

## RESEARCH ARTICLE

## A secreted proteomic footprint for stem cell pluripotency

Philip A. Lewis<sup>1</sup><sup>1a</sup>, Edina Silajdžić<sup>1</sup><sup>1b</sup>, Helen Smith<sup>1</sup>, Nicola Bates<sup>1</sup><sup>1</sup>, Christopher A. Smith<sup>1</sup><sup>1c</sup>, Fabrizio E. Mancini<sup>1</sup>, David Knight<sup>1</sup>, Chris Denning<sup>2</sup>, Daniel R. Brison<sup>3</sup>, Susan J. Kimber<sup>1</sup><sup>1\*</sup>

**1** Division of Cell Matrix Biology and Regenerative Medicine, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, United Kingdom, **2** Biodiscovery Institute, Division of Cancer & Stem Cells, School of Medicine, University of Nottingham, University Park, Nottingham, United Kingdom, **3** Royal Manchester Children's Hospital, Manchester, United Kingdom

<sup>1a</sup> Current address: Proteomics Facility, Biomedical Sciences, University of Bristol, Bristol, United Kingdom

<sup>1b</sup> Current address: Department of Psychosis Studies, Institute of Psychiatry Psychology & Neuroscience, Kings College London, London, United Kingdom

<sup>1c</sup> Current address: Medicines Discovery Catapult, Alderley Park, Macclesfield, United Kingdom

\* [Sue.Kimber@manchester.ac.uk](mailto:Sue.Kimber@manchester.ac.uk)

 OPEN ACCESS

**Citation:** Lewis PA, Silajdžić E, Smith H, Bates N, Smith CA, Mancini FE, et al. (2024) A secreted proteomic footprint for stem cell pluripotency. PLoS ONE 19(6): e0299365. <https://doi.org/10.1371/journal.pone.0299365>

**Editor:** Shinsuke Yuasa, Okayama University, Okayama Daigaku, JAPAN

**Received:** May 31, 2023

**Accepted:** February 8, 2024

**Published:** June 14, 2024

**Copyright:** © 2024 Lewis et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Data are available via ProteomeXchange with identifiers PXD045347 and PXD045525.

**Funding:** This work was funded by MRC grant MR/M017344 and an MRC Confidence in Concept (CIC) grant to SJK. Funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. <https://www.ukri.org/councils/mrc/>.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

With a view to developing a much-needed non-invasive method for monitoring the healthy pluripotent state of human stem cells in culture, we undertook proteomic analysis of the waste medium from cultured embryonic (Man-13) and induced (Rebl.PAT) human pluripotent stem cells (hPSCs). Cells were grown in E8 medium to maintain pluripotency, and then transferred to FGF2 and TGF $\beta$  deficient E6 media for 48 hours to replicate an early, undirected dissolution of pluripotency. We identified a distinct proteomic footprint associated with early loss of pluripotency in both hPSC lines, and a strong correlation with changes in the transcriptome. We demonstrate that multiplexing of four E8- against four E6- enriched secretome biomarkers provides a robust, diagnostic metric for the pluripotent state. These biomarkers were further confirmed by Western blotting which demonstrated consistent correlation with the pluripotent state across cell lines, and in response to a recovery assay.

## Introduction

Although the clinical potential of regenerative medicine, including approaches using human pluripotent stem cells (hPSCs), remains unrivalled, major hurdles exist in translating promising protocols for treatments into scalable, reproducible commercial processes. Maintaining hPSCs in their pluripotent state has proved challenging and labour intensive. Advances have been made in automating the culture and expansion of high quality hPSCs, such as by the development of the SelecT (TAP biosystems—Hertfordshire, UK), Freedom EVO (TECAN Trading AG, Switzerland) systems, or technologies developed by Tokyo Electron (Kyoto, Japan). However, a key issue in hPSC culture that remains to be overcome is the absence of a rapid, reproducible, non-invasive, and quantitative metric for pluripotent stem cells.

All of the most commonly used techniques have major drawbacks when it comes to their use in hPSC production. For instance, using immunofluorescence for assessing the core pluripotency transcription factors and cell surface markers does not accurately pick-up initial loss of pluripotency in entire cultures, but rather relatively late loss of pluripotency in individual cells. Immunofluorescence is also imprecise, generally non-quantitative, cannot easily be integrated into the manufacturing process, and the tested cells are lost from culture in the assay. Using flow cytometry for identifying loss of pluripotency markers is helpful but shows profound run variability and again sacrifices cell product. At present the most quantifiable metrics of pluripotency available are RNA-based assays on cells, such as Pluri-test [1] and ScoreCard [2], or teratoma assays [3]. However, these metrics are costly and require sacrifice of pluripotent cell product. Moreover, they also take too long to provide results for 'at-risk' cultures, which would likely become irreversibly compromised by the time the results are available. The most rapid and cost-effective tool researchers currently have is analysis of cell morphology, however this is inherently subjective. While attempts have been made to automate and quantify the characteristics of hPSC morphology [4, 5], it remains a low-resolution way to measure pluripotency [6]. Thus, there are no fit for purpose quantitative assays which can rapidly pick up very early loss of pluripotency.

To address these limitations, we sought to identify secreted protein biomarkers in conditioned culture medium within 48 hours of the onset of hPSC pluripotency dissolution. Conditioned medium would be particularly suitable for inline monitoring in a commercial setting such as for therapeutic manufacture, allowing regular, rapid, repeated media sampling for MS or other evaluation without disturbing the growth of the culture. Although the secretome of hPSCs has been studied in the past [7–11], there is yet to be a comparative study aiming to identify proteins in medium indicative of the pluripotent state, and its loss. It is well established that hPSC self-renewal relies on the signalling pathways down stream of FGF-2 and TGF $\beta$  family members [12–14]. After removal of these growth factors from standard hPSC culture medium the cells lose their core pluripotency network, a prerequisite for lineage differentiation [15]. Therefore, we used the removal of FGF-2 and TGF $\beta$  from the medium to trigger loss of the pluripotent state, aiming to identify protein biomarkers which could be detected in conditioned medium while it is still possible to rescue the stem cell state. Using LC-MS/MS (Liquid Chromatography with tandem Mass Spectrometry) of conditioned medium from two pluripotent stem cell lines (one hESC, one hiPSC), we identified 117 secreted proteins which were highly significantly changed in their abundance ( $q < 0.01$ ) within 48 hours of E8 medium being substituted for the FGF-2/TGF $\beta$  deficient E6. Of these 74 were identified as lower in E6 medium than in E8, and 38 higher.

Additionally, we address the fundamental limitation in the use of secreted proteins in media as biomarkers, that the concentration of such proteins will vary greatly dependent upon medium volume, cell concentration, timing of passage and cell health, in a way which cannot routinely be normalised for. As such we identified a panel of biomarkers which were highly significantly changed in both cell lines ( $q < 0.05$ ) in secreted medium and, for all but one marker, also at the RNA level. The relative abundances of these biomarkers when multiplexed provide a more robust metric for this change than the individual protein abundances. We propose that the use of protein abundance ratios provides a more reliable metric for pluripotency loss that is not affected by changes in overall protein concentration in the media caused by cell seeding and variability in cell viability and health, or applied medium volume. We further demonstrate the utility of these markers by returning E6 cultures to E8 medium, leading to a full recovery of pluripotency. Such indicators in the conditioned medium are amenable to simple detection assays and will be particularly useful in a scale up culture setting to prevent or reverse culture deterioration while avoiding loss of product.

## Results

### Incipient pluripotency loss after 48 hours of FGF2/TGF $\beta$ removal is at the limits of detection by flow cytometry and immunocytochemistry

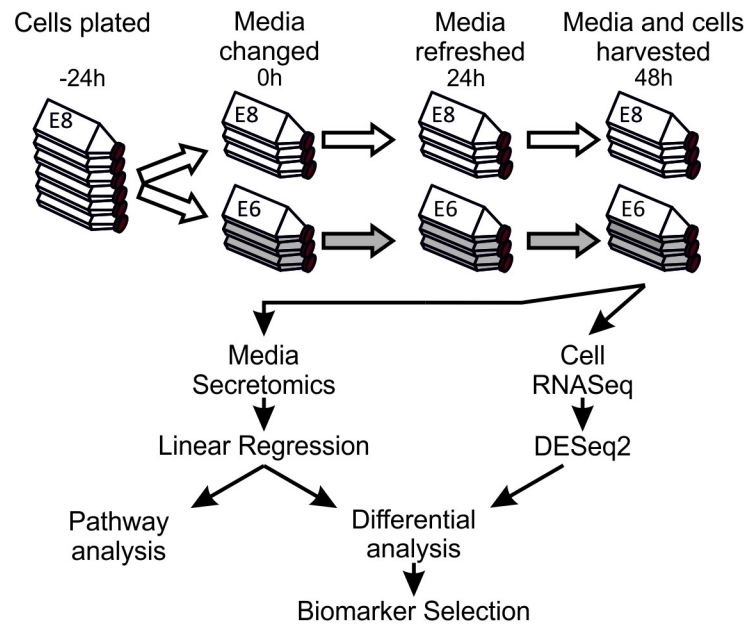
We first confirmed the time scale of loss of conventional human pluripotency-associated markers. HESC (Man-13: [16]) and hiPSC (Rebl.PAT: [17]) lines (collectively referred to as hPSCs) were cultured in either pluripotency maintenance medium (E8 [18]), or identical medium lacking the pluripotency maintenance factors FGF2 and TGF $\beta$ 1 (E6 [19]) for 48 hours, before analysis (Fig 1). A previous study indicated that mRNA for lineage specific marker proteins is not definitively expressed until at least four days in E6. However, loss of pluripotency-associated markers was detectable after as little as 48 hours following removal of FGF2/TGF $\beta$  [15]. Therefore, it was hypothesised that it should be possible to identify secretome changes associated with the loss of pluripotency at this early time point, enabling cultures to be identified at a point at which they can be rescued. In this manuscript we use the word secretome to refer to all proteins identified in the conditioned media in both cell lines, which may include proteins shed through means beyond canonical cell secretion pathways.

In E8 or 48 hours after FGF2/TGF $\beta$  removal (in E6), MAN13 cells exhibited broadly comparable immunofluorescence and staining for pluripotency markers OCT4 and SSEA3. In Man-13 cells cultured in E6 for 48 hours there was a slight decrease in immunofluorescence for NANOG and TRA160, and appearance of the early differentiation marker, SSEA1 (S1A and S1B Fig). Flow cytometric data shows similar NANOG distribution in both Man-13 and Rebl.PAT cells grown for 48 hours in E6 compared to E8 medium, whereas TRA1-81 marginally decreased in Rebl.PAT cells and marginally increased in Man-13 cells (S1B Fig). Only SOX2 was consistently observed to be moderately increased after 48 hours in E6 (S1A and S1B Fig) confirming transcript data (S1C Fig). At 48hours, OCT4 was observed to be slightly increased, as is common on initial differentiation, at the transcript level in Rebl.PAT and by flow cytometry in Man-13 (S1B and S1C Fig). Additionally, no significant change in viability of cells between E8 and E6 was observed (S2 Fig).

Although it is well established that feeder-free pluripotent stem cell maintenance requires FGF2 and TGF $\beta$ , without which cells lose pluripotency, our results indicate that current standard methods cannot unambiguously detect the incipient loss of pluripotency 48h after growth factor withdrawal. Whilst routine flow cytometry, RT-PCR and immunofluorescence can identify some subtle changes in pluripotency-associated markers at this time, the changes are only distinguishable when comparing these samples with high quality, stable pluripotent controls cultured in parallel. In isolation these changes would be unreliable if used to identify incipient pluripotency loss.

