Data Descriptor

Proteomic analysis data of human induced pluripotent stem cells and human fibroblast cells

Mio Iwasaki[1*], Yasushi Ishihama[2*]

1.   Center for iPS Cell Research and Application, Kyoto University, Kyoto 606-8507, Japan
2.   Graduate school of Pharmaceutical Sciences, Kyoto University, Kyoto 606-8501, Japan

ORCID: Mio Iwasaki: https://orcid.org/0000-0003-2085-732X
ORCID: Yasushi Ishihama: https://orcid.org/0000-0001-7714-203X

* Corresponding authors.
E-mail addresses: omio@cira.kyoto-u.ac.jp (M. Iwasaki) and yishiham@pharm.kyoto-u.ac.jp (Y. Ishihama)

Dataset summary

Specific subject area	Human iPS cell proteome and fibroblast cell proteome
Data acquisition	Data-dependent acquisition mode by TripleTOF 5600 (Sciex)
Dataset repository	jPOST
Dataset identifiers	JPST000081, JPST000082, JPST000083, JPST000085, JPST000086, JPST000087, JPST000088, JPST000089

Abstract

The molecular mechanisms for the self-renewal and pluripotency of human induced pluripotent stem cells (hiPSCs) remain unclear. Obtaining a deep proteome profile of hiPSCs will help clarify these mechanisms. Here we applied a long gradient time (8 hours) with meter-long column analysis to achieve rapid and deep profiling of the hiPSC proteome. The data accompanying this paper have been deposited to jPOST with the identifiers JPST000081, JPST000082, JPST000083, JPST000085, JPST000086, JPST000087, JPST000088, and JPST000089.

1. Materials and Methods

Here we provide the identified protein catalog of hiPSCs and human dermal fibroblasts (HDFs) analyzed by nanoLC/MS/MS using a 2-meter long monolithic silica capillary column[1].

1.1. Samples

Three HDF lines (aHDF1388, aHDF1419 and Tig120slc) were maintained in Dulbecco's modified eagle medium (DMEM, Nacalai Tesque, Japan) containing 10% fetal bovine serum (FBS, JapanSerum) and 0.5% penicillin and streptomycin (Invitrogen). Five hiPSC lines (201B7, 32R1, 414C2,585A1 and 606A1) were maintained on feeder cells. After removal of feeder cells, iPS cells were washed with PBS(−) and collected with scrapers. The pellets were stored at −80°C.

1.2. Sample pretreatment for MS analysis

Protein extraction was performed according to the phase transfer surfactant (PTS) protocol[2,3] with some modifications. The cell pellets were solubilized with PTS buffer (12 mM SDC, 12 mM SLS, and 50 mM ammonium bicarbonate). Then the protein samples were reduced with 10 mM

dithiothreitol at room temperature for 30 min, and alkylated with 55 mM iodoacetamide in the dark at room temperature for 30 min. After alkylation, the samples were 5-fold diluted with 50 mM ammonium bicarbonate and digested by Lys-C/trypsin at an enzyme/substrate ratio of 1:100. To remove the PTS surfactant, an equal volume of ethyl acetate was added to the eluent solution, and the mixture was acidified with 0.5% trifluoroacetic acid (final concentration). The mixture was shaken for 1 min and centrifuged at 15700 × g for 2 min, and the aqueous phase was collected and desalted with C18 Stage Tips[4].

