OXFORD

BIOLOGY Methods & Protocols

Biology Methods and Protocols, 2024, bpae039

https://doi.org/10.1093/biomethods/bpae039

Advance Access Publication Date: 29 May 2024

Methods Article

Mapping adipocyte interactome networks by HaloTag-enrichment-mass spectrometry

Junshi Yazaki [1,2,†,*], Takashi Yamanashi[1,3,4], Shino Nemoto[5], Atsuo Kobayashi[1], Yong-Woon Han [1], Tomoko Hasegawa[1], Akira Iwase [6], Masaki Ishikawa[7], Ryo Konno[7], Koshi Imami [8,‡], Yusuke Kawashima[7,†,‡] and Jun Seita[1,3,4,†,‡]

[1]Laboratory for Integrative Genomics, RIKEN Center for Integrative Medical Sciences, Yokohama, 230-0045, Japan
[2]Faculty of Agriculture, Laboratory for Genome Biology, Setsunan University, Osaka, 573-0101, Japan
[3]Medical Data Deep Learning Team, Advanced Data Science Project, RIKEN Information R&D and Strategy Headquarters, RIKEN, Tokyo, 103-0027, Japan
[4]School of Integrative and Global Majors, University of Tsukuba, Tsukuba, 305-8577, Japan
[5]Laboratory for Intestinal Ecosystem, RIKEN Center for Integrative Medical Sciences, Yokohama, 230-0045, Japan
[6]Cell Function Research Team, RIKEN Center for Sustainable Resource Science, Yokohama, 230-0045, Japan
[7]Department of Applied Genomics, Technology Development Team, Kazusa DNA Research Institute, Kisarazu, 292-0818, Japan
[8]Proteome Homeostasis Research Unit, RIKEN Center for Integrative Medical Sciences, Yokohama, 230-0045, Japan

*Correspondence address. Faculty of Agriculture, Laboratory for Genome Biology, Setsunan University, Hirakata City, Osaka, 573-0101, Japan. Tel: +81 728-96-5402; Fax: +81 72-896-6003; E-mail: junshi.yazaki@setsunan.ac.jp

[†]These authors contributed equally to this work.
[‡]Senior author.

Abstract

Mapping protein interaction complexes in their natural state *in vivo* is arguably the Holy Grail of protein network analysis. Detection of protein interaction stoichiometry has been an important technical challenge, as few studies have focused on this. This may, however, be solved by artificial intelligence (AI) and proteomics. Here, we describe the development of HaloTag-based affinity purification mass spectrometry (HaloMS), a high-throughput HaloMS assay for protein interaction discovery. The approach enables the rapid capture of newly expressed proteins, eliminating tedious conventional one-by-one assays. As a proof-of-principle, we used HaloMS to evaluate the protein complex interactions of 17 regulatory proteins in human adipocytes. The adipocyte interactome network was validated using an *in vitro* pull-down assay and AI-based prediction tools. Applying HaloMS to probe adipocyte differentiation facilitated the identification of previously unknown transcription factor (TF)–protein complexes, revealing proteome-wide human adipocyte TF networks and shedding light on how different pathways are integrated.

Keywords: interaction mapping; network analysis; HaloMS; adipocyte

Introduction

Lifestyle-related diseases, such as diabetes, are estimated to affect hundreds of millions of people worldwide and are a major obstacle to extending healthy life expectancy [1]. Many such diseases are caused by metabolic disorders. Numerous studies have evaluated the molecular regulators of metabolism. For example, studies on transcription factors (TFs) and their interacting partners (i.e. the transcriptional network) have identified CCAAT/enhancer-binding protein and peroxisome proliferator-activated receptors (CEBPs, PPARs) as critical TFs [2]. To date, most systematic transcriptional network studies have focused on nucleic acids and metabolites [3, 4], while few have concentrated on proteins [5, 6]. Nevertheless, elucidating the protein networks directly related to disease is essential to understanding which molecular factors regulate metabolism, including those associated with lifestyle-related diseases. Considering the existing one-to-one protein–protein interaction detection technology, at least 400 million experiments would be required to determine the protein–protein interactions in an organism with approximately 20 000 genes, such as humans [7]. Therefore, the development of a nonbiased, high-throughput technology capable of detecting protein complex interactions is an important challenge. Furthermore, with the increasing focus on targeted drug discovery, integrating artificial intelligence (AI) in proteomics research has become a major focus [8–14].