### Consistent, detectable changes in the secretome are observable after 48 hours of FGF2/TGF $\beta$ removal

Medium conditioned by stem cells cultured in either E8 or E6 was collected according to the workflow shown in Fig 1, after 24 and 48 hours. Briefly, equal volumes of media were spun to remove any floating cells and debris, then concentrated in 10kDa centrifugal filter units, before 50ug (Man-13) or 25ug (Rebl.PAT) of protein from each sample was prepared for label-free LC-MS/MS using by Filter-Assisted Sample Preparation (FASP). Acquired MS data for each MS experiment were analysed in Progenesis LC-MS and peak lists were searched against Swissprot/Trembl database using Mascot v2.5.1 to give protein abundance measurements. In total we analysed 29 Rebl.PAT samples (5 batches of 3 biological replicate flasks for E8 & E6, 1 QC failure removed) and 24 Man-13 samples (4 batches of 3 biological replicate flasks for E8 & E6,



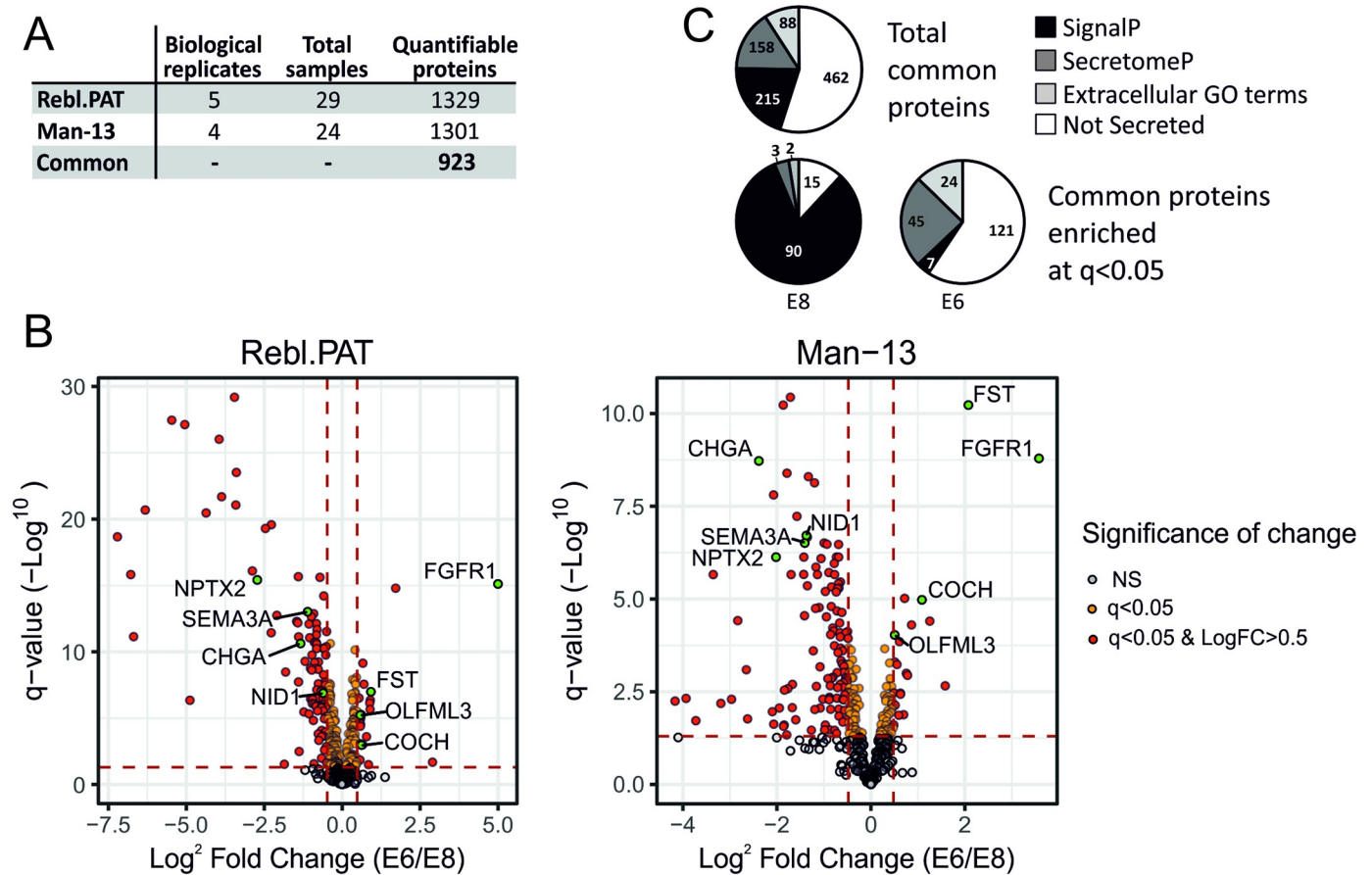
**Fig 1. Experimental design for sample collection and data processing.** Cells were cultured in E8 for at least two passages prior to the onset of the experiment. Cells were then plated in E8 medium containing Rock inhibitor onto Vitronectin-N coated flasks and were allowed to settle for 24 hours before onset of experimental conditions. Cells were washed in PBS, then half were cultured in E8, and half in E6, with full medium replacement after 24 hours. After 48 hours of culture under experimental conditions the medium was collected, spun, concentrated, and processed for LC-MS/MS. Cells were also collected and processed for quality control, and RNA-Seq analysis.

<https://doi.org/10.1371/journal.pone.0299365.g001>

1 QC failure removed) by label-free LC-MS/MS, identifying a total of 923 common proteins detected at quantifiable levels (a Progenesis confidence score of  $>30$ ) in both cell lines (Fig 2A). Principal Component Analysis (PCA) of the proteomics data correctly separated the samples by culture medium type. However, it also identified the cell lines, and the MS batch effects as strong sources of variance (S3 Fig). As such differential analysis of the protein abundance between E8 and E6 samples was performed separately for each cell line using linear regression models, fitted independently for each protein using protein abundance as the dependent variable, and medium as the independent variable, with biological replicate included as a covariate. A common trend in both cell lines was that although protein abundance was normalised to total protein amount, a larger number of proteins were significantly enriched ( $q < 0.05$ ) in E8 than in E6 (Fig 2B). We also identified that 84% of E8-enriched proteins ( $q < 0.05$ ) were secreted, as identified by the presence of an amino-acid secretion tag (identification through SignalP [20]), or by an amino-acid sequence indicative of non-classical secretion (identification through SecretomeP [21]). By contrast, a method of secretion could not be identified for 64% of E6 enriched proteins, including by presence of annotation using the cellular component gene ontology (GO) tags; extracellular space, extracellular vesicle, cell surface, or extracellular region (Fig 2C).

### Correlation between changes in protein and RNA abundance

RNASeq was performed on cells from a selection of the replicates to investigate whether changes in secretome correlated with changes at the transcript level (Fig 3). We observed a significant positive correlation between the LogFCs of significantly changed ( $q < 0.05$ ) secreted proteins and RNA transcripts (M13  $R^2 = 0.44$ ,  $p < 0.001$ ; Rebl.PAT  $R^2 = 0.54$ ,  $p < 0.001$ ).



<https://doi.org/10.1371/journal.pone.0299365.g002>

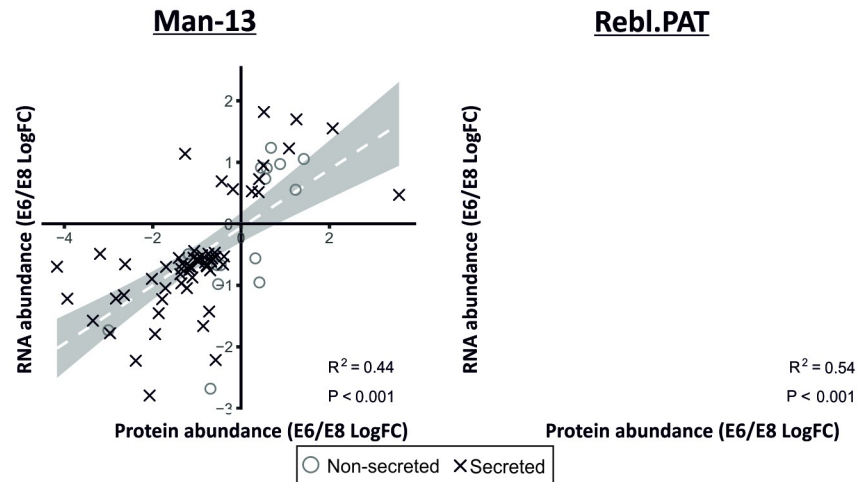
Further investigation into the minority of proteins which correlated negatively identified that these tended not to be identified as secreted.

### Pathway analysis

Secretomic data were uploaded to Ingenuity Pathway Analysis (IPA) for pathway analysis of significantly changed proteins ( $q < 0.01$ ,  $> 1.4\text{FC}$ ). Though in-depth pathway analysis on secreted proteomics data is limited in its utility as large portions of most signalling pathways are intracellular, and as such not detectable in this assay, IPA analysis identified elements of the Human Embryonic Stem Cell Pluripotency pathway were significantly enriched in both cell lines (Fig 4).

### Selection of biomarker panel

As human PSC culture media contain high concentrations of serum proteins such as Transferin, or Serum Albumin, the proportion of proteins in conditioned medium that originates from cells can vary dramatically dependent upon variances in culture volume, evaporation,



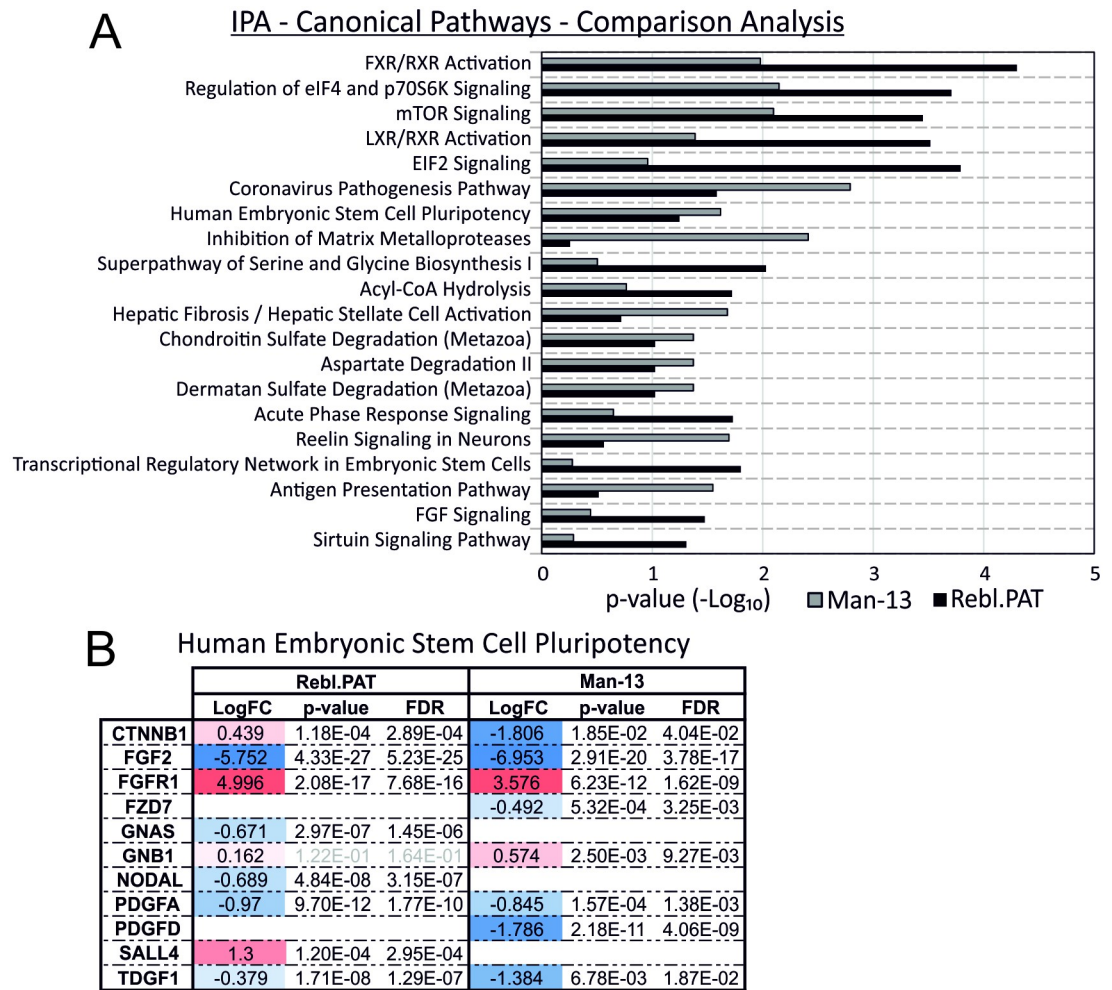
**Fig 3. Correlation between E6/E8 log fold-changes between mRNA and protein data.** Protein data was paired to RNA-Seq data using Entrez IDs retrieved from Uniprot and Bioconductor [22–24]. These data show a correlation between the changes observed in significantly changed ( $q < 0.05$ ) secreted proteins and the corresponding RNA-seq data (Man-13  $R^2 = 0.44$  & Rebl.PAT  $R^2 = 0.54$ ).

