DR, Xie Z, et al. Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. *Nature.* 2011;480:547–51.
7. Kirkeby A, Grealish S, Wolf DA, Nelander J, Wood J, Lundblad M, et al. Generation of regionally specified neural progenitors and functional neurons from human embryonic stem cells under defined conditions. *Cell Rep.* 2012;1:703–14.
8. Nolbrant S, Heuer A, Parmar M, Kirkeby A. Generation of high-purity human ventral midbrain dopaminergic progenitors for in vitro maturation and intracerebral transplantation. *Nat Protoc.* 2017;12:1962–79.
9. Grealish S, Diguët E, Kirkeby A, Mattsson B, Heuer A, Bramouille Y, et al. Human ESC-derived dopamine neurons show similar preclinical efficacy and potency to fetal neurons when grafted in a rat model of Parkinson's disease. *Cell Stem Cell.* 2014;15:653–65.
10. Kikuchi T, Morizane A, Doi D, Magotani H, Onoe H, Hayashi T, et al. Human iPS cell-derived dopaminergic neurons function in a primate Parkinson's disease model. *Nature.* 2017;548:592–6.
11. Wu H, Xu J, Pang ZP, Ge W, Kim KJ, Bianchi B, et al. Integrative genomic and functional analyses reveal neuronal subtype differentiation bias in human embryonic stem cell lines. *Proc Natl Acad Sci USA.* 2007;104:13821–6.
12. Osafune K, Caron L, Borowiak M, Martinez RJ, Fitz-Gerald CS, Sato Y, et al. Marked differences in differentiation propensity among human embryonic stem cell lines. *Nat Biotechnol.* 2008;26:313–5.
13. Hu B-Y, Weick JP, Yu J, Ma L-X, Zhang X-Q, Thomson JA, et al. Neural differentiation of human induced pluripotent stem cells follows developmental principles but with variable potency. *Proc Natl Acad Sci USA.* 2010;107:4335–40.
14. Devine MJ, Ryten M, Vodicka P, Thomson AJ, Burdon T, Houlden H, et al. Parkinson's disease induced pluripotent stem cells with triplication of the  $\alpha$ -synuclein locus. *Nat Commun.* 2011;2:440.
15. Iwashita H, Shiraki N, Sakano D, Ikegami T, Shiga M, Kume K, et al. Secreted cerberus1 as a marker for quantification of definitive endoderm differentiation of the pluripotent stem cells. *PLoS ONE.* 2013;8:e64291.
16. Tateno H, Onuma Y, Ito Y, Hiemori K, Aiki Y, Shimizu M, et al. A medium hyperglycosylated podocalyxin enables noninvasive and quantitative detection of tumorigenic human pluripotent stem cells. *Sci Rep.* 2014;4:4069.
17. Tiklová K, Björklund ÅK, Lahti L, Fiorenzano A, Nolbrant S, Gillberg L, et al. Single-cell RNA sequencing reveals midbrain dopamine neuron diversity emerging during mouse brain development. *Nat Commun.* 2019;10:581.