While the search for biomarkers related to diseases is progressing rapidly, a systematic interaction network of proteins involved in lifestyle-related diseases remains elusive. Here, we aimed to report HaloTag-based affinity purification MS (HaloMS), a high-affinity capture method that combines the fabrication of *in situ* synthesized proteins for affinity purification with mass spectrometry (APMS). We applied this improved APMS to map protein complex interaction networks in adipocytes, which are closely related to lifestyle-related diseases [15], using 17 human regulatory proteins as bait. Subsequent in-depth analysis of the adipocyte interactome network revealed heretofore unknown specific and interacting networks of adipocyte signaling pathways regulated by multiple human proteins.

Received: 01 May 2024. **Revised:** 19 May 2024. **Editorial decision:** 20 May 2024. **Accepted:** 28 May 2024

Materials and methods

Arabidopsis thaliana

Arabidopsis thaliana plants used in this study were in the ecotype Col-0 background. Arabidopsis was used because of the availability of inexpensive, cost-effective ORF clones and the technical ease of protein preparation from plant material. We also used Arabidopsis because of the results obtained from protein arrays for comparison with the new technology, HaloMS. Plants were grown on 0.6% (w/v) gelzan (Sigma-Aldrich, cat. no. G1910) plates containing Murashige and Skoog salt (Fujifilm Wako, cat. no. 392-00591) and 1% sucrose (Fujifilm Wako, cat. no. 196-00015) medium at 22°C with a photoperiod of 16 h white light (40–50 μmol m^{-2} s^{-1}) and 8 h darkness.

HEK 293-F cells

Commercially available HEK 293-F cells (Thermo Fisher Scientific, cat. no. R79007) were cultured in a 125-ml flask in 30 ml of FreeStyle 293 Expression Medium, according to the manufacturer's recommendations. Subsequently, 1×10^{10} cells were harvested 48 h after seeding when cultures were in the exponential growth phase. After discarding the culture medium, cells were detached mechanically by scrapping with 10 ml of cold phosphate-buffered saline (PBS; Nacalai Tesque, cat. no. 05150-45), and the cell solution was washed twice by centrifugations at 1200*g* and 4°C for 3 min. The cell pellet was immediately placed in liquid nitrogen and stored at −80°C until use.

HeLa cells

HeLa cells (JCRB Cell Bank, cat. no. JCRB9004) were cultured in a 15-cm dish to 90% confluence in Dulbecco's Modified Eagle Medium (DMEM, high glucose, Nacalai Tesque, cat. no. 08459-64) containing 10% (v/v) fetal bovine serum (FBS; GIBCO, cat. no. 10437-028) and 1% penicillin/streptomycin (Nacalai Tesque, cat. no. 26253-84) at 37°C in a 5% CO_2 incubator. The cultured cells were detached using trypsin-EDTA (Nacalai Tesque, cat. no. 32777-44) at 37°C for 5 min, and 1 ml of PBS (Nacalai Tesque, cat. no. 05150-45) was added. Cells were collected by centrifugation at 1200*g* and 4°C for 3 min; the cell pellet was stored at −80°C until use.