<https://doi.org/10.1371/journal.pone.0299365.g003>

seeding density, starting viability, or growth rate. For the presented assays we aimed to limit the effect of these variables, however we acknowledge that they present a limitation to the accurate relative quantification of secreted protein abundance by affecting normalisation to total protein abundance. To overcome this, we sought to identify a biomarker panel which would allow for discrimination between sample types in a robust fashion and which could be quantified without having to normalise to total protein quantity.

As biomarkers which both increase and decrease during loss of pluripotency are responding to the same stimulus but in different directions, we anticipate that the change in their relative abundance upon loss of pluripotency should be consistent, regardless of variance in total protein concentration. Therefore, we sought to identify a marker set which could be internally normalised against each other rather than relying on volume or total protein abundance normalisation.

We selected a list of secreted proteins which demonstrated highly consistent changes in both RNA and protein abundance between conditions, and tested whether the relative abundances of these proteins formed ratios which were consistent and discriminatory between E8 and E6 conditions, thus, they should be robust irrespective of the normalisation challenges of conditioned medium (Table 1). From this list, the relative abundance of each E8-enriched protein compared to each E6-enriched protein was calculated and the resultant ratios were filtered for those which demonstrated the most statistical separation between culture media ( $p < 1E-8$  &  $> 1$  LogFC). This resulted in a panel of four proteins which increased in medium when cells were cultured in E6 compared with E8—Cochlin (COCH), FGF Receptor-1 (FGFR1), Follistatin (FST) and Olfactomedin-like 3 (OLFML3)—, and four proteins which were higher when cells were cultured in E8 compared E6—Chromogranin A (CHGA), Nidogen 1 (NID1), Neuronal Pentraxin-2 (NPTX2) and Semaphorin 3A (SEMA3A). The relative abundances of these proteins provide an internally normalised set of ratios which each discriminate between E8 and E6 medium after 48 hours, in a manner which avoids the variables which thwart quantification of conditioned medium biomarkers (Fig 5).



**Fig 4. Canonical Pathways identified by Ingenuity Pathway Analysis (IPA) of secreted proteomics data.** A) Independent pathway analyses were performed for Man-13 and Rebl.PAT comparing proteins with a  $q < 0.05$  change between E6/E8 to total secretome of each cell line. A comparison analysis is shown of the top canonical pathways identified as significantly enriched in both cell lines. B) Individual proteins identified as changed in the Human Embryonic Stem Cell Pluripotency pathway for both Man-13 and Rebl.PAT.

<https://doi.org/10.1371/journal.pone.0299365.g004>

### Confirmation of marker proteins by Western blotting

Based upon the strength of these markers in the LC-MS/MS data, equal quantities of protein from conditioned media samples (collected as described in Fig 1A) were probed by Western blot for the relative abundances of these marker proteins where suitable antibodies were available (Fig 6). In the absence of a conditioned medium loading control, each E8-marker protein was probed for on the same membrane with a corresponding E6-marker protein to confirm the utility and validity of the internal normalisation method. The Western blot data was consistent with MS results: NID1, SEMA3A, CHGA and NPTX2 bands were stronger in E8 than E6 while COCH, FGFR1, FST and OLFML3 bands were stronger in E6 medium.

### Recovery of protein-pair ratio abundances after 48 hours in E6 medium

An important component of an assay for loss of pluripotency is the ability to detect a degradation of culture quality while the cells are recoverable. Therefore, we repeated the experiment,

Table 1. Secreted proteins significantly changed in both RNA and protein abundances in both Rebl.PAT and Man-13 cell lines.

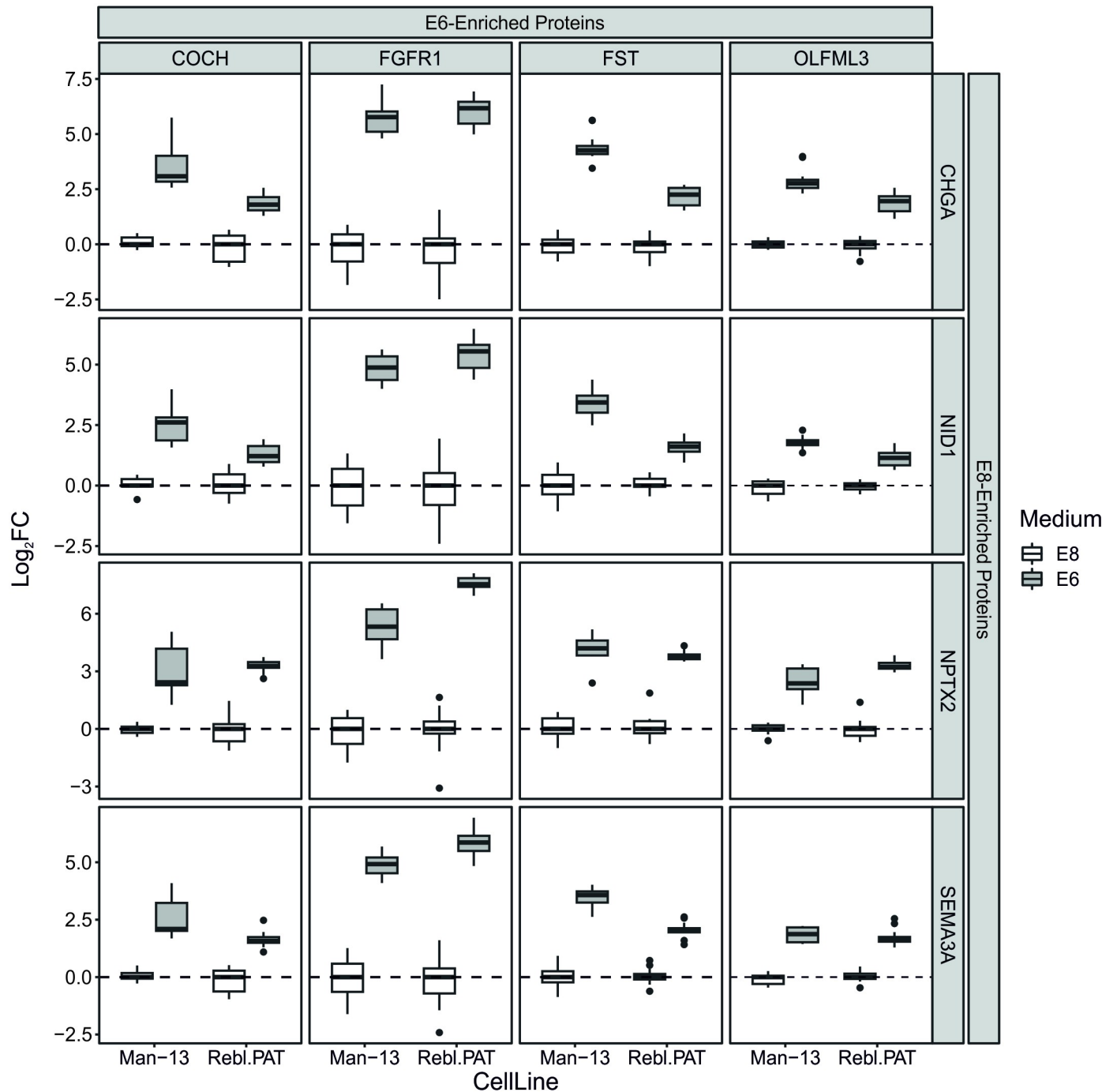
		LogFC (E6/E8)				Adj. P-value			
		Protein		RNA		Protein		RNA	
		Man-13	Rebl. PAT	Man-13	Rebl. PAT	Man-13	Rebl.PAT	Man-13	Rebl.PAT
E6-Enriched	COCH	1.09	0.62	1.23	0.54	1.05E-05	1.06E-03	1.02E-10	5.17E-06
	FGFR1	3.58	5.00	0.47	0.24	1.62E-09	7.68E-16	2.57E-02	8.59E-02
	FST	2.07	0.92	1.55	1.04	5.91E-11	1.03E-07	1.50E-18	9.79E-24
	OLFML3	0.51	0.58	0.95	0.76	9.31E-05	5.93E-06	4.25E-07	1.47E-03
	SERPINB9	1.25	1.71	1.70	1.89	3.95E-05	1.61E-15	1.91E-07	1.90E-22
E8-Enriched	ADAMTS8	-3.36	-0.42	-1.57	-0.66	2.19E-06	2.23E-03	1.65E-20	2.14E-06
	APP	-1.00	-0.71	-0.57	-0.43	3.09E-07	2.43E-16	2.73E-03	7.73E-06
	B3GNT7	-1.17	-0.81	-0.75	-0.62	1.38E-06	2.09E-12	1.20E-03	7.72E-09
	CHGA	-2.38	-1.34	-2.23	-1.37	1.89E-09	2.34E-11	7.75E-31	2.50E-21
	COL5A2	-2.65	-0.89	-1.16	-0.80	8.07E-04	8.52E-07	1.95E-09	8.06E-16
	EFEMP1	-2.83	-1.07	-1.22	-0.67	3.83E-05	8.08E-13	9.72E-10	3.72E-12
	EPHA1	-1.33	-0.61	-0.65	-0.69	5.04E-09	5.49E-11	2.32E-04	1.96E-16
	ERVMER34-1	-1.71	-0.89	-1.05	-0.53	3.64E-11	1.93E-11	7.13E-08	2.55E-05
	FBLN2	-1.14	-0.53	-0.70	-0.55	2.19E-06	5.11E-07	4.20E-05	1.77E-08
	IGFBP2	-1.08	-0.60	-0.55	-0.49	8.69E-03	5.38E-09	3.45E-03	1.58E-04
	IGFBP4	-1.70	-0.67	-0.70	-0.30	2.19E-06	8.01E-08	9.67E-05	4.30E-03
	NID1	-1.37	-0.61	-0.82	-0.53	1.95E-07	1.29E-07	1.08E-04	1.49E-05
	NPTX2	-2.02	-2.73	-0.90	-0.93	7.42E-07	3.82E-16	1.85E-07	3.48E-18
	NTS	-2.07	-1.81	-2.79	-1.60	1.56E-08	3.35E-09	3.18E-48	3.46E-20
	PLTP	-0.58	-0.52	-0.54	-0.43	9.59E-04	1.44E-06	4.07E-03	1.59E-06
	PODXL2	-3.19	-1.38	-0.49	-0.40	6.52E-03	3.25E-03	1.64E-02	6.06E-04
	SEMA3A	-1.41	-1.11	-0.56	-0.47	3.09E-07	9.58E-14	1.11E-02	2.25E-03
	SEMA3F	-1.06	-0.60	-0.44	-0.47	8.11E-07	9.54E-07	3.14E-02	7.86E-06
	SFRP2	-1.87	-0.84	-1.45	-0.81	5.91E-11	1.79E-09	3.85E-16	4.84E-21
	STC2	-1.23	-0.79	-1.04	-0.76	2.50E-04	2.84E-06	2.65E-06	5.19E-14
	THY1	-0.86	-0.83	-1.66	-1.46	9.11E-05	1.99E-11	1.39E-28	3.02E-50
	TNFRSF8	-4.17	-0.83	-0.70	-0.52	5.65E-03	7.47E-13	2.69E-05	8.61E-06
WFDC2	-1.35	-0.83	-0.97	-0.84	4.40E-06	5.29E-12	1.09E-08	2.59E-12	

<https://doi.org/10.1371/journal.pone.0299365.t001>

but replated cells into E8 medium for 7 days after the 2 days in E6. In cell lines WIGW2 (Fig 7) and Man1 (S4 Fig), we show that the change in the E6/E8 marker protein abundance ratio caused by 48 hours in E6 is reversed after 6 days of culture in E8, both in WIGW2 hiPSCs and Man-1 hESCs, using FST/NID1 and OLFML3/NID1 marker pairs respectively.