18. Uhlén M, Karlsson MJ, Hober A, Svensson A-S, Scheffel J, Kotol D, et al. The human secretome. *Sci Signal*. 2019;12.
19. Nakagawa M, Taniguchi Y, Senda S, Takizawa N, Ichisaka T, Asano K, et al. A novel efficient feeder-free culture system for the derivation of human induced pluripotent stem cells. *Sci Rep*. 2014;4:3594.
20. Doi D, Samata B, Katsukawa M, Kikuchi T, Morizane A, Ono Y, et al. Isolation of human induced pluripotent stem cell-derived dopaminergic progenitors by cell sorting for successful transplantation. *Stem Cell Reports*. 2014;2:337–50.
21. Arber C, Precious SV, Cambray S, Risner-Janiczek JR, Kelly C, Noakes Z, et al. Activin A directs striatal projection neuron differentiation of human pluripotent stem cells. *Development*. 2015;142:1375–86.
22. Eiraku M, Watanabe K, Matsuo-Takasaki M, Kawada M, Yonemura S, Matsumura M, et al. Self-organized formation of polarized cortical tissues from ESCs and its active manipulation by extrinsic signals. *Cell Stem Cell*. 2008;3:519–32.
23. Chambers SM, Fasano CA, Papapetrou EP, Tomishima M, Sadelain M, Studer L. Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. *Nat Biotechnol*. 2009;27:275–80.
24. Iwasaki M, Kawahara Y, Okubo C, Yamakawa T, Nakamura M, Tabata T, et al. Multi-omics approach reveals posttranscriptionally regulated genes are essential for human pluripotent stem cells. *iScience*. 2022;25:104289.
25. Meier F, Brunner A-D, Frank M, Ha A, Bludau I, Voytik E, et al. diaPASEF: parallel accumulation-serial fragmentation combined with data-independent acquisition. *Nat Methods*. 2020;17:1229–36.
26. Demichev V, Messner CB, Vernardis SI, Lilley KS, Ralser M. DIA-NN: neural networks and interference correction enable deep proteome coverage in high throughput. *Nat Methods*. 2020;17:41–4.
27. Okuda S, Watanabe Y, Moriya Y, Kawano S, Yamamoto T, Matsumoto M, et al. jPOSTrepo: an international standard data repository for proteomes. *Nucleic Acids Res*. 2017;45:D1107–11.
28. Fasano CA, Chambers SM, Lee G, Tomishima MJ, Studer L. Efficient derivation of functional floor plate tissue from human embryonic stem cells. *Cell Stem Cell*. 2010;6:336–47.
29. Morizane A, Doi D, Kikuchi T, Nishimura K, Takahashi J. Small-molecule inhibitors of bone morphogenic protein and activin/nodal signals promote highly efficient neural induction from human pluripotent stem cells. *J Neurosci Res*. 2011;89:117–26.
30. Xi J, Liu Y, Liu H, Chen H, Emborg ME, Zhang S-C. Specification of midbrain dopamine neurons from primate pluripotent stem cells. *Stem Cells*. 2012;30:1655–63.
31. Cook DB, McLucas BC, Montoya LA, Brotski CM, Das S, Miholits M, et al. Multiplexing protein and gene level measurements on a single Luminex platform. *Methods*. 2019;158:27–32.
32. Bye CR, Jönsson ME, Björklund A, Parish CL, Thompson LH. Transcriptome analysis reveals transmembrane targets on transplantable midbrain dopamine progenitors. *Proc Natl Acad Sci USA*. 2015;112:E1946-55.
33. Samata B, Doi D, Nishimura K, Kikuchi T, Watanabe A, Sakamoto Y, et al. Purification of functional human ES and iPSC-derived midbrain dopaminergic progenitors using LRTM1. *Nat Commun*. 2016;7:13097.
34. Chung CY, Seo H, Sonntag KC, Brooks A, Lin L, Isacson O. Cell type-specific gene expression of midbrain dopaminergic neurons reveals molecules involved in their vulnerability and protection. *Hum Mol Genet*. 2005;14:1709–25.

35. Toledo EM, La Manno G, Rivetti di Val Cervo P, Gyllborg D, Islam S, Villaescusa C, et al. Molecular analysis of the midbrain dopaminergic niche during neurogenesis. *BioRxiv*. 2017.
36. You Z, Wang L, He H, Wu Z, Zhang X, Xue S, et al. Mapping of clonal lineages across developmental stages in human neural differentiation. *Cell Stem Cell*. 2023;30:473-487.e9.
37. Jiang J, Wu S, Wang W, Chen S, Peng J, Zhang X, et al. Ectodomain shedding and autocleavage of the cardiac membrane protease corin. *J Biol Chem*. 2011;286:10066–72.
38. Rifés P, Isaksson M, Rusimbi C, Santoja AR, Wahlestedt JN, Laurell T, et al. Identifying secreted biomarkers of dopaminergic ventral midbrain progenitor cells. *Res Sq*. 2023.