Primary human adipose-derived stem cells

Human adipose-derived stem cells (hADSCs) obtained from healthy donors were purchased from Lonza Inc. (Lonza Inc. cat. no. PT5006). Cells at passage two were expanded thrice in culture (i.e. to passage 5) in DMEM/F-12 medium (Wako Pure Chemical Industries, Ltd cat. no. 042-30795) supplemented with 10% (v/v) FBS, 100 U/ml penicillin, and 100 μg/ml streptomycin (basal medium). When the cells reached 80% (v/v) confluence, adipogenic differentiation was induced with insulin (10 μg/ml, Sigma, cat. no. I9278), dexamethasone (1 μM, Nacalai Tesque, cat. no. 11107-64), 3-isobutyl-1-methylxanthine (50 μM, Nacalai Tesque, cat. no. 19624-44), and rosiglitazone (5 μM, Wako, cat. no. 180-02653; differentiation medium). hADSCs were cultured for 21 days in basal medium ("pre-adipocytes") or differentiation medium ("adipocytes") with twice-a-week medium change. All cells were incubated at 37°C in a 5% CO_2 humidified incubator. A total of 1×10^8 cells were harvested by scraping, rinsed three times with ice-cold PBS, and centrifuged at 1200*g* at 4°C for 3 min. The cell pellets were stored at −80°C until use.

ORF clones

The initial HaloMS testing and comparison with protein array system data were performed using five pIX-HALO expression clones (AT1G32640, AT1G71930, AT3G62420, AT5G28770, and AT5G65210), as described previously [16–18]. To test human cultured cells, human pENTR-ORFs clones (Dnaform, National Institute of Technology and Evaluation) in the Gateway-compatible entry vector were recombined using Gateway LR clonase II (Thermo Fisher Scientific) into the pIX-Halo: ccdb destination vector [17, 19].

Preparation of plant proteins

Protein isolation from mature leaves of *A. thaliana* wild-type Columbia was performed as previously described [20] with slight modifications. Briefly, 15-day-old mature leaves were frozen in liquid nitrogen, ground using MB1200 multi-beads shocker (Yasui Kikai), and homogenized in an extraction buffer (50 mM Tris–HCl pH 7.5, 150 mM NaCl, 10% glycerol, 2 mM EDTA, 5 mM dithiothreitol, 2% (v/v) IGEPAL CA-630, 2 mM sodium molybdate, 2.5 mM NaF, 1 mM phenylmethylsulfonyl fluoride, 1 mM sodium orthovanadate, and 1 Roche cOmplete ULTRA tablet/50 ml solution). After incubation at 4°C for 30 min with gentle rolling, samples were sonicated using a Bioruptor (UCD-250, Cosmo Bio) with high mode for 10 s sonication plus 10 s without sonication (interval) as one set. After 10 sets as above, the crude protein solution was incubated at 4°C for an additional 30 min. Phenylmethylsulfonyl fluoride was added for a final concentration of 2.44 mM. The samples were centrifuged at 15 310*g* and 4°C for 10 min. The supernatant was transferred to a new tube, and the centrifugation step was repeated twice more. The protein in the final sample was quantified using the Qubit Protein Assay Kit with the Qubit version 2.0 Fluorometer (Thermo Fisher Scientific).

Protein extraction from human cells

Proteins from human cell lines were extracted in IP lysis buffer (Thermo Fisher Scientific, Pierce IP Lysis Buffer, cat. no. 87788) containing protease inhibitors (cOmplete ULTRA tablet, Sigma-Aldrich, cat. no. 5892791001) and phosphatase inhibitors (PhosSTOP tablet, Sigma-Aldrich, cat. no. 4906837001). The buffer was added to cell pellets obtained from 1×10^{10} cells, and the samples were agitated at 4°C for 30 min. The cell lysate was then centrifuged at 18 000*g* at 4°C for 30 min, and the supernatant was collected. Protein concentration in the extract was determined using a BCA protein assay kit (Thermo Fisher Scientific, cat. no. 23225) and adjusted to 2 μg/μl using the IP lysis buffer.