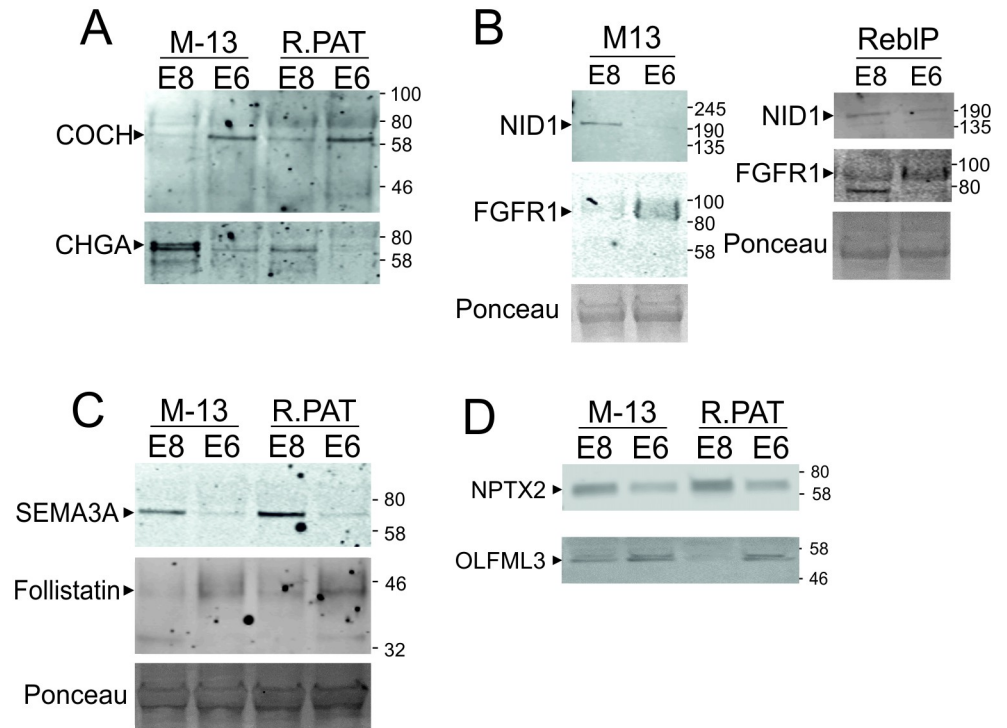
Moreover, cells cultured for 48 hours in E6 and recovered for 7 days in E8 media can still form teratomas (S5 Fig) demonstrating differentiation to tissues of all three germ layers. This suggests not only that the pluripotency of these cells can be fully recovered after 48h in E6, but also that our marker proteins can robustly track this recovery. That the changes we observe are not simply molecular markers of the immediate loss of growth factor is supported by the fact that the proteins take several days to return to their baseline, rather than returning within 12–24 after the FGF2 and TGFβ1 are restored. Thus, at the time early loss of pluripotency is recognised by these early marker changes, cultures are still plastic and can recover.





**Fig 5. Box plots demonstrating efficacy of multiplexing marker proteins.** For each boxplot, the LogFCs of two protein abundances were calculated independently for each sample. These LogFCs are calculated from two marker proteins, one identified as enriched in E6 (top) over one identified as enriched in E8 (right). By normalising the protein abundances against each other in this way we are able to demonstrate the consistent changes in the relative abundance of these proteins in all combinations of markers, and all samples across both cell lines. For ease of comparison across different protein combinations, these LogFCs are scaled such that the median LogFC in samples grown in E8 media is 0. For every marker protein combination, the difference between mean ratios for E8 and E6 media was below  $p < 1E-8$  &  $> 1$  LogFC.

<https://doi.org/10.1371/journal.pone.0299365.g005>



**Fig 6. Western blot confirmation of marker proteins.** Equal quantities of protein from media supernatant from cells cultured as described in Fig 1 were run on Western blot and probed with antibodies for an E8-enriched marker protein (CHGA, NID1, SEMA3 or NPTX3) and an E6-enriched marker protein (COCH, FGFR1, Follistatin or OLFML3). Membranes were imaged using LICOR Odyssey system.

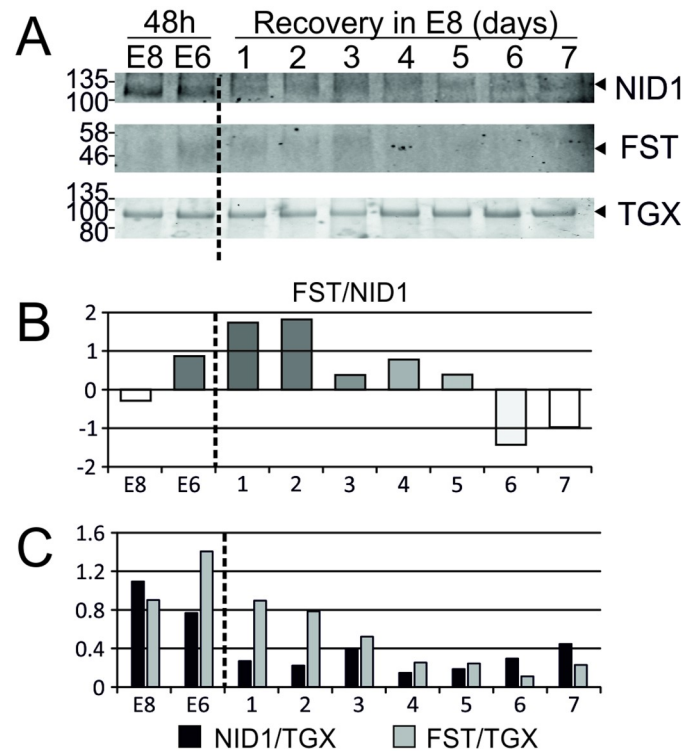
<https://doi.org/10.1371/journal.pone.0299365.g006>

### Assessment in TeSR1 and mesodermal differentiation medium

In order to further validate that some of these secretome markers could be predictive with cells cultured in other commonly used media we grew Man13 in the pluripotent stem cell culture medium, TeSR1, (S6 Fig) and chondrogenic differentiation medium [25]. TeSR1 medium has high HSA which interferes with Western blotting and so this had to be removed before running the gel (see Methods). Three separate 5ml medium samples of TeSR1 were concentrated to contain 50 $\mu$ g, 100 $\mu$ g and 200 $\mu$ g of protein in total (after HSA removal) and quantified using a BCA assay prior to western blotting. NPX2 was observed in all lanes in the TeSR1 medium after HSA extraction (S6B Fig). A clear Follistatin band was detected strongly in the E6 medium using only 20 $\mu$ g protein in comparison to, and stronger than, 50 $\mu$ g protein from the TeSR1 pluripotency medium. This confirms the enhancement of Follistatin on transfer of stem cells to a pro-differentiation medium (S6A Fig). Similarly, when cells were transferred to a chondrogenic differentiation medium for 48h after culture in TeSR1, there was an increase in Follistatin and decrease of NPTX2 (S6B Fig). NID1 was detected in TeSR1 after hESC culture but only at 200 $\mu$ g total protein and it is likely that this may be partly removed with the HSA.

### Discussion

We first investigated changes in commonly employed pluripotency associated proteins after only 48 hours following FGF2/TGF $\beta$  removal from the stem cell culture medium, confirming



**Fig 7. Re-establishment of relative marker abundances after a 7-day rescue.** For WIGW2 cells cultured as described in (Fig 1), but after media collection at 48 hours. E6 cultured cells were passaged and returned to E8 for 7 days with media collection every 24 hours. A) TGX Stain-free gels were loaded with 20ug of protein and imaged before transfer to membrane—the most prominent gel band is shown as a demonstration of equal loading. The membrane was probed with antibodies FST & NID1. B) Quantification shows the Log<sub>2</sub>-transformed ratio of FST/NID1 band intensity. C) Individual FST & NID1 band intensity, normalised to most prominent TGX band. Abundances calculated from densitometry in ImageJ. Membranes were imaged using the LICOR Odyssey system.

<https://doi.org/10.1371/journal.pone.0299365.g007>

that neither immunofluorescence nor flow cytometry for these markers could confidently detect pluripotency loss at this time. Although it is a pluripotency-associated marker, SOX2 increased slightly on transfer to E6 lacking FGF2 and TGFβ. This is to be expected as SOX2 expression has previously been found to increase upon SMAD2/3 inhibition and FGF2 deprivation [26]. Moreover, increases in SOX2 in hPSCs can initiate differentiation [27, 28] and SOX2 has an important role in specification and differentiation of PSCs towards the neuroectoderm lineage [29–31].

In this study we identified a series of proteins which function as secreted, early marker proteins for incipient loss of the pluripotent stem cell state in culture. To examine if the identified protein changes indicated a change in protein synthesis, we also investigated gene transcription and found that the proteome changes were indeed also generally reflected in changes in the respective gene transcripts. From this list of marker proteins (Table 1) a shortlist of eight proteins was identified for further analysis by Western blotting.

### Selected marker protein biology

Although the cellular proteomes of both hESC and hiPSC lines have been investigated by a variety of mass spectrometry techniques in the past [7–10] the protein secretome of these cells

has been under studied [11]. Using a combination of LC-MS/MS and bioinformatic analysis with confirmation by Western blotting, we identified several markers in the secretome indicative of healthy pluripotent stem cells and early incipient pluripotency loss. Moreover, we have also shown that cells returned to E8 after 48h in E6 show recovery and exhibit secretome marker proteins correlating with the pluripotent state again.

The transmembrane FGF2 receptor FGFR1 was one of the most highly represented secreted markers identified to increase upon growth factor withdrawal. There are several possibilities to explain the increased incidence of FGFR1 in the conditioned E6 medium. Firstly, FGFR1 is known to be released into the extracellular space by cleavage in its membrane-adjacent extracellular domain by MMP2 [32]. As FGFR1 proteins are endocytosed by the cell upon substrate binding and degraded in lysosomes [33–35], it is likely that decreased FGF2 abundance will result in an increased surface abundance of FGFR1, and a corresponding increase in the amount of FGFR1 which is cleaved by surface metalloproteinases. Indeed, a majority of the FGFR1 peptides identified as significantly increased in E6 in our proteomics study were N-terminal to the proposed MMP2 cleavage site. As it has previously been reported that a secreted form of the extracellular region of FGFR1 inhibits FGF2 activity [36] it is possible that the FGFR1 accumulation in FGF2 deficient media may further inhibit residual FGF2 signalling.

Indeed, the biological functions of other E6-enriched protein biomarkers are consistent with a hypothesis that cells triggered towards differentiation release factors into the medium that bind and modulate growth factors, sequestering pluripotency promoting factors such as FGF2 and TGF $\beta$  1, or stabilising other growth factors, to regulate their local availability. This will generate a feed-forward amplification of the original pluripotency dissolution signal which further highlights the need for early warning of incipient pluripotency loss. Follistatin is a well-known stem cell differentiation factor, and TGF $\beta$  inhibitor, functioning through the binding and neutralisation of TGF $\beta$  superfamily members [14]. Similarly, OLFML3, also enriched in E6, has been shown to bind to and stabilise BMP4, enhancing SMAD1/5/8 signalling in endothelial cells [37]. As BMP4 initiates differentiation of human embryonic stem cells to trophoblast in the absence of FGF [38–40] and to mesoderm and chondrocytes [41, 42], it could be concluded that after 48 hours of FGF2/TGF $\beta$  loss, these cells are secreting factors to facilitate specification to differentiated lineages [43]. In murine ES cells COCH has been reported to be expressed in response to BMP4 signalling [44], in our study COCH was enhanced in E6 medium possibly reflecting the loss of FGF which is known to suppress BMP signalling in hESCs [12]. The molecular differences in the pluripotent ESC state between human and murine reflect an earlier naïve murine ESC phenotype and later primed human ESC phenotype [45] and are supported by the wealth of data on the differences in signalling required for murine and human stem cell maintenance [12, 13, 46, 47]. Thus, BMP signalling has different effects in hESCs and murine ESCs, being active in stem cell maintenance in the latter [39, 47]. Since BMP signalling is strongly associated with differentiation to mesodermal and trophodermal lineages in hESCs [38, 41, 48–50], this would be in keeping with a correlation between targets such as COCH down stream of BMP and the dissolution of pluripotency precipitating incipient lineage differentiation.