1.3. MS analysis

NanoLC/MS/MS analysis was conducted using the TripleTOF 5600 system (Sciex, Foster City, CA) equipped with a Dionex Ultimate 3000 pump (Germering, Germany) and HTC-PAL autosampler (CTC Analytics, Zwingen, Switzerland). Monolithic silica columns (100 μm i.d., 2 m long) were prepared as described previously[5]. The coiled monolithic capillary columns were connected to a self-pulled emitter (20 μm i.d., 5 μm tip) formed with a Sutter P-2000 (Novato, CA) with a conductive distal coating end applied with an Ion Coater Model IB-2 (Eiko Engineering, Ibaraki, Japan), at which the spray voltage was applied (24 A, 2300 V). The column temperature was controlled at 25°C. The injection volume was 5 μL, and the flow rate was 500 nL/min. The mobile phases consisted of (A) 0.5% acetic acid and (B) 0.5% acetic acid in 80% acetonitrile. A three-step gradient condition of 5–40% B in 480 min, 40–100% B in 5 min and 100% B for 10 min was employed. The MS scan range was m/z 300–1500. The top 10 precursor ions were selected in each MS scan for subsequent MS/MS scans. MS scans were performed for 0.25 s, and subsequently 10 MS/MS scans were performed for 0.1 s each. To minimize repeated scanning, previously scanned ions were excluded for 12 s. The CID energy was automatically adjusted by the rolling CID function of Analyst TF 1.5. Triplicate analyses were done for each sample, and blank runs were inserted between different samples.

1.4. Data analysis

The raw data files were analyzed using AB SCIEX MS Data Converter to create peak lists on the basis of the recorded fragmentation spectra. Peptides and proteins were identified using Mascot v2.3 (Matrix Science, London, U.K.) against IPI human database v3.87 (91464 sequences) with a precursor mass tolerance of 20 ppm, a fragment ion mass tolerance of 0.1 Da and a strict trypsin specificity allowing for up to 2 missed cleavages. The carbamidomethylation of cysteine was set as a fixed modification, and methionine oxidation was allowed as a variable modification. Peptides were initially rejected if the Mascot score was below the 95% confidence limit based on the "identity" score of each peptide and the peptide length was less than 7 amino acid residues. For protein identification, peptides were grouped into "protein groups" based on the rules previously established[6]. Then, at least two confidently identified peptides per protein were used for the protein identification. In addition, single peptides with higher confidence (p < 0.01) were allowed for the protein identification. False discovery rates (FDR) were estimated by searching against a randomized decoy database created by Mascot.

2. Data description

In the present work, 24 LC/MS/MS runs were conducted for 8 different samples with triplicate injections. As a result, 76,762 unique peptides and 8179 protein groups were identified from 5 cell lines of hiPSCs. From this hiPSC proteome data, the translational products of 1091 out of 6244 genes (missing proteins in the year of 2012) were newly confirmed. For 3 HDF lines, 60124 unique peptides and 7202 protein groups were identified. In total, 98,977 unique peptides and 9510 protein groups were identified from these 8 samples. The total MS/MS events were 3783765 and the number of MS/MS spectra matched to peptides was 1012918 (27%).

Acknowledgments

References

[1] Yamana, R.; Iwasaki, M.; Wakabayashi, M.; Nakagawa, M.; Yamanaka, S.; Ishihama, Y. *J. Proteome Res.* **2013**, *12 (1)*, 214–21.
[2] Masuda, T.; Saito, N.; Tomita, M.; Ishihama, Y. *Mol. Cell. Proteomics* **2009**, *8 (12)*, 2770–7.
[3] Masuda, T.; Tomita, M.; Ishihama, Y. *J. Proteome Res.* **2008**, *7 (2)*, 731–40.
[4] Rappsilber, J.; Ishihama, Y.; Mann, M. *Anal. Chem.* **2003**, *75 (3)*, 663–70.
[5] Miyamoto, K.; Hara, T.; Kobayashi, H.; Morisaka, H.; Tokuda, D.; Horie, K.; Koduki, K.; Makino, S.; Núñez, O.; Yang, C.; Kawabe, T.; Ikegami, T.; Takubo, H.; Ishihama, Y.; Tanaka, N. *Anal. Chem.* **2008**, *80 (22)*, 8741–50.
[6] Nesvizhskii, A. I.; Aebersold, R. *Mol. Cell. Proteomics* **2005**, *4 (10)*, 1419–40.

Supporting Information

Supporting information is available online at https://doi.org/10.14889/jpdm.2019.0002.