First-generation HaloTag ligand plate assay

The initial HaloMS was performed using a commercially available HaloTag ligand plate (Promega, HaloLink 96 Well Plate, cat. no. CS180802). For HaloMS plate preparation, 500 ng of HaloTag fusion ORF plasmid were translated, and the proteins were expressed in a HaloTag ligand plate at 30°C for 2 h, using the TNT wheat germ system (Promega, cat. no. L4140). The reaction mixture contained 9.5 μl of nuclease-free water, 12.5 μl (50%) of wheat germ extract, 1.0 μl of 10× TNT buffer, 0.5 μl of amino acid mix without methionine, 0.5 μl of amino acid mix without lysine and cysteine (Promega, cat. no. L4140), 0.5 μl (50 unit) of T7 RNA polymerase, and 0.5 μl of RNase inhibitor (Promega cat. no. N2615). The wells were washed containing PBS with 0.1% (v/v) Tween 20 (PBST) and incubated with a human cell lysate containing 50 mg of proteins per 25 μl IP lysis buffer at 4°C for 2 h to allow for protein interaction. The wells were washed in PBST and MS-grade water and used for on-plate protein digestion.

On-plate protein digestion

One hundred microliters of 50 mM Tris–HCl (pH 8.0) and 500 ng of trypsin/Lys-C mix (Promega, cat. no. V5072) were added to the washed HaloTag ligand plate following the first-generation assay and mixed gently at 37°C overnight to digest the proteins. The digested sample (supernatant) was collected into a new 96-well microtiter plate by multi-channel pipetting, treated with 20 mM tris (2-carboxyethyl) phosphine at 80°C for 10 min, and alkylated using 30 mM iodoacetamide at 25°C for 30 min in the dark. Subsequently, the alkylated sample was acidified with 20 µl of 5% (v/v) trifluoroacetic acid and desalted using a STAGE tip (GL Sciences Inc. cat. no. 7820-11200) according to the manufacturer's protocol. This was followed by drying in a centrifugal evaporator (miVac Duo concentrator, Genevac Ltd). The dried sample was resuspended in 10 µl of 2% (v/v) acetonitrile containing 0.1% trifluoroacetic acid. Finally, 4 µl of samples were used for liquid chromatography–tandem MS (LC–MS/MS) analysis.

LC–MS/MS

Peptides were directly injected into a 75 µm × 12 cm nanoLC nanocapillary column (Nikkyo Technos Co. Ltd Tokyo, Japan) at 40°C and separated over a 60 min gradient (A = 0.1% formic acid in water, B = 0.1% formic acid in 80% ACN) consisting of 0 min 8% B, 48 min 28%, 60 min 65% B at a flow rate of 200 nl/min using an UltiMate 3000 RSLCnano LC system (Thermo Fisher Scientific). Peptides eluted from the column were analyzed using a Q Exactive HF-X or Orbitrap Exploris 480 (Thermo Fisher Scientific) for overlapping window DIA-MS [21, 22]. MS1 spectra were collected in the range of 495–745 *m/z* at 30 000 resolution to set an automatic gain control target of 3e6 and maximum injection time of 55 ms. MS2 spectra were collected in the range of >200 *m/z* at a 30 000 resolution to set an automatic gain control target of 3e6 with an "auto" maximum injection time, along with stepped normalized collision energies of 22%, 26%, and 30%. The isolation width for MS2 was 4 *m/z* and overlapping window patterns in the range 500–740 *m/z* were used as previously reported [23]. MS files were searched against a spectral library of human proteins and HaloTag using Scaffold DIA (Proteome Software, Inc.). The spectral library was generated from the amino acid sequence of the human protein database (UniProt, Proteome ID UP000005640, reviewed, canonical) by Prosit [24, 25]. The Scaffold DIA search parameters were as follows: experimental data search enzyme, trypsin; maximum missed cleavage sites, 1; precursor mass tolerance, 10 ppm; fragment mass tolerance, 10 ppm; and static modification, cysteine carbamidomethylation. The protein identification threshold was set at both peptide and protein false-discovery (FDR) rates <1%. Peptide quantification was performed using the EncyclopeDIA algorithm in Scaffold DIA [26]. For each peptide, the four highest quality fragment ions were selected for quantitation. MS data generated in the current study were deposited in JPOST under the accession number PXD041085 (https://repository.jpostdb.org/).