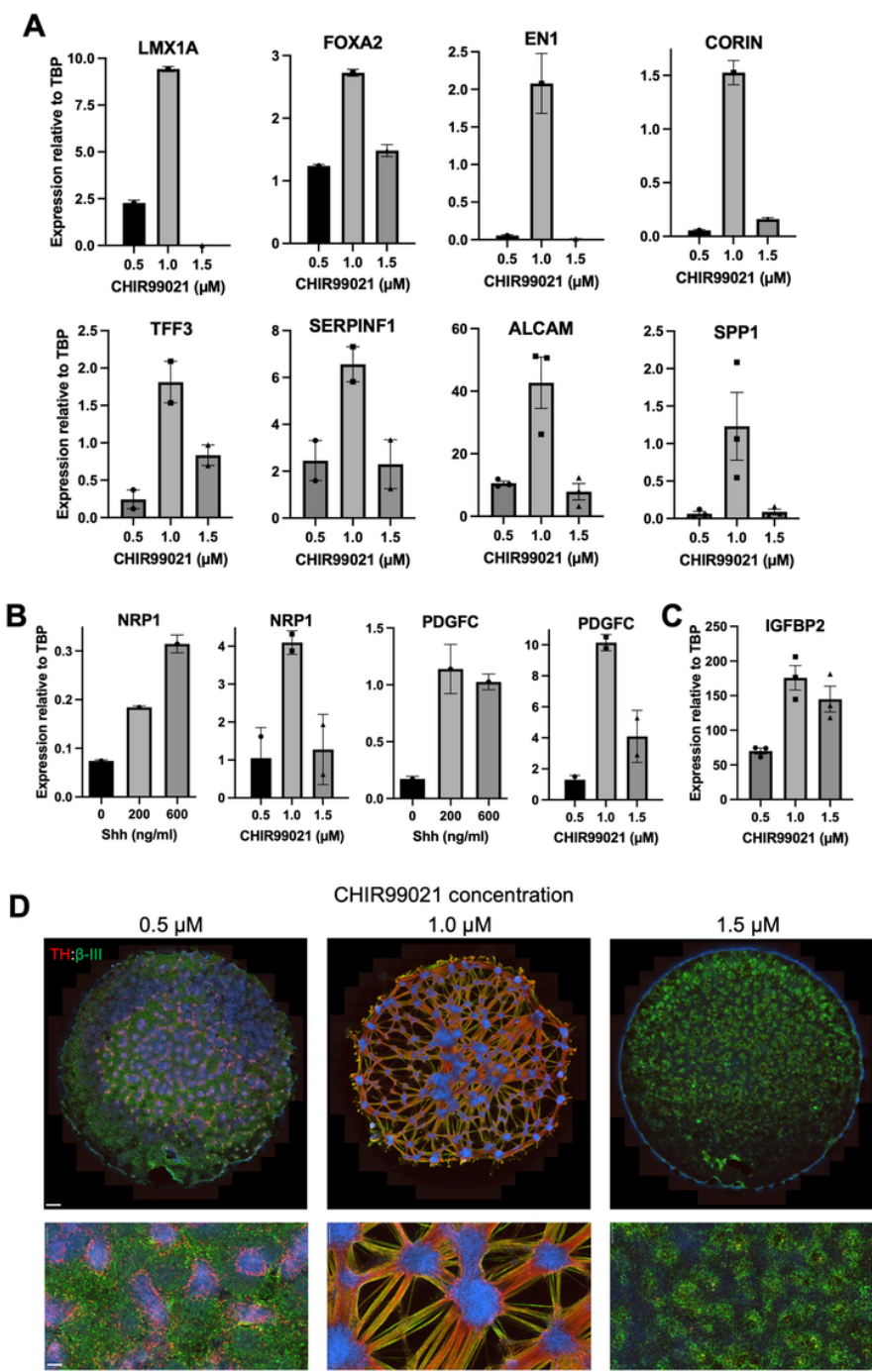
## Figures



**Figure 1**

*Candidate secreted biomarkers for human pluripotent stem cell differentiation into midbrain dopaminergic (mDA) progenitor cells and their responsiveness to SHH signaling. A.* Re-analysis of microarray data from Kriks et al., 2011. Up-regulated secreted biomarkers expressed at day 11 and day 25 of mDA progenitor cell induction comparing (i) full mDA protocol to dual-Smad inhibition (LSB), and (ii) full mDA protocol to the protocol without CHIR99021 (ShhF8). The overlap of the two gene lists for each time-point are indicated. **B.** Gene expression analysis by RT-qPCR of RNA from day 11 mDA progenitor cells differentiated from RC17 hESCs in three concentrations of SHH-C24II (0 ng/ml, 200 ng/ml, 600 ng/ml). Known ventral midbrain markers, *LMX1A*, *FOXA2*, *CORIN*, and *SHH* demonstrated a SHH-C24II-dose-dependent increase in expression. Eight (8) genes (*TFF3*,

