Sample Article

Data Descriptor

https://dx.doi.org/10.14889/jpdm.2020.xxxx

JPDM
Journal of Proteome Data and Methods

Data for proteomic analysis of DNA-binding proteins

Taro Suzuki[1], Jack Smith[1], Sun-Moon Lake[2], Andrew Baker[1*]

1. Laboratory of Proteome Research, National Institute of Biomedicine, Osaka, Japan
2. Graduate School of Science, University of Proteomics, New York, NY, USA

ORCID: Andrew Baker: https://orcid.org/0000-000X-XXXX-XXXX

* Corresponding author.
E-mail address: abaker@nib.go.jp

© 2019 The Authors. This is an open access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/).

Keywords

Cell line, DNA-binding, Transcription factor

Dataset summary

Specific subject area	DNA-binding proteins and mass spectrometry
Data acquisition	Data-dependent acquisition acquired on Q-Exactive (Thermo)
Dataset repository	jPOST
Dataset identifiers	JPST100999

Abstract

Interactions between DNA and DNA-binding proteins are required for most cellular processes. Thus, it is clearly important to identify and quantify these interactions for understanding basic cellular mechanisms. We used an oligonucleotide probe to enrich specific DNA-binding proteins. As a result, we identified 512 proteins, including 120 novel DNA-binding protein candidates. The data accompanying this paper have been deposited to jPOST with identifier JPST100999.

1. Materials and Methods

In the present work, we provide the DNA-binding protein catalog obtained by proteomics experiments using the affinity purification with the oligonucleotide probe followed by LC/MS/MS analysis [1].

1.1. Samples

HeLa-S3 cells were grown in DMEM with 10% fetal bovine serum plus antibiotics in 10% CO_2 at 37 °C. For SILAC labeling, HeLa-S3 cells were cultured in DMEM supplemented with 10% dialyzed fetal bovine serum and either 28.0 mg/L normal isotopic abundance arginine and 48.7 mg/L normal isotopic abundance lysine (Light) or 28.0 mg/L arginine with six ^{13}C and four ^{15}N atoms and 48.7 mg/L lysine with six ^{13}C and two ^{15}N atoms (Heavy) [2]. Labeling efficiency was confirmed after five passages.

1.2. Sample pretreatment for MS analysis

Cell lysate was diluted with reaction buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 0.1% Triton X100) to a final protein concentration of 4 mg/mL. $MnCl_2$ was added to 250 µL of lysates at a final concentration of 20 mM. Affinity purification of DNA-binding proteins was carried out at different concentrations of the oligonucleotide probe (1 nM – 10 µM) with SILAC-labeled “Light” lysate at room temperature with gentle shaking for 10 min. The control samples were prepared in the same way using the control probe with SILAC-labeled “Heavy” lysate. After the incubation, these samples were mixed, denatured by 5 M urea, reduced with DTT (5 mM final concentration), and alkylated with iodoacetamide (20 mM final concentration). After the alkylation step, the solution was substituted by digestion buffer by gel filtration followed

Received 19 June 2099; Received in revised form 7 October 2099; Accepted 15 October 2099; Published 26 October 2099

Sample Article

by digestion at 37 °C overnight with sequencing-grade trypsin at an enzyme/substrate ratio of 1:100. Peptides were dissolved in 50 µL of 2 M urea and 1% TFA and desalted using StageTips [3]. Each experiment was performed in duplicate.

1.3. MS analysis

All nanoflow LC/MS/MS experiments were performed on a Q Exactive mass spectrometer (Thermo Scientific). Data were acquired in data-dependent mode using Xcalibur software. The precursor ion scan MS spectra (*m/z* 350–1800) were acquired in the Orbitrap with 70,000 resolution after accumulation of ions to a 3 x 10^6 target value. The twelve most intensive ions were isolated and fragmented in the collision cell by higher-energy collisional dissociation (HCD) with a maximum injection time of 120 ms and 35,000 resolution. In data-dependent LC/MS/MS experiments dynamic exclusion was used with 20-s exclusion duration. Typical mass spectrometric conditions were as follows: spray voltage, 2 kV; no sheath and auxiliary gas flow; heated capillary temperature, 250 °C; normalized HCD collision energy 25%. The MS/MS ion selection threshold was set to 5 x 10^4 counts. A 3.0 Da isolation width was chosen.

1.4. Data analysis

All raw files were processed by MaxQuant software suite (version 1.3.0.5) supported by the Andromeda search engine for peptide identification [4]. MaxQuant was used to score peptides for identification based on a search with an initial allowed mass deviation of the precursor ion of up to 7 ppm. The allowed fragment mass deviation was 20 ppm. Peak lists were searched against a concatenated forward and reversed version of the UniProt human database (release 2011_11) combined with 262 common contaminants was performed using the Andromeda search engine. Enzyme specificity was set as C-terminal to Arg and Lys with allowed cleavage proline bonds and a maximum of three missed cleavages. Carbamidomethylation of cysteine was set as a fixed modification and methionine oxidation as a variable modification. Peptides and proteins were accepted with a false discovery rate of 1.0%.

2. Data description

In the present work, 10 LC/MS/MS runs were conducted for 5 different samples (probe conc: 1 nM, 10 nM, 100 nM, 1 µM and 10 µM) with duplicate injections. The SILAC light-labeled lysate was used for the control, while the heavy-labeled lysate was used for the samples with the oligonucleotide probe. The details on the relationship between samples, the peak list files, MS raw files and the result files are described in Supplementary Table 1. As a result, 512 proteins including 120 novel DNA-binding protein candidates were identified.

Acknowledgments

This work was supported by Grants-in-Aid 8787878787 to T. S. and 78787878 and 1919191919 to A. B. from the Ministry of Science of Japan. We thank Jiro Sato for advice on MS analysis.

References

[1] Suzuki, T.; Baker, A., *J. Proteome Res.* **2014**, *13 (2)*, 5461–70. DOI:10.1021/pr5012685.
[2] Ong, S.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M., *Mol. Cell. Proteomics* **2002,** *1*, 376–86. DOI:10.1074/mcp.m200025-mcp200.
[3] Rappsilber, J.; Mann, M.; Ishihama, Y., *Nat. Protoc.* **2007**, *2 (8)*, 1896-1906. DOI:10.1038/nprot.2007.261.
[4] Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M., *J. Proteome Res.* **2011**, *10*, 1794–1805. DOI:10.1021/pr101065j.

Supporting Information

Supporting information is available online at http://www.jhupo.org/jpdm/.