Pull-down assay

The assay was performed using HaloLink magnetic beads (Promega), according to the manufacturer's recommendations. The corresponding human ORF clones were transferred by Gateway LR recombination into pIX-Halo: ccdB and pIX-3×HA: ccdB destination vectors [17]. Competent bacteria (*Escherichia coli*, strain TOP10, Thermo Fisher Scientific, cat. no. C404010) were transformed with the resulting recombination products. The transformants were selected in liquid Plusgrow II medium (Nacalai Tesque, cat. no. 08202-75) containing 50 mg/ml ampicillin, and plasmid DNA was extracted and purified using the NucleoSpinPlasmid kit (Macherey-Nagel). Bait proteins encoded by pIX-Halo were expressed using the TNT T7-coupled wheat germ extract system (Promega, cat. no. L4140) according to the manufacturer's recommendations. Further, 20 µl of bait proteins (pIX-Halo-ORFs) was mixed and agitated with rotation using 5 µl of Halo magnetic beads in a total volume of 50 µl of PBST at 25°C for 1 h. Subsequently, beads with the HaloTag fusion protein baits were washed with PBST, mixed with 20 µl of prey protein fused with a triple hemagglutinin (3× HA), and agitated with rotation at 25°C for 2 h. Next, the mixture was washed three times with 200 µl of PBST. The washed beads were heated in 20 µl of SDS sampling buffer (Thermo Fisher Scientific, cat. no. NP0007) at 90°C for 5 min. The bait and prey proteins (2.5 µl; 10% input) were used as a loading control for the original protein amount.

Second-generation HaloTag ligand plate assay

HaloTag-PEG-biotin ligand (Promega, cat. no. G859A; 500 pmol in 100 µl of PBS buffer) was added to the wells of avidin-coated plates (Thermo Fisher Scientific, cat. no. 15507) and incubated with moderate agitation for 1 h at 25°C. The plates were then washed thrice with nanopure water and incubated with 200 µl of PBSB for 1 h at 25°C with moderate agitation. Finally, the plates were placed in a clean environment (Labconco, purifier class II biosafety cabinet) for 60 min to dry, and the coated plate was stored at –80°C until use.

Protein–protein interaction AI analysis

AlphaFold (version 2.1.1) was downloaded from https://github.com/deepmind/alphafold onto Ubuntu 20.04 comprising 28 cores, 128 GB memory, 2 NVIDIA RTX A5000, and 3 4TB SSD [14]. The list of human and Arabidopsis protein sequences was obtained from the UniProt database [27] and The Arabidopsis Information Resources (on 11 July 2022 at https://www.uniprot.org/, https://www.arabidopsis.org/), respectively. AFM was performed as the default implementation condition. AFM provides five model confidence scores, the interface-predicted template modeling (ipTM) scores. A weighted combination of pTM and ipTM was used to compute the interaction confidence, with reference to previous reports, where model confidence = 0.8 × ipTM + 0.2 × pTM [14]. The mean and SD values were calculated based on these five scores. *P*-values were obtained by comparing pairs of test sets. HaloTag pairs that differed at $P < .05$ were considered positive. LocalColabFold (version 1.4.0) was downloaded from https://github.com/YoshitakaMo/localcolabfold into Ubuntu 20.04 comprising eight cores, 64 GB memory, NVIDIA GeForce RTX 3090, 8TB SSD, and 12TB HDD. ColabFold (version 1.3.0) was downloaded by running the local ColabFold script [11]. In addition, AlphaFold2-multimer-v2, a trained model parameter of AFM, was downloaded from https://github.com/deepmind/alphafold. LocalColabFold were performed using the following settings:—amber,—use-gpu-relax,—templates,—num-recycle 3,—random-seed 0,—num-models 1,—model-order 1,—model-type AlphaFold2-multimer-v2.