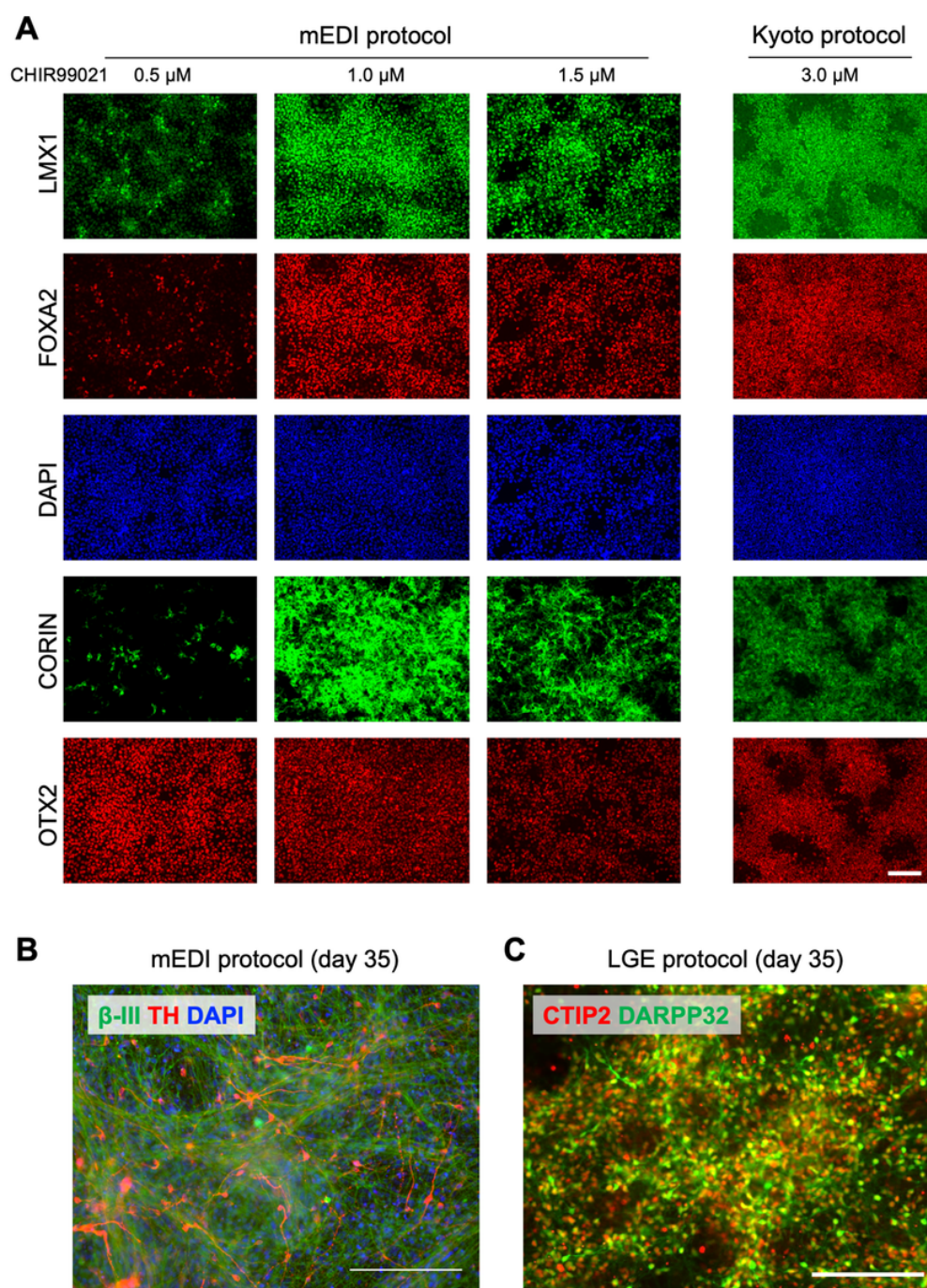
*SERPINF1*, *PKDCC*, *NPY*, *BMP7*, *NTNG1*, *NRCAM*, *FLRT2*) out of 12 investigated also exhibited a dose-dependent increase in expression in response to SHH-C24II.



**Figure 2**

*WNT* signaling responsiveness of candidate secreted biomarkers. **A.** Gene expression analysis by RT-qPCR of day 11 mDA progenitor cells differentiated from RC17 hESCs in three concentrations of CHIR99021 (0.5 μM, 1.0 μM, 1.5 μM). Known ventral midbrain markers, *LMX1A*, *FOXA2*, *EN1*, and *CORIN*, exhibited highest expression at 1.0 μM CHIR99021. Candidate secreted biomarkers TFF3, *SERPINF1*, *ALCAM*, and *SPP1* also had a peak of expression at 1.0 μM CHIR99021. **B.** Expression of Neuropilin-1 (*NRP1*) and *PDGFC* was SHH-C24II-dependent