SEMA3A, a neuronal signalling protein involved in axon guidance, is highly enriched in the E8 condition by both secretome proteomics and RNAseq, whilst its transmembrane receptor, NRP1 is highly enriched ( $q < 0.01$ ) in the E6 transcriptome [51–53]. NRP1 also binds to both FGF2 and TGF $\beta$  and modulates their signalling. In HUVEC cells, increased NRP1 expression was demonstrated to suppress Smad2/3 activation upon TGF $\beta$  stimulation, and expression of SMAD target genes in response to TGF $\beta$  was increased in NRP1 deficient HUVECs [54], a relationship which has been replicated *in vivo* in murine endothelial cells [55]. NRP1 is a highly promiscuous receptor, and ligands compete for NRP1 binding such that

competitive inhibition between ligands is common [56]. So NRP1 may also play a role in the modulating TGF $\beta$  and other growth factor activity in the transition from pluripotent stem cells to early differentiation.

CHGA is a precursor protein for several neuroendocrine signalling proteins. Peptides from across the whole CHGA molecule were identified as being highly enriched in E8 conditioned medium over those in E6 conditioned medium. Generally speaking, CHGA drives formation and release of secretory granules [57, 58], and its decreased abundance in stem cells during pluripotency dissolution could reflect a reduction in overall secretion. This would be consistent with the greater number of secreted proteins identified in E8 conditioned medium compared with in E6 though further data is needed to confirm this association.

NID1 is a secreted glycoprotein with two principal protein-binding domains (G2 and G3), separated by a flexible chain [59]. By domain-specific binding of different components of the basement membrane, NID1 stabilises the basement membrane, cross-linking its multiple components [60, 61]. NID1 is abundant in locations where there is a requirement for additional resilience against mechanical stress, and conversely, less abundant where more flexibility is required and tissue disposition is not fixed [62] for instance being actively degraded during basement membrane disassembly [63]. It has been demonstrated by proteomic analysis to be expressed at significant amounts by hESCs [11, 64]. Here we observed a dramatic reduction in the abundance of NID1 in E6 medium compared with E8. As NID1 has a stabilising effect on the basement membrane, it is likely that its abundance is decreased prior to the epithelial-mesenchymal transition inherent in the progression from a pluripotent stem cell to an early progenitor. This will involve the disassembly of several extracellular matrix complexes: in ovarian cancer cells NID1 plays a key role in this EMT transition [65].

NPTX2 binds to  $\alpha$ -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) receptor subunit Glutamate receptor 4 (GRIA4) in neuronal tissues [66] and has a presynaptic role in Pyramidal neurons with loss/loss of function associated with Alzheimer's disease [67] and Schizophrenia [68]. It is noteworthy that the expression of *GRIA4* is also much higher in E8 than E6 (Supplementary RNAseq data) as for NPTX2, suggesting some as-yet unidentified function for this ligand-receptor pair in hPSCs. NPTX2 is also a circulating biomarker of pancreatic cancer [69]. In glioblastoma NPTX2 has been shown to inhibit AKT which in turn represses NFKB activity [70]. There is evidence for AKT and Nf $\kappa$ B involvement in pluripotent stem cell maintenance [71] with PI3k/AKT being involved in switching Smad2/3 signalling between a pluripotent self-renewal and differentiation role through controlling ERK and GSK3 $\beta$  signalling [72].

Although we are unclear about the precise role of some of our pluripotent stem cell and early pluripotency secretome markers in the switch from self-renewal to early differentiation, we have robust evidence of their utility to predict the switch, suggesting that these marker proteins will be of commercial interest and in generation of cell therapies, where efficiency, robust and reliable cell output and cost effectiveness are paramount.

## Development of paired-ratio method

A caveat of these analyses, common to all secretome studies, is the lack of a suitable normalisation standard against which to compare proteins to robustly identify changes between conditions. Typically, 'omics' studies normalise to total protein or mRNA, or to house-keeping genes or proteins such as GAPDH or  $\beta$ -actin. These controls allow for reliable, reproducible comparisons across conditions or experiments. When analysing the secreted complement of the cell proteome, however, no such house-keeping protein exists, and the overall protein abundance between conditions is highly susceptible to variability due to the presence and

concentration of medium proteins, the density of cells, differing cell viability, cytoplasmic leakage, and a host of other factors. In this study, we endeavoured to limit the effect of these variables, and further sought to identify marker proteins which could be multiplexed against each other to provide ratios which are much more discriminatory between the experimental conditions. Regardless of other changes in the cell secretome, the relative abundances of these proteins should remain relatively constant within an individual genetic background, culture method, and cell state. As such, once baseline ratios are identified under control conditions, it is expected that these ratios would provide a rapid, quantitative warning of imminent pluripotency loss in cultured hPSCs, above and beyond any current method in terms of sensitivity, affordability, and rapidity of timescale.

In summary we have shown that there are several proteins in the secretome of hPSCs that are rapidly decreased or increased during dissolution of pluripotency. By employing a matrix of four of each such species, robust ratios of proteins could be generated which were indicative of stable pluripotency or its incipient loss. Additionally, other proteins were identified by MS/MS and will make excellent candidate biomarkers and provide biological interest for follow-up analyses. Such proteins will be invaluable as markers allowing intermittent or continuous culture monitoring using rapid turnaround techniques such as ELISA or high-throughput targeted MS, during the scale up and manufacture of hPSCs for cell therapy or for their use in pharmaceutical drug development and toxicity testing. For commercial use, in-line MS monitoring of media samples for specific proteins and their ratios would be possible to give early warning of loss of stem cell status provided that MS quantitation is a reliable output. The development of aptamers or nanobodies is exciting as they can be used with a variety of biosensor technologies. Optical sensors such as multi-parametric surface plasmon resonance, and bio-layer interferometry may be sensitive enough to accurately quantify the ratio of biomarkers directly but electrochemical detectors can enhance detection limits. They can be developed relatively cheaply giving a cost-effective means of monitoring with such markers. These devices will be equally valuable in the research laboratory, allowing rapid evaluation of culture status without loss of cell product.

## Materials and methods

### Cell culture

Cells were maintained in Essential 8 (E8) medium (Life Technologies) on Vitronectin-N (Life Technologies) and passaged using TrypLE dissociation reagent (Life Technologies). Prior to sample collection for secretome analysis, cells were plated in E8 medium, supplemented for 24 hours with 10 $\mu$ M ROCK inhibitor, at a density of 20,000 (Rebl.PAT) or 25,000 (Man-13) cells per cm<sup>2</sup> in 75cm<sup>2</sup> culture flasks (Corning). Twenty-four hours after plating (experimental time 0), cells were either fed with fresh E8 medium, or transferred to Essential 6 (E6) medium (Life Technologies). Cells were fed at 48 and 72 hours (experimental time 24h and 48h respectively) after plating and at these time-points medium was collected, spun at 1,000 x g at -9°C for 5 minutes then aliquoted into lo-bind 2ml centrifuge tubes (Eppendorf) and snap frozen in liquid nitrogen. At 72 hours after plating (experimental time 48h), cells were dissociated with TrypLE, spun at 1,000 x g in PBS for 5 minutes, the supernatant removed, and the pellet snap frozen in liquid nitrogen. Viability assays were conducted at medium harvest using Apotex-Glo assay (Promega) [73] or with the NucleoCounter NC-200 and Via-1 cassettes (Chemometec) according to manufacturers' instructions.

A subset of the cells at this time-point were also collected and processed for RNA-Seq analysis. Five biological replicates were conducted for Rebl.PAT, each of which was composed of three technical replicates, and all samples were analysed by LC-MS/MS in a single run. To

ensure quantitative reproducibility of the marker proteins identified across multiple analyses, four Man-13 biological replicates were collected, and these were analysed in two independent LC-MS/MS runs consisting of two biological replicates each. Some Man13 cells were also cultured in TeSR1 medium (Stem Cell Technologies) and then transferred to E6, or chondrogenic medium (DMEM:F12, 2mM L-glutamine, 1% (vol/vol) ITS, 1% (vol/vol) nonessential amino acids, 2% (vol/vol) B27, 90 $\mu$ M  $\beta$ -mercaptoethanol (all Life Technologies) + 20ng/ml BMP2 (R&D systems) 10ng/ml ActivinA (Peprotech, 20ng/ml FGF2 (Peprotech), 3 $\mu$ M CHIR99021 (Tocris), 100 $\mu$ M Pik90 (Selleckchem), modified from Wang et al 2019 [25].

An important component of an assay for decreased PSC culture quality is the ability to detect incipient pluripotency loss while the cells are still recoverable. The experiments above were repeated but after 2 days in E6, cells were restored to E8 medium for 7 days to assess whether full pluripotency could be recovered at the time that protein ratios detected early changes. Medium was refreshed every 24hrs with samples of conditioned medium being taken for analysis by Western Blot every 48 hours.

### Immunochemical staining of pluripotency markers

Cells were seeded at a density of 20,000 (Rebl.PAT) or 25,000 (Man-13) cells per cm<sup>2</sup> in a 24 well plate and treated identically to samples prepared for proteomic and RNA-Seq analysis (Fig 1A). Cell surface markers and transcription factors characteristic of pluripotent hESCs were detected using immunofluorescence. The cells were fixed with 4% paraformaldehyde and incubated with antibodies against stage specific embryonic antigens SSEA-4, SSEA-1 (R&D Systems), TRA-1-60, TRA-1-81 (Abcam) and transcription factors SOX2, NANOG (Cell signalling Technologies), and OCT-4 (BD Biosciences) at 4°C overnight. Secondary antibodies (Life Technologies), specific for the species and isotype of the primary antibody, conjugated to Alexafluors 488 or 594 were used for detection using a BX51 microscope (Olympus, Hertfordshire, UK) equipped with a Q-Imaging camera (Micro Imaging Applications Group, Inc, Buckinghamshire, UK). Image processing was done with the aid of Q-Capture Pro software package (Micro Imaging Applications Group, Inc).

### Gene expression analysis

Total RNA was extracted using mirVana™ miRNA Isolation Kit (Life Technologies), reverse transcribed using M-MLV reverse transcriptase (Promega) and candidate genes expression (normalised to GAPDH) assessed using SYBR Green PCR Master Mix (Applied Biosystems) with an ABI PRISM 7500 Real Time System (Applied Biosystems). Primers for rtPCR were as follows: GAPDH -FW (5' - ATGGGGAAGGTGAAGGTCG), GAPDH -RV (3' - TAAAAGCAGCCCTGGTGACC), Sox2 -FW (5' - AACCAGCGCATGGACAGTTAC), Sox2 -RV (3' - TGGTCCTGCATCATGCTGTAG), Oct4 -FW (5' - AGACCATCTGCCGCTTTGAG), Oct4 -RV (3'-), NANOG -FW (5' - GGCTCTGTTTTGCTATATCCCCTAA) and NANOG -RV (3' - CATTACGATGCAGCAAATACAAGA).