Uniform manifold approximation and projection and hierarchical clustering

Protein homology determined in CFM was visualized based on the metagenome visualization method [28]. First, each sequence was embedded as a 256-dimensional vector using localcolabfold [11] and visualized using the umap version 0.5.3 and scanpy version 1.9.3 Python packages [28, 29], where dimensionality reduction was applied directly to the embedding vectors (use_rep="X" in scanpy.tl.umap) with default parameters (3 or 4-nearest-neighbor graph via approximate Euclidean distance, uniform

manifold approximation and projection (UMAP) min_dist = 0.5) [30]. Hierarchical clustering and dendrogram generation were performed using embedding vectors obtained from ColabFold using the hierarchical clustering algorithm ("single" method, "euclidean" metric) from the scipy version 1.9.3 package in Python [31, 32].

Statistical analysis

Data were compared using Student's *t*-tests for most analyses if no specification was indicated. Statistical analysis of data was performed using BellCurve for Excel (Social Survey Research Information Co. Ltd.) and https://www.socscistatistics.com/tests/ . The employed quantification methods and web-based tools are described in the legend of each relevant figure/table.

Results

Development of HaloMS: a HaloTag-based affinity purification–MS system

In conventional APMS, tagged protein-encoding sequences are first transformed into a cell and retrieved by binding to ligand-capturing beads in a reaction tube. The interacting protein complex is then pulled down from a biological sample, and the proteins are identified one by one via MS. We modified this system by incorporating a high-affinity capture tag, the HaloTag (Promega), to immobilize nascent proteins expressed with the cell-free system in a 96-well plate (Fig. 1) [16]. In the assay, *in vitro* expressed and translated HaloTag-labeled proteins are covalently bound to a small chemical ligand chloroalkane that coats the plate surface. This allows for rapid capture of newly expressed fusion proteins and eliminates the cumbersome process of purifying individual proteins on ligand beads, avoiding the possibility of compromising protein functional integrity.

To test the activity of the HaloTag-fused protein bound to the chloroalkane ligand-coated 96-well plate, we detected protein–protein interactions in the model organism *A. thaliana* using a query protein set previously used in a similar protein immobilization method with HaloTag [16]. After HaloTag-based affinity purification MS (hereafter, HaloMS), we assessed the expression of five query TFs, namely: AT1G32640, AT1G71930, AT3G62420, AT5G28770, and AT5G65210, based on the MS intensity. The MS intensities of the query proteins were between 8.41 and 9.52 ($\log_{10}$ scale), indicating that the levels of protein produced from the five ORFs were consistent (Supplementary Table S1). We then probed the lysate of Arabidopsis leaves containing 9999 identified proteins, including splice variants, for complex interactions with the five TFs, which were translated and expressed using the wheat germ expression system and bound to HaloTag ligand plates (Supplementary Table S2). We set the protein identification threshold for the peptide and protein at an FDR of <1%. An interaction was scored as positive when the log ratio of the signal intensity obtained from the negative control (i.e. 33 kDa HaloTag protein) was >1 standard deviation (SD) above the median. By using this criterion, HaloMS for the five query proteins produced a dataset of 740 interactions and 306 proteins (Supplementary Table S3). We then calculated the overlap between the interactions revealed by HaloMS and those in a published protein array dataset containing 1544 interactions and 1234 proteins that use the same HaloTag protein immobilization method [16]. HaloMS recapitulated a statistically significant 10 of 709 protein array interactions in the shared protein space ($P < .00001$, χ^2 test with Yates' correction; hereafter "χ^2"). This significant association between the two approaches suggested that these interactions are not affected by the test method (HaloMS versus protein array).

We next investigated the degree to which the identified protein–protein interactions could be classified using gene ontology (GO) functional annotation terms to assign a higher probability of biological relevance. The resultant enriched terms for all target proteins interacting with the five query proteins, as identified by HaloMS, are provided in Supplementary Table S4. Except for "mRNA binding," "RNA binding," and "binding," a significant proportion (*P*-value > 1.64E-14) of HaloMS targets of the five TFs tested were related to the enzyme and carbon-binding activity. By contrast, according to the enrichment analysis based on the protein array data [16], most targets were related to transcription, "binding," and "protein binding" (Supplementary Table S5). The considerable differences in GO enrichment of the protein–protein interactions detected using the two methods could be explained by a wider dynamic range of the expressed proteins in the cell lysate used in HaloMS compared to that in the input used for protein arrays (Supplementary Fig. S1A and B).