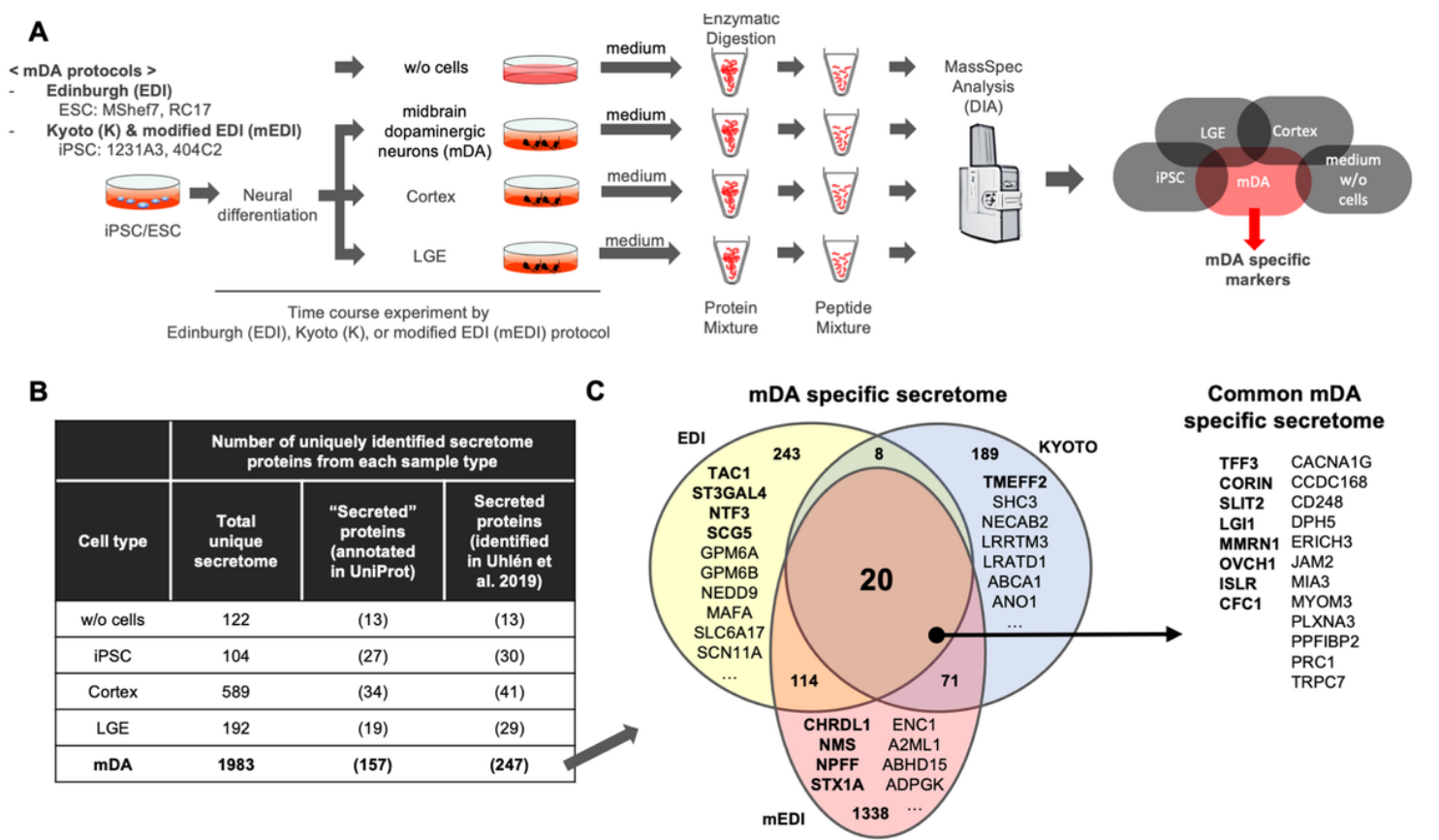
and had highest expression at 1.0  $\mu\text{M}$  CHIR99021. **C.**IGFBP2, a Luminex-compatible biomarker, also demonstrated a peak of expression at 1.0  $\mu\text{M}$  CHIR99021. **D.** Day 11 mDA progenitor cells induced in 0.5  $\mu\text{M}$ , 1.0  $\mu\text{M}$ , or 1.5  $\mu\text{M}$  CHIR99021 were differentiated to day 39 and immunostained for tyrosine hydroxylase (TH; red) and  $\beta$ -III tubulin ( $\beta$ -III; green) and entire well was imaged with the Opera Phenix. Upper panels scale bar 350  $\mu\text{m}$ , lower panels scale bar, 100  $\mu\text{m}$ .