### Flow cytometry

Cells were collected at the time of plating, and at the 48-hour time-point for each experiment, as well as at the 24-hour time-point for a subset of experiments. All flow cytometry analyses were performed on a BD LSRFortessa™ cell analyser (Becton-Dickinson, San Jose, CA). Cells were dissociated using TrypLE Express (Life Technologies), fixed in the dark for 7 minutes at room temperature in 4% paraformaldehyde, blocked for thirty minutes to two hours in 5% FBS and were permeabilised for intracellular staining in ice-cold 70% methanol. Species and

fluorochrome-matched isotype controls were used for each antibody to control for non-specific binding.

Cells were incubated with directly conjugated primary antibodies for 30 minutes at room temperature, at the following dilutions; Phycoerythrin-conjugated mouse anti-TRA 1–81 (Ebioscience 12-8883-80), 1:300; Phycoerythrin-conjugated mouse isotype control (Santa-Cruz, SC2870) 1:75; Alexa 488-conjugated Mouse anti-Nanog (BD Bioscience, BD560791), 1:10; Alexa 488-conjugated mouse isotype control (BD Bioscience, BD557702) 1:100; Alexa 647-conjugated mouse anti-Sox2 (BD bioscience, BD56139) 1:80; Alexa 488-conjugated mouse anti-OCT4 (R&D systems, IC17591G) 1:50; Alexa 647-conjugated mouse isotype control (BD Bioscience, BD557714) 1:80. Staining for TRA-1-60 was performed using a non-directly conjugated antibody. Cells were incubated with the primary antibody mouse anti-TRA-1-60 (Invitrogen, MA1-023, dilution 1:400), followed by incubation with Alexa 488-conjugated goat-anti-mouse IgM secondary antibody (Invitrogen, A-21042, dilution 1:2500). Flow cytometry data was collected and preliminarily processed in FACSDIVA™ software followed by analysis and figure-preparation in FlowJo (both Becton-Dickinson, San Jose, CA).

### Proteomics sample preparation

Five biological repeats each consisting of three technical replicates of Rebl.PAT cells were performed, and all analysed together by label-free mass-spec (Fig 2A). Man-13 samples were analysed by LC-MS/MS in two batches of two biological repeats (each consisting of three technical replicates) to ensure that any identified biomarkers would be resilient to run variation (Fig 2A). A subset of the cells was also collected and processed for RNA-Seq analysis at the 48-hour time-point.

Media samples were digested using a modified Filter-assisted Sample Preparation (FASP) method [74] with the following modification: Equal volumes of conditioned media were concentrated to approximately 50  $\mu$ L with Microcon—10 kDa centrifugal filter units (Merck Millipore) at a speed of 14,000  $\times$  g. This was then washed and centrifuging three times with the addition of 100 mM phosphate buffer pH 7.4 before the proteins were reconstituted in 50  $\mu$ L of phosphate buffer. The protein concentration was determined using a Millipore Direct Detect<sup>®</sup> spectrometer and 50  $\mu$ g (Man-13) or 25  $\mu$ g (Rebl.PAT) of protein was added to a fresh 10 kDa filter tube with reduction, alkylation and digestion occurring using the filter tubes. After digestion peptides were collected by centrifugation and the samples were desalted with OLIGO™ R3 reversed-phase media [75] on a microplate system dried to completion and reconstituted just before analysis in 5% acetonitrile and 0.1% formic acid.

### LC-MS/MS

Digested samples were analysed by LC-MS/MS using an UltiMate 3000 Rapid Separation LC (RSLC, Dionex Corporation, Sunnyvale, CA) coupled to an Orbitrap Elite (Thermo Fisher Scientific, Waltham, MA) mass spectrometer. Peptide mixtures were separated using a multistep gradient from 95% A (0.1% FA in water) and 5% B (0.1% FA in acetonitrile) to 7% B at 1 min, 18% B at 35 min, 27% B in 43 min and 60% B at 44 min at 300 nL min<sup>-1</sup>, using a 75 mm  $\times$  250  $\mu$ m i.d. 1.7  $\mu$ M CSH C18, analytical column (Waters). Positive Polarity source voltage (ESI) voltage was 1.7Kv. Peptides were selected for fragmentation automatically by data dependent and the top 9 peaks were selected MS2 analysis. The total run time for each sample was 90 minutes.



## Identification and quantification of peptides

The acquired MS data was analysed using Progenesis LC-MS (v4.1, Nonlinear Dynamics). The retention times in each sample were aligned using one LC-MS run as a reference, then the “Automatic Alignment” algorithm was used to create maximal overlay of the two-dimensional feature maps. Where necessary a minimal amount of manual adjustment was employed to increase alignment score to above 80%. Features with charges  $\geq +5$  were masked and excluded from further analyses, as were features with less than 3 isotope peaks. The resulting peak lists were searched against the SwissProt (release 2016–04) and Trembl (release 2016–04) databases using Mascot v2.5.1, (Matrix Science). Search parameters included a precursor tolerance of 5 ppm and a fragment tolerance of 0.6 Da. Enzyme specificity was set to trypsin and one missed cleavage was allowed. Carbamidomethyl modification of cysteine was set as a fixed modification while methionine oxidation was set to variable. The Mascot results were imported into Progenesis LC-MS for annotation of peptide peaks. The Proteomic datasets generated in this study have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifiers PXD045525 & PXD045347.

## Proteomics data processing

SwissProt and Trembl IDs were used to align data between independent proteomics experiments. Normalised protein abundance data from Progenesis were log<sub>2</sub> transformed to improve the normality of distribution. All proteins used for statistical analysis had one or more unique peptides, and a Progenesis confidence score of 30 or more in their MS experiment. Statistical significance was calculated using linear regression models, fitted independently for each protein using the *lm* package in R and the formula  $ProteinAbundance \sim Media + BiologicalReplicate$ . The p-values from these models were then adjusted for multiple comparisons using the Benjamini & Hochberg (q) method.

Peptide sequences for all proteins were uploaded to the SignalP 5.0 server [20] to identify the proportion of proteins bearing a classical secretion signal-peptide sequence. Non-classical secretion was identified by uploading peptide sequences to the SecretomeP 2.0 server [21]; and the recommended cut-off score of 0.6 was used to quantify the proportion of proteins which were likely to be secreted through non-classical pathways. Protein GO-term associations were obtained by from Uniprot, or by use of WebGestalt’s GO-Slim output [76].

## RNASeq data analysis

BAM files were aligned to the human genome using BowTie [77]. In Linux, BAMs were sorted in the command interface using samtools [78], and the number of read counts per gene was obtained using featureCounts [79] (supplemental script 1). For each cell line, a count matrix table was constructed from all the counted BAMs.

The following analysis was conducted in R [80]. Genes with less than 10 reads in all three replicates were filtered out. Counts were normalised using TMM [81] and normalisation factors for each cell line generated. Subsets of filtered counts and normalisation factors were created to represent samples relating to loss of pluripotency. Differential analysis was performed using DESeq2 [82]. Statistically significant differentially expressed genes are defined using FDR corrected p-value  $< 0.05$ .

## Western blotting

Medium from a variety of cell lines, including WIGW2, Man-1, Man-7 and H9 cells in addition to MAN-13 and Rebl.PAT, was collected as described in Fig 1A, and 3ml of medium was

concentrated to 50–100 $\mu$ l in Microcon– 10 kDa centrifugal filter units (Merck Millipore). Samples were run on 10% Bis-Tris gels (Thermo #NW00100BOX), or TGX Stain-Free gels (BIO RAD #4568021) on a Mini Gel Tank (Thermo #A25977) using MES SDS Running Buffer (Thermo #B0002) alongside broad range markers (11–245KDa, NEB #P7712S). Each lane contained 20  $\mu$ g of protein, heated in Pierce lane marker reducing buffer (Thermo #39000) at 95°C for 10 minutes. TGX Stain-Free gels were activated prior to membrane transfer (45 seconds, UV Light Chemidoc MP imaging system). Gels were transferred using the iBlot2 Gel Transfer device (Thermo #IB21001), using iBlot 2 Transfer stacks (nitrocellulose membrane, Thermo #IB23001). Cells were analysed by immunoblotting using the following antibodies: 1/200 Mouse anti-Chromogranin A (Novus Biologicals, NBP2-44774); 1/100 Rabbit anti-COCH (Abcam, ab170266); 1/500 mouse anti-FGFR1 (R&D systems, MAB658); 1/1000 rabbit anti-Follistatin (Abcam, ab157471); 1/1000 rabbit anti-Entactin/NID1 (Abcam, ab133686); 1:1000 rabbit anti-Neuronal Pentaxtrin 2 (Abcam, ab191563); 1/1000 rabbit anti-Neurotensin (Abcam, ab172114); 1/300 rabbit anti-Olfactomedin-like 3 (Abcam, ab111712); 1/500 mouse anti-Semaphorin 3A (R&D, MAB1250); 1/500 mouse anti-Secreted Frizzled Related Protein 2 (R&D, MAB6838). Secondary antibody staining was performed either using 1/20,000 IRDye 800CW Goat anti-mouse, or 1/20,000 IRDye 680RD Donkey anti-Rabbit (P/N 925–32210 and 925–68071 respectively, both from LI-COR). These fluorescent secondary antibodies were imaged using the Odyssey CLx imaging system (LI-COR) typically in pairs of rabbit and mouse antibodies together. Image brightness/contrast adjustment and densitometry quantification was all performed in ImageJ (NIH).

### Protein extraction method for Western blotting

Media was concentrated using 10KDa Micron filter spin columns (Millipore), centrifuged at 14,000xg for 40mins at 4C, followed by 2 washes with PBS (Life Tech) and spinning at 14,000xg for 15mins at 4C. Protein was resuspended to a total volume of 200 $\mu$ l with 1x PBS. Protein concentration was determined using a BCA Assay kit (Pierce). Concentrated proteins were mixed and vortexed for 2 mins with 1% Trichloroacetic acid- Isopropanol (TCA-IPA) solution (1:10 ratio) and centrifuged 1500xg for 5mins at 5C. Pellets were washed twice with methanol, centrifuged at 800xg for 2mins and resuspended in 6x SDS-PAGE sample buffer, boiled at 96C 10mins and loaded onto an SDS gel for analysis.

### Teratoma formation

Teratoma generation was by subcutaneous injection into Scid Beige mice of approximately 1 million pelleted cells in Matrigel™ [83]. Animals were killed within 11 weeks of injection; skin biopsies fixed in 4% paraformaldehyde and embedded in paraffin wax for histological analysis. Some sections were stained with antibodies to early endoderm (GATA6- CST 5851 1:1600) neurectoderm (Beta3 tubulin R&D Systems MAB1195 1:200) or smooth muscle actin in mesoderm (alpha SMA R&D Systems MAB1420 1:200) followed by incubation with biotinylated secondary antibodies (Vector Labs biotinylated secondaries, anti-mouse, BP-9200 and anti-rabbit IgG BP-9100) and developed using the ABC peroxidase kit (Vector labs PK-6100). Slides were imaged using the 3D Histec Panoramic250 slide scanner and assessed using CaseViewer with Histoquant licence.

All mice were housed and handled in accordance with applicable Home Office guidelines. Local ethical approval was granted, and the work was done under UK home office licence 70/7838. Surgery was performed under isoflurane anaesthesia, and all efforts were made to minimize suffering.