We then determined the overlap of interactions between the HaloMS dataset and the literature-curated interactions in the BioGrid database (https://thebiogrid.org/). HaloMS recapitulated a statistically significant 1 of 79 BioGrid interactions in the shared protein space ($P = .019044$, χ^2 test). This moderate overlap may be associated with differences in, for example, the biochemical experiments (fusion tags, etc.) leading to the detection of different subsets of true interactions in the specific assay [33]. Alternatively, this could be due to the relatively small number of Arabidopsis interactions available in the dataset (81 804 in BioGrid on 27 August 2021). Collectively, the low overlap suggested that HaloMS is highly complementary to the public dataset for Arabidopsis, generated primarily using a yeast two-hybrid (Y2H) system. Furthermore, HaloMS has several advantages. For example, it does not require the preparation of a large number of clones and allows for *in vivo* screening of protein complexes in a variety of physiological states. It also avoids the cumbersome stages of individual transgenic cell line preparation and carrier bead purification of conventional APMS, as the query proteins are synthesized cell-free and captured on plates, on demand.

HaloMS reliably detects protein complex interactions in cultured human cells

We tested HaloMS for the detection of protein–protein interactions in human cells. We transferred a set of 16 Gateway pENTR ORF clones, including *ATF3*, *GATA2*, and *FOS*, which encode well-studied human protein-coding genes into a Gateway convertible HaloTag expression vector. We then *in vitro* expressed and captured the proteins on a HaloLink 96-well plate (see Materials and methods section). The intensity of the expressed 16 query proteins was 7.45–9.67 ($\log_{10}$ scale), indicating consistent levels of produced and captured proteins (Supplementary Fig. S1C).

These 16 proteins were then used to probe the mixed lysates of HeLa and HEK293F cells, which were made by culturing both cells separately and lysing each cell culture containing 7130 proteins to detect their protein complex interactions (Supplementary Table S6). The same criterion for FDRs of peptide, protein, and interaction scoring as positive was applied as with the initial Arabidopsis experiment. Thus, the generated HaloMS dataset contained 3302 interactions and 913 proteins (Supplementary Table S7). We determined its interaction overlaps with the 1 057 434 interactions documented in BioGrid (on 29 June 2022, at https://thebiogrid.org/). HaloMS recapitulated a statistically significant 183 of 4220 literature-curated