**Figure 3**

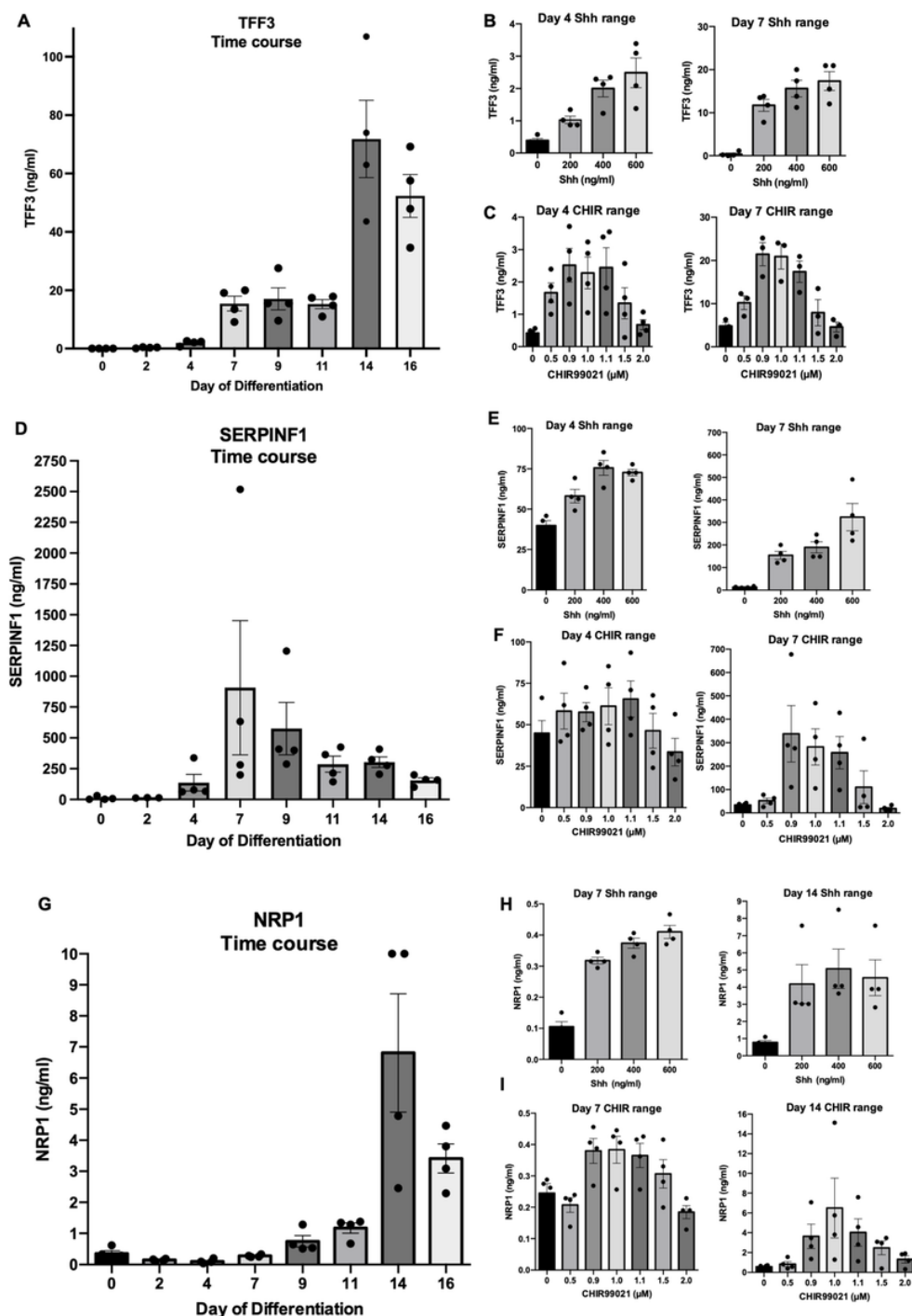
*Differentiation from Kyoto iPSCs in various neural induction protocols. A.* Midbrain floor plate induction of 1231A3 iPSCs in mEDI at three CHIR99021 concentrations (0.5  $\mu\text{M}$ , 1.0  $\mu\text{M}$ , 1.5  $\mu\text{M}$ ) and Kyoto protocol at day

14. Midbrain floor plate identity was assessed by immunostaining for LMX1, FOXA2, CORIN, and OTX2. Scale bar, 100  $\mu$ m. **B.** Prolonged differentiation in the mEDI protocol to day 35 produced neurons immune-positive for the DA marker tyrosine hydroxylase (TH) and the pan-neuronal marker  $\beta$ -III tubulin ( $\beta$ -III) Scale bar, 200 $\mu$ m. **C.** LGE neuronal induction and striatal neuron identify confirmed by DARPP32:CTIP2 immunostaining. Scale bar, 200 $\mu$ m.