## Supporting information

**S1 Fig.** A) Immunocytochemistry of Man-13 cells cultured in E8 (control) or after 48 hours of culture in E6. B) Flow cytometry of Rebl.PAT and Man-13 cells cultured in E8, or after 48 hours of culture in E6. For Rebl.PAT cells, extra flasks were cultured alongside experimental flasks, and these were used for observation of cell state at the 24-hour time-point (shown for Rebl.PAT). C) QRT-PCR of Man-13 and Rebl.PAT cells cultured in parallel to those used for medium collection, either maintained in E8, or after being incubated for 48 hours in E6. QRT-PCR for pluripotency associated markers Oct4 Sox2 and Nanog relative to GAPDH. (TIF)

**S2 Fig. Viability of cells after 48 hours of culture in E6.** At the 48-hour timepoint (Fig 1), viability of E8 and E6 cultured cells was assessed using Via-1 cassettes (Chemometec). All three flasks of cells for each condition were assayed, and no significant change between the conditions was observed in any experiment. (TIF)

**S3 Fig. Principal Component Analysis (PCA) of all secreted proteomics samples.** PC1 and PC3 are shown to demonstrate firstly the necessity of controlling for MS run in the statistical analysis (PC1, 67.2% of variance), and secondly that despite the other sources of variance, the experimental conditions are still clearly separated at PC3. (TIF)

**S4 Fig. Re-establishment of relative marker abundances in Man-1 after a 7-day rescue.** Man-1 cells were cultured under the same conditions as those in Fig 7, however the Western blot membrane was probed with OLFML3 & NID1 antibodies. Quantification shows the Log<sub>2</sub> ratio of E6/E8 marker abundance (OLFML3/NID1). Densitometry was performed in ImageJ. Membranes were imaged using the LICOR Odyssey system. (TIF)

**S5 Fig. Paraffin sections (7µm) of teratomas generated by subcutaneous implantation of human pluripotent stem cells cultured in in E8 (Man13) or after 2 days in E6 Man13 and Rebl.PAT) stained for morphological and protein markers of the 3 germ layers.** Paraffin sections were stained with Haematoxylin and eosin or with antibodies to early endoderm (GATA6) neurectoderm (Beta3 tubulin) or smooth muscle actin in mesoderm (alpha SMA) followed by a peroxidase labelled secondary antibody. Scale bar 100µm. (TIF)

**S6 Fig.** A) NPTX2 / FST Ratio of concentrated TeSR1 pluripotency media from cells grown for 48hrs, and medium subjected to HSA removal using 1%TCA-IPA or without HSA removal. B) NPTX2 / FST Ratio in media concentrated from cells grown in TeSR1 pluripotency medium versus mesodermal (stage1 chondrogenesis) differentiation medium showing that the change in protein ratios for NPTX2 and Follistatin holds in both experiments if HSA is removed. (TIF)

**S1 Raw images.**  
(PDF)

## Acknowledgments

We thank Qi Wang for the staining in [S6 Fig](#), as well as the Histology Facility for equipment used in this study. Special thanks go to Peter Walker & Grace Bako for their help with the Histology.

## Author Contributions

**Conceptualization:** Philip A. Lewis, Edina Silajdžić, Chris Denning, Daniel R. Brison, Susan J. Kimber.

**Data curation:** Philip A. Lewis, Helen Smith, Susan J. Kimber.

**Formal analysis:** Philip A. Lewis, Helen Smith.

**Funding acquisition:** Philip A. Lewis, Chris Denning, Daniel R. Brison, Susan J. Kimber.

**Investigation:** Philip A. Lewis, Edina Silajdžić, Nicola Bates.

**Methodology:** Philip A. Lewis, Edina Silajdžić, Nicola Bates, David Knight, Susan J. Kimber.

**Project administration:** Philip A. Lewis, Edina Silajdžić, Daniel R. Brison, Susan J. Kimber.

**Resources:** Philip A. Lewis, Edina Silajdžić, David Knight.

**Software:** Philip A. Lewis.

**Supervision:** Philip A. Lewis, Daniel R. Brison, Susan J. Kimber.

**Validation:** Philip A. Lewis, Edina Silajdžić, Nicola Bates, Christopher A. Smith, Fabrizio E. Mancini.

**Visualization:** Philip A. Lewis.

**Writing – original draft:** Philip A. Lewis, Daniel R. Brison, Susan J. Kimber.

**Writing – review & editing:** Philip A. Lewis, Edina Silajdžić, David Knight, Daniel R. Brison, Susan J. Kimber.

## References

1. Muller F.J., et al., A bioinformatic assay for pluripotency in human cells. *Nat Methods*, 2011. 8(4): p. 315–7. <https://doi.org/10.1038/nmeth.1580> PMID: 21378979
2. Bock C., et al., Reference Maps of human ES and iPS cell variation enable high-throughput characterization of pluripotent cell lines. *Cell*, 2011. 144(3): p. 439–52. <https://doi.org/10.1016/j.cell.2010.12.032> PMID: 21295703
3. Avior Y., Biancotti J.C., and Benvenisty N., TeratoScore: Assessing the Differentiation Potential of Human Pluripotent Stem Cells by Quantitative Expression Analysis of Teratomas. *Stem Cell Reports*, 2015. 4(6): p. 967–74. <https://doi.org/10.1016/j.stemcr.2015.05.006> PMID: 26070610
4. Kerz M., et al., A Novel Automated High-Content Analysis Workflow Capturing Cell Population Dynamics from Induced Pluripotent Stem Cell Live Imaging Data. *J Biomol Screen*, 2016. 21(9): p. 887–96. <https://doi.org/10.1177/1087057116652064> PMID: 27256155
5. Perestrelo T., et al., Pluri-IQ: Quantification of Embryonic Stem Cell Pluripotency through an Image-Based Analysis Software. *Stem Cell Reports*, 2017. 9(2): p. 697–709.
6. Smith D., Glen K., and Thomas R., Automated image analysis with the potential for process quality control applications in stem cell maintenance and differentiation. *Biotechnol Prog*, 2016. 32(1): p. 215–23. <https://doi.org/10.1002/btpr.2199> PMID: 26560993
7. Munoz J., et al., The quantitative proteomes of human-induced pluripotent stem cells and embryonic stem cells. *Mol Syst Biol*, 2011. 7: p. 550. <https://doi.org/10.1038/msb.2011.84> PMID: 22108792
8. Phanstiel D.H., et al., Proteomic and phosphoproteomic comparison of human ES and iPS cells. *Nat Methods*, 2011. 8(10): p. 821–7. <https://doi.org/10.1038/nmeth.1699> PMID: 21983960

9. Yamana R., et al., Rapid and deep profiling of human induced pluripotent stem cell proteome by one-shot NanoLC-MS/MS analysis with meter-scale monolithic silica columns. *J Proteome Res*, 2013. 12 (1): p. 214–21. <https://doi.org/10.1021/pr300837u> PMID: 23210603
10. Pripuzova N.S., et al., Development of a protein marker panel for characterization of human induced pluripotent stem cells (hiPSCs) using global quantitative proteome analysis. *Stem Cell Res*, 2015. 14 (3): p. 323–38. <https://doi.org/10.1016/j.scr.2015.01.009> PMID: 25840413
11. Soteriou D., et al., Comparative proteomic analysis of supportive and unsupportive extracellular matrix substrates for human embryonic stem cell maintenance. *J Biol Chem*, 2013. 288(26): p. 18716–31. <https://doi.org/10.1074/jbc.M113.463372> PMID: 23658023
12. Xu R.H., et al., Basic FGF and suppression of BMP signaling sustain undifferentiated proliferation of human ES cells. *Nat Methods*, 2005. 2(3): p. 185–90. <https://doi.org/10.1038/nmeth744> PMID: 15782187
13. Beattie G.M., et al., Activin A maintains pluripotency of human embryonic stem cells in the absence of feeder layers. *Stem Cells*, 2005. 23(4): p. 489–95. <https://doi.org/10.1634/stemcells.2004-0279> PMID: 15790770
14. Vallier L., Alexander M., and Pedersen R.A., Activin/Nodal and FGF pathways cooperate to maintain pluripotency of human embryonic stem cells. *J Cell Sci*, 2005. 118(Pt 19): p. 4495–509. <https://doi.org/10.1242/jcs.02553> PMID: 16179608
15. Gonzales K.A., et al., Deterministic Restriction on Pluripotent State Dissolution by Cell-Cycle Pathways. *Cell*, 2015. 162(3): p. 564–79. <https://doi.org/10.1016/j.cell.2015.07.001> PMID: 26232226
16. Ye J., et al., High quality clinical grade human embryonic stem cell lines derived from fresh discarded embryos. *Stem Cell Res Ther*, 2017. 8(1): p. 128.
17. Hammad M., et al., Identification of polymer surface adsorbed proteins implicated in pluripotent human embryonic stem cell expansion. *Biomater Sci*, 2016. 4(9): p. 1381–91. <https://doi.org/10.1039/c6bm00214e> PMID: 27466628
18. Chen G., et al., Chemically defined conditions for human iPSC derivation and culture. *Nat Methods*, 2011. 8(5): p. 424–9. <https://doi.org/10.1038/nmeth.1593> PMID: 21478862
19. Lippmann E.S., Estevez-Silva M.C., and Ashton R.S., Defined human pluripotent stem cell culture enables highly efficient neuroepithelium derivation without small molecule inhibitors. *Stem Cells*, 2014. 32(4): p. 1032–42. <https://doi.org/10.1002/stem.1622> PMID: 24357014
20. Almagro Armenteros J.J., et al., SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol*, 2019. 37(4): p. 420–423. <https://doi.org/10.1038/s41587-019-0036-z> PMID: 30778233
21. Bendtsen J.D., et al., Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng Des Sel*, 2004. 17(4): p. 349–56. <https://doi.org/10.1093/protein/gzh037> PMID: 15115854
22. Durinck S., et al., Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. *Nat Protoc*, 2009. 4(8): p. 1184–91. <https://doi.org/10.1038/nprot.2009.97> PMID: 19617889
23. Durinck S., et al., BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. *Bioinformatics*, 2005. 21(16): p. 3439–40. <https://doi.org/10.1093/bioinformatics/bti525> PMID: 16082012
24. UniProt C., UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res*, 2019. 47(D1): p. D506–D515. <https://doi.org/10.1093/nar/gky1049> PMID: 30395287
25. Wang T., et al., Enhanced chondrogenesis from human embryonic stem cells. *Stem Cell Res*, 2019. 39: p. 101497. <https://doi.org/10.1016/j.scr.2019.101497> PMID: 31326745
26. Greber B., et al., Conserved and divergent roles of FGF signaling in mouse epiblast stem cells and human embryonic stem cells. *Cell Stem Cell*, 2010. 6(3): p. 215–26. <https://doi.org/10.1016/j.stem.2010.01.003> PMID: 20207225
27. Kopp J.L., et al., Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells*, 2008. 26(4): p. 903–11. <https://doi.org/10.1634/stemcells.2007-0951> PMID: 18238855
28. Boer B., et al., Elevating the levels of Sox2 in embryonal carcinoma cells and embryonic stem cells inhibits the expression of Sox2:Oct-3/4 target genes. *Nucleic Acids Res*, 2007. 35(6): p. 1773–86. <https://doi.org/10.1093/nar/gkm059> PMID: 17324942
29. Zhao S., et al., SoxB transcription factors specify neuroectodermal lineage choice in ES cells. *Mol Cell Neurosci*, 2004. 27(3): p. 332–42. <https://doi.org/10.1016/j.mcn.2004.08.002> PMID: 15519247
30. Thomson M., et al., Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell*, 2011. 145(6): p. 875–89. <https://doi.org/10.1016/j.cell.2011.05.017> PMID: 21663792