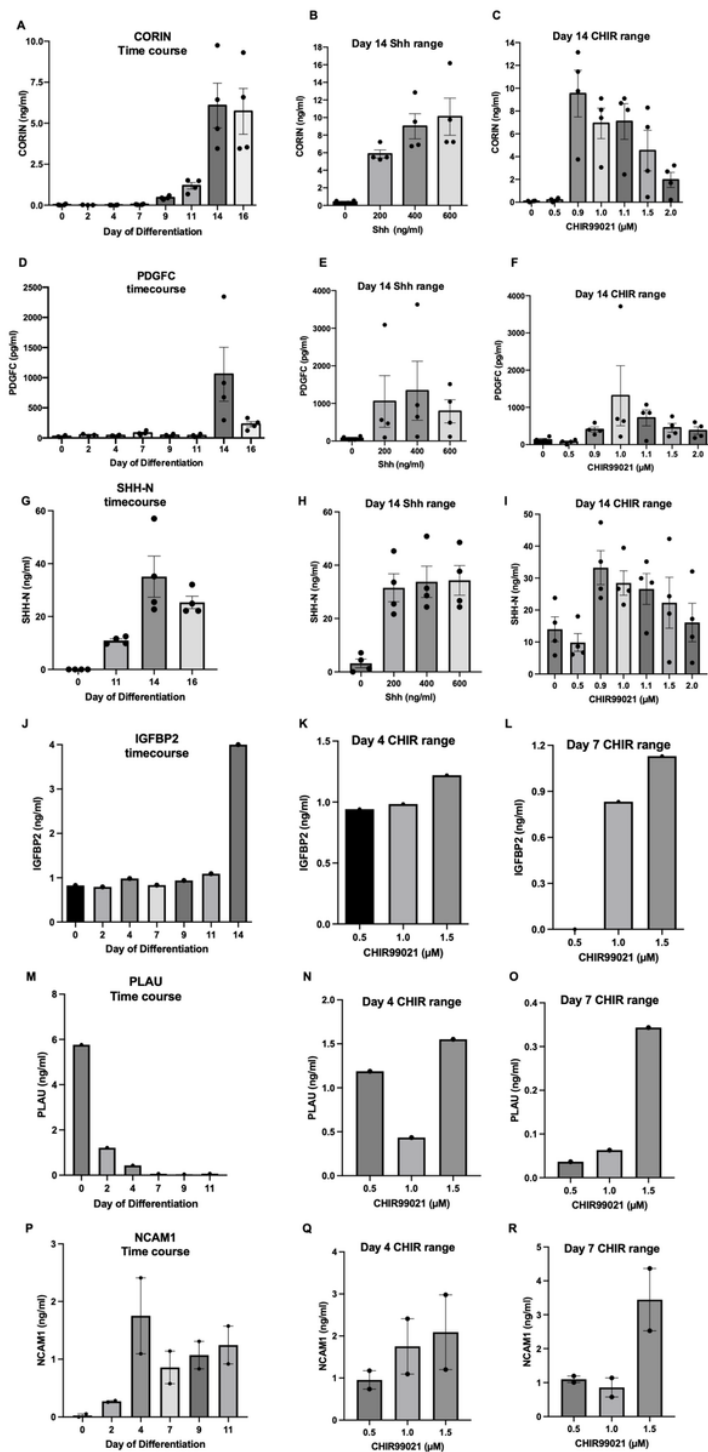
**Figure 4**

*Unbiased proteomics identified unique mDA secretome among EDI, Kyoto, mEDI mDA protocols.* **A.** A flow diagram of the unbiased proteomics used in this study. Two hESC lines from Edinburgh (RC17, MShef7) and two iPSC lines from Kyoto (404C2, 1231A3) were used. The time course differentiation experiments were performed by Edinburgh (EDI), Kyoto (K), or modified EDI (mEDI) protocols and conditioned media collected for mass spectrometry analysis. **B.** Number of uniquely identified secretome proteins from each cell type. The overlapping number of identified proteins was shown in parentheses to that of "secreted" proteins annotated in UniProt Knowledgebase (downloaded at 2023-02-20) and Uhlén et al., 2019. **C.** The Venn diagram of mDA specific secretome identified in EDI, Kyoto, mEDI mDA protocols. Gene names in bold are annotated as "secreted" in UniProt Knowledgebase.



**Figure 5**

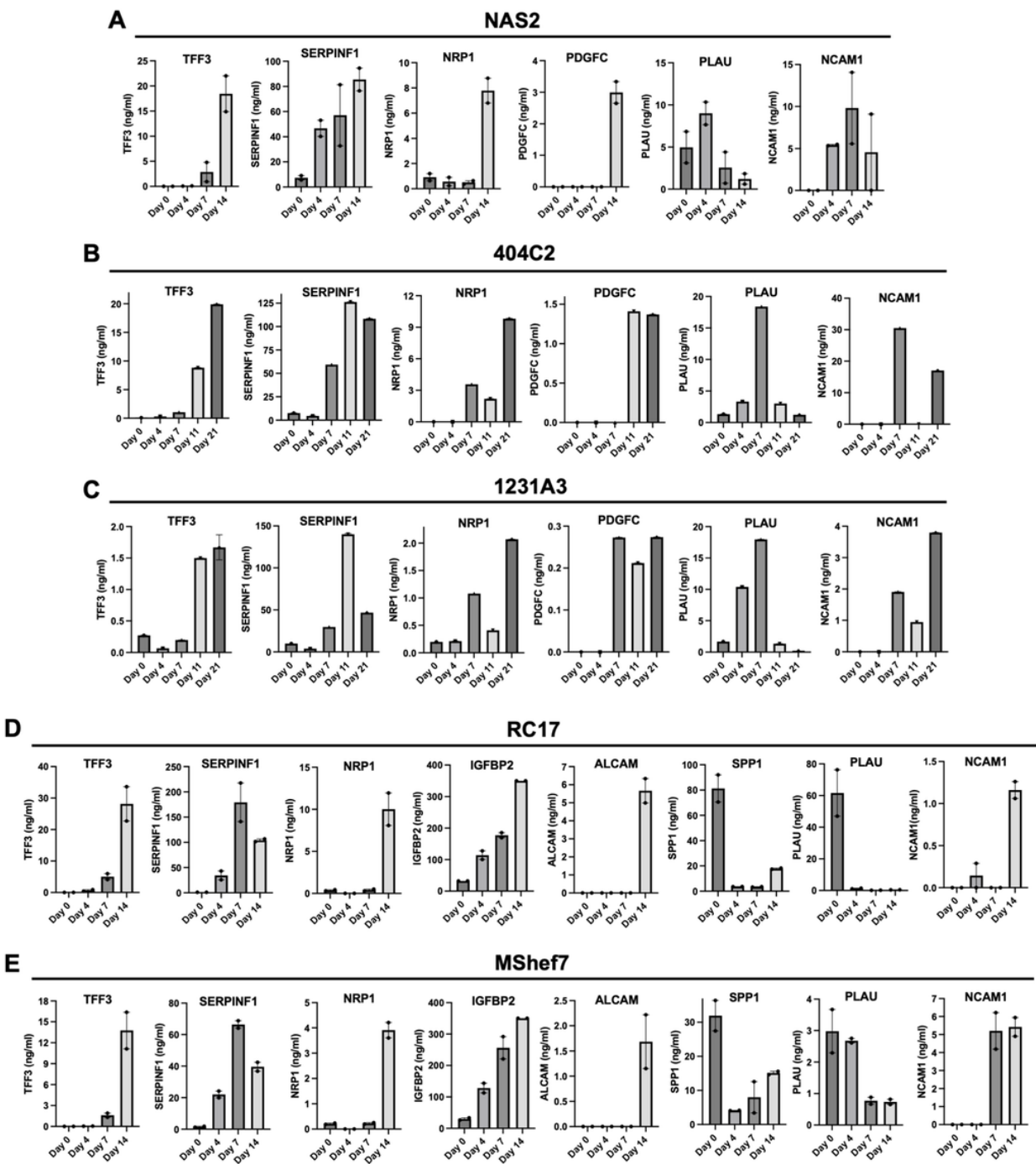
*TFF3, SERPINF1, and NRP1 ELISAs of conditioned media (CM) from mDA progenitor cell induction of RC17 hESCs. A.* Time-course of secreted TFF3 from day 0 to day 16 of mDA induction. **B.** Day 4 and day 7 TFF3 levels in CM for differing concentrations of SHH-C24II. **C.** Day 4 and day 7 TFF3 levels in CM for differing concentrations of CHIR99021. **D.** Time-course of secreted SERPINF1 from day 0 to day 16 of mDA induction. **E.** Day 4 and day 7 SERPINF1 levels in CM for differing concentrations of SHH-C24II. **F.** Day 4 and day 7 SERPINF1 levels in CM for differing concentrations of CHIR99021. **G.** Time-course of secreted NRP1 from day 0 to day 16 of mDA induction. **H.** Day 7 and day 14 NRP1 levels in CM for differing concentrations of SHH-C24II. **I.** Day 7 and day 14 NRP1 levels in CM for differing concentrations of CHIR99021.



**Figure 6**

*ELISAs of conditioned media (CM) from mDA progenitor cell induction for selected candidate secreted biomarkers. A.*Time-course of secreted CORIN from day 0 to day 16 of mDA induction. *B.*Day 14 CORIN levels in CM for differing concentrations of SHH-C24II. *C.*Day 14 CORIN levels in CM for differing concentrations of CHIR99021. *D.*Time-course of secreted PDGFC from day 0 to day 16 of mDA induction. *E.*Day 14 PDGFC levels in CM for differing concentrations of SHH-C24II. *F.*Day 14 PDGFC levels in CM for differing concentrations of CHIR99021. *G.*Time-course of secreted SHH-N from day 0 to day 16 of mDA induction. *H.*Day 14 SHH-N levels in CM for differing concentrations of SHH-C24II. *I.*Day 14 SHH-N levels in CM for differing concentrations of CHIR99021. *J.*Time-course of secreted IGFBP2 from day 0 to day 16 of mDA induction. *K,L.*Day 4 and day 7

IGFBP2 levels in CM for 0.5  $\mu$ M, 1.0  $\mu$ M, or 1.5  $\mu$ M CHIR99021. **M.**Time-course of secreted PLAU from day 0 to day 16 of mDA induction. **N,O.**Day 4 and day 7 PLAU levels in CM for 0.5  $\mu$ M, 1.0  $\mu$ M, or 1.5  $\mu$ M CHIR99021. **P.**Time-course of secreted NCAM1 from day 0 to day 16 of mDA induction. **Q,R.**Day 4 and day 7 NCAM1 levels in CM for 0.5  $\mu$ M, 1.0  $\mu$ M, or 1.5  $\mu$ M CHIR99021.



**Figure 7**

*Luminex assays (6-plex and 8-plex) of conditioned media (CM) from a time-course of mDA progenitor cell induction for five pluripotent stem cell lines. **A.** Luminex 6-plex assays (TFF3, SERPINF1, NRP1, PDGFC, PLAU, NCAM1) for NAS2 iPSCs from day 0 to day 14 of mDA progenitor cell induction. **B,C.** Luminex 6-plex assays for*

404C2 (B) and 1231A3 (C) iPSCs from day 0 to day 21 of mDA progenitor cell induction. **D,E.** Luminex 8-plex assays (TFF3, SERPINF1, NRP1, IGFBP2, ALCAM, SPP1, PLAUI, NCAM1) for RC17 (D) and MShel7 (E) hESCs from day 0 to day 14 of mDA progenitor cell induction.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryTable1Day11SecretedKriks.xlsx](#)
- [SupplementaryTable2Day25SecretedKriks.xlsx](#)
- [SupplementaryTable3SecretomeProteinList.xlsx](#)