31. Wang Z., et al., Distinct lineage specification roles for NANOG, OCT4, and SOX2 in human embryonic stem cells. *Cell Stem Cell*, 2012. 10(4): p. 440–54. <https://doi.org/10.1016/j.stem.2012.02.016> PMID: 22482508
32. Levi E., et al., Matrix metalloproteinase 2 releases active soluble ectodomain of fibroblast growth factor receptor 1. *Proc Natl Acad Sci U S A*, 1996. 93(14): p. 7069–74. <https://doi.org/10.1073/pnas.93.14.7069> PMID: 8692946
33. Haugsten E.M., et al., Different intracellular trafficking of FGF1 endocytosed by the four homologous FGF receptors. *J Cell Sci*, 2005. 118(Pt 17): p. 3869–81. <https://doi.org/10.1242/jcs.02509> PMID: 16091423
34. Auciello G., et al., Regulation of fibroblast growth factor receptor signalling and trafficking by Src and Eps8. *J Cell Sci*, 2013. 126(Pt 2): p. 613–24. <https://doi.org/10.1242/jcs.116228> PMID: 23203811
35. Sorokin A., et al., Internalization of fibroblast growth factor receptor is inhibited by a point mutation at tyrosine 766. *J Biol Chem*, 1994. 269(25): p. 17056–61. PMID: 7516330
36. Guillonneau X., et al., Fibroblast growth factor (FGF) soluble receptor 1 acts as a natural inhibitor of FGF2 neurotrophic activity during retinal degeneration. *Mol Biol Cell*, 1998. 9(10): p. 2785–802. <https://doi.org/10.1091/mbc.9.10.2785> PMID: 9763444
37. Mijlkovic-Licina M., et al., Targeting olfactomedin-like 3 inhibits tumor growth by impairing angiogenesis and pericyte coverage. *Mol Cancer Ther*, 2012. 11(12): p. 2588–99. <https://doi.org/10.1158/1535-7163.MCT-12-0245> PMID: 23002094
38. Pera M.F., et al., Regulation of human embryonic stem cell differentiation by BMP-2 and its antagonist noggin. *J Cell Sci*, 2004. 117(Pt 7): p. 1269–80. <https://doi.org/10.1242/jcs.00970> PMID: 14996946
39. Onishi K., et al., Local BMP-SMAD1 signaling increases LIF receptor-dependent STAT3 responsiveness and primed-to-naïve mouse pluripotent stem cell conversion frequency. *Stem Cell Reports*, 2014. 3(1): p. 156–68. <https://doi.org/10.1016/j.stemcr.2014.04.019> PMID: 25068129
40. Amita M., et al., Complete and unidirectional conversion of human embryonic stem cells to trophoblast by BMP4. *Proc Natl Acad Sci U S A*, 2013. 110(13): p. E1212–21. <https://doi.org/10.1073/pnas.1303094110> PMID: 23493551
41. Oldershaw R.A., et al., Directed differentiation of human embryonic stem cells toward chondrocytes. *Nat Biotechnol*, 2010. 28(11): p. 1187–94. <https://doi.org/10.1038/nbt.1683> PMID: 20967028
42. Cheng A., et al., Cartilage repair using human embryonic stem cell-derived chondroprogenitors. *Stem Cells Transl Med*, 2014. 3(11): p. 1287–94. <https://doi.org/10.5966/sctm.2014-0101> PMID: 25273540
43. Xu R.H., et al., BMP4 initiates human embryonic stem cell differentiation to trophoblast. *Nat Biotechnol*, 2002. 20(12): p. 1261–4. <https://doi.org/10.1038/nbt761> PMID: 12426580
44. Zhang J., et al., BMP induces cochlin expression to facilitate self-renewal and suppress neural differentiation of mouse embryonic stem cells. *J Biol Chem*, 2013. 288(12): p. 8053–60. <https://doi.org/10.1074/jbc.M112.433995> PMID: 23344953
45. Tesar P.J., et al., New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature*, 2007. 448(7150): p. 196–9. <https://doi.org/10.1038/nature05972> PMID: 17597760
46. Kunath T., et al., FGF stimulation of the Erk1/2 signalling cascade triggers transition of pluripotent embryonic stem cells from self-renewal to lineage commitment. *Development*, 2007. 134(16): p. 2895–902. <https://doi.org/10.1242/dev.02880> PMID: 17660198
47. Ying Q.L., et al., BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell*, 2003. 115(3): p. 281–92. [https://doi.org/10.1016/S0092-8674\(03\)00847-X](https://doi.org/10.1016/S0092-8674(03)00847-X) PMID: 14636556
48. Roberts R.M., et al., Specification of trophoblast from embryonic stem cells exposed to BMP4. *Biol Reprod*, 2018. 99(1): p. 212–224. <https://doi.org/10.1093/biore/iy070> PMID: 29579154
49. Kurek D., et al., Endogenous WNT signals mediate BMP-induced and spontaneous differentiation of epiblast stem cells and human embryonic stem cells. *Stem Cell Reports*, 2015. 4(1): p. 114–28. <https://doi.org/10.1016/j.stemcr.2014.11.007> PMID: 25544567
50. Loh K.M., et al., Mapping the Pairwise Choices Leading from Pluripotency to Human Bone, Heart, and Other Mesoderm Cell Types. *Cell*, 2016. 166(2): p. 451–467. <https://doi.org/10.1016/j.cell.2016.06.011> PMID: 27419872
51. Janssen B.J., et al., Neuropilins lock secreted semaphorins onto plexins in a ternary signaling complex. *Nat Struct Mol Biol*, 2012. 19(12): p. 1293–9. <https://doi.org/10.1038/nsmb.2416> PMID: 23104057
52. Lepelletier Y., et al., Immunosuppressive role of semaphorin-3A on T cell proliferation is mediated by inhibition of actin cytoskeleton reorganization. *Eur J Immunol*, 2006. 36(7): p. 1782–93. <https://doi.org/10.1002/eji.200535601> PMID: 16791896

53. Catalano A., The neuroimmune semaphorin-3A reduces inflammation and progression of experimental autoimmune arthritis. *J Immunol*, 2010. 185(10): p. 6373–83. <https://doi.org/10.4049/jimmunol.0903527> PMID: 20937848
54. Aspalter I.M., et al., Alk1 and Alk5 inhibition by Nrp1 controls vascular sprouting downstream of Notch. *Nat Commun*, 2015. 6: p. 7264. <https://doi.org/10.1038/ncomms8264> PMID: 26081042
55. Hirota S., et al., Neuropilin 1 balances beta8 integrin-activated TGFbeta signaling to control sprouting angiogenesis in the brain. *Development*, 2015. 142(24): p. 4363–73.
56. Bachelder R.E., et al., Competing autocrine pathways involving alternative neuropilin-1 ligands regulate chemotaxis of carcinoma cells. *Cancer Res*, 2003. 63(17): p. 5230–3. PMID: 14500350
57. Stettler H., et al., Determinants for chromogranin A sorting into the regulated secretory pathway are also sufficient to generate granule-like structures in non-endocrine cells. *Biochem J*, 2009. 418(1): p. 81–91. <https://doi.org/10.1042/BJ20071382> PMID: 18973469
58. Hendy G.N., et al., Targeted ablation of the chromogranin a (Chga) gene: normal neuroendocrine dense-core secretory granules and increased expression of other granins. *Mol Endocrinol*, 2006. 20(8): p. 1935–47. <https://doi.org/10.1210/me.2005-0398> PMID: 16556729
59. Fox J.W., et al., Recombinant nidogen consists of three globular domains and mediates binding of laminin to collagen type IV. *EMBO J*, 1991. 10(11): p. 3137–46. <https://doi.org/10.1002/j.1460-2075.1991.tb04875.x> PMID: 1717261
60. Ho M.S., et al., Nidogens-Extracellular matrix linker molecules. *Microsc Res Tech*, 2008. 71(5): p. 387–95. <https://doi.org/10.1002/jemt.20567> PMID: 18219668
61. McKee K.K., et al., Role of laminin terminal globular domains in basement membrane assembly. *J Biol Chem*, 2007. 282(29): p. 21437–47. <https://doi.org/10.1074/jbc.M702963200> PMID: 17517882
62. Poschl E., et al., Collagen IV is essential for basement membrane stability but dispensable for initiation of its assembly during early development. *Development*, 2004. 131(7): p. 1619–28. <https://doi.org/10.1242/dev.01037> PMID: 14998921
63. Mayer U., et al., Sites of nidogen cleavage by proteases involved in tissue homeostasis and remodeling. *Eur J Biochem*, 1993. 217(3): p. 877–84. <https://doi.org/10.1111/j.1432-1033.1993.tb18316.x> PMID: 8223643
64. Evseenko D., et al., Identification of the critical extracellular matrix proteins that promote human embryonic stem cell assembly. *Stem Cells Dev*, 2009. 18(6): p. 919–28. <https://doi.org/10.1089/scd.2008.0293> PMID: 19021502
65. Zhou Y., et al., NID1, a new regulator of EMT required for metastasis and chemoresistance of ovarian cancer cells. *Oncotarget*, 2017. 8(20): p. 33110–33121. <https://doi.org/10.18632/oncotarget.16145> PMID: 28416770
66. von Roemeling C.A., et al., Neuronal pentraxin 2 supports clear cell renal cell carcinoma by activating the AMPA-selective glutamate receptor-4. *Cancer Res*, 2014. 74(17): p. 4796–810. <https://doi.org/10.1158/0008-5472.CAN-14-0210> PMID: 24962026
67. Xiao M.F., et al., NPTX2 and cognitive dysfunction in Alzheimer's Disease. *Elife*, 2017. 6. <https://doi.org/10.7554/eLife.23798> PMID: 28440221
68. Xiao M.F., et al., A biomarker-authenticated model of schizophrenia implicating NPTX2 loss of function. *Sci Adv*, 2021. 7(48): p. eabf6935. <https://doi.org/10.1126/sciadv.abf6935> PMID: 34818031
69. Garcia-Ortiz M.V., et al., Circulating NPTX2 methylation as a non-invasive biomarker for prognosis and monitoring of metastatic pancreatic cancer. *Clin Epigenetics*, 2023. 15(1): p. 118. <https://doi.org/10.1186/s13148-023-01535-4> PMID: 37481552
70. Shukla S., et al., A DNA methylation prognostic signature of glioblastoma: identification of NPTX2-PTEN-NF-kappaB nexus. *Cancer Res*, 2013. 73(22): p. 6563–73.
71. Armstrong L., et al., The role of PI3K/AKT, MAPK/ERK and NFkappabeta signalling in the maintenance of human embryonic stem cell pluripotency and viability highlighted by transcriptional profiling and functional analysis. *Hum Mol Genet*, 2006. 15(11): p. 1894–913. <https://doi.org/10.1093/hmg/ddl112> PMID: 16644866
72. Singh A.M., et al., Signaling network crosstalk in human pluripotent cells: a Smad2/3-regulated switch that controls the balance between self-renewal and differentiation. *Cell Stem Cell*, 2012. 10(3): p. 312–26. <https://doi.org/10.1016/j.stem.2012.01.014> PMID: 22385658
73. Schultz, S. *Determining the Predictive Mechanism of Toxicity Using a Single-Well Multiplexed Assay*. 2009 [cited 2019; <http://www.promega.co.uk/resources/pubhub/determining-the-predictive-mechanism-of-toxicity-with-apotox-glo/>].
74. Wisniewski J.R., et al., Universal sample preparation method for proteome analysis. *Nat Methods*, 2009. 6(5): p. 359–62. <https://doi.org/10.1038/nmeth.1322> PMID: 19377485

75. Horton E.R., et al., Modulation of FAK and Src adhesion signaling occurs independently of adhesion complex composition. *J Cell Biol*, 2016. 212(3): p. 349–64. <https://doi.org/10.1083/jcb.201508080> PMID: 26833789
76. Liao Y., et al., WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res*, 2019. 47(W1): p. W199–W205. <https://doi.org/10.1093/nar/gkz401> PMID: 31114916
77. Langmead B., et al., Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*, 2009. 10(3): p. R25. <https://doi.org/10.1186/gb-2009-10-3-r25> PMID: 19261174
78. Li H., et al., The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 2009. 25(16): p. 2078–9. <https://doi.org/10.1093/bioinformatics/btp352> PMID: 19505943
79. Liao Y., Smyth G.K., and Shi W., featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 2014. 30(7): p. 923–30. <https://doi.org/10.1093/bioinformatics/btt656> PMID: 24227677
80. Team, R.C. *R: A language and environment for statistical computing*. 2020; <https://www.R-project.org/>.
81. Robinson M.D. and Oshlack A., A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*, 2010. 11(3): p. R25. <https://doi.org/10.1186/gb-2010-11-3-r25> PMID: 20196867
82. Love M.I., Huber W., and Anders S., Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*, 2014. 15(12): p. 550. <https://doi.org/10.1186/s13059-014-0550-8> PMID: 25516281
83. Bantounas I., et al., Generation of Functioning Nephrons by Implanting Human Pluripotent Stem Cell-Derived Kidney Progenitors. *Stem Cell Reports*, 2018. 10(3): p. 766–779. <https://doi.org/10.1016/j.stemcr.2018.01.008> PMID: 29429961