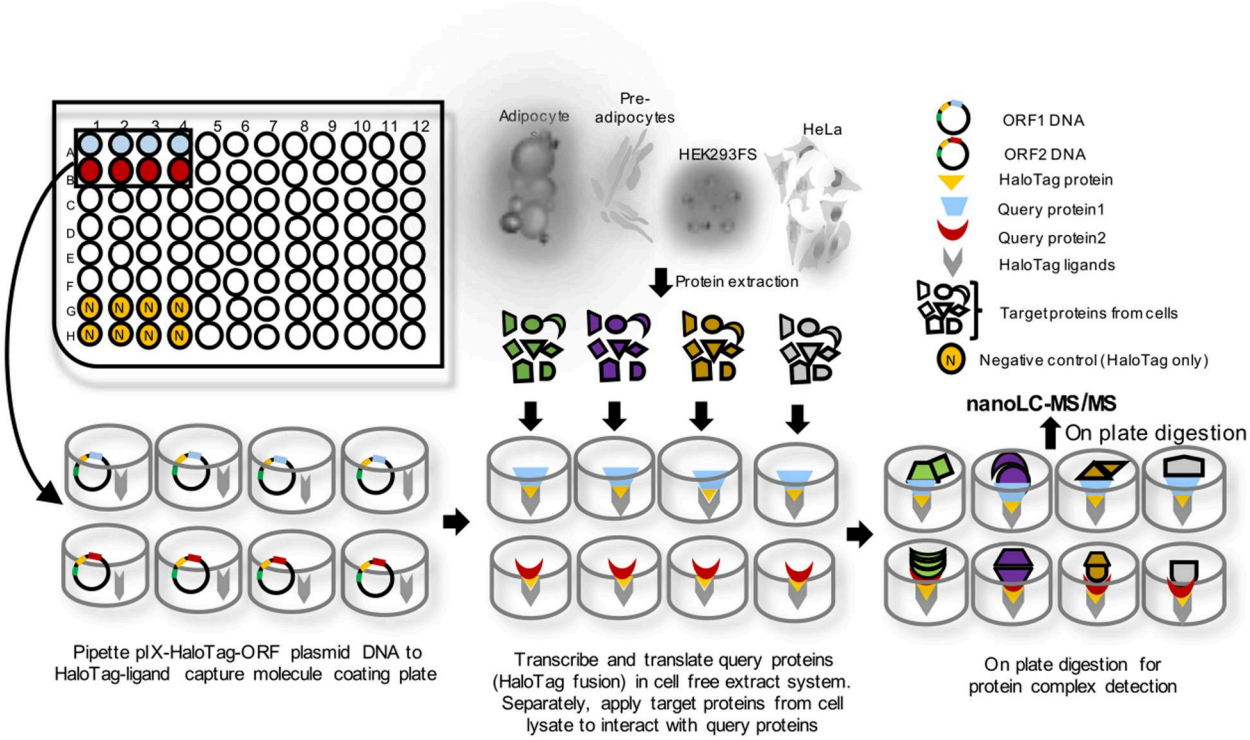

Figure 1. Overview of the HaloMS assay. HaloTag-ORF plasmid DNA is placed in a well of a HaloTag-ligand-coated plate. The addition of coupled transcription–translation reagent results in protein expression and probe protein capture, accordingly. Target protein lysates from various cells can be added to the wells, and protein complex interactions are detected by LC–MS/MS.

interactions in the shared protein space ($P < .00001$, χ^2 test). Hence, the *in vitro* expressed proteins immobilized on the 96-well plate engaged in specific interactions with the thousands of proteins in the lysate. Thus, the interaction overlap analysis indicated that HaloMS yields reliable protein–protein interaction data, statistically indistinguishable from the literature-–curated interaction dataset.

In vitro pull-down assay validates the HaloMS dataset

A critical concern for any new technology is the quality of the obtained data. Accordingly, we systematically validated a subset of interactions from a new dataset in a second interaction assay, that is a pull-down assay [33–36]. As a caveat, no assay can detect all protein–protein interactions, and each has a different interaction-detection profile [35, 37]. To validate the HaloMS dataset generated for the human cell lines, we selected 32/3302 pairs of the detected interactions, corresponding to approximately 1% of all interactions in the dataset, and we subsequently tested them using a pull-down assay (Fig. 2A and Supplementary Fig. S2 and Supplementary Table S8) [15]. We compared the obtained data with benchmarking data generated previously [16, 38]. We benchmarked the new data against (i) a set of known positive interactions (PRS: 49 pairs) and (ii) a set of randomized pairs of interactions that, to date, have not been supported by experimental evidence (RRS: 69 pairs). Of the 32 pairs, 7 (21.8%) scored positive (Fig. 2B and Supplementary Table S8). This proportion differed significantly from that for PRS and RRS ($P = .0003$ and .0041, respectively, Fisher's exact test; Fig. 2B).

The validation rate for the HaloMS positive interactions suggested that the obtained dataset reflects a different aspect of the well-documented interactions reported in the literature and detected by direct protein–protein interaction analysis. In contrast, HaloMS yielded protein interaction data of a reliable quality that was statistically different from the random reference set. As anticipated, the reproducibility (21.8%) of HaloMS, which includes indirect interactions in a protein complex, was low compared with PRS, which is composed of binary protein–protein interactions. Taken together, the analysis of the pull-down assay data against the benchmark dataset indicated that HaloMS can be used to reveal interactions that are statistically independent of binary interaction data, literature-curated interactions, and random protein pairs.

AI analysis supports protein–protein interactions detected by HaloMS

Although we were able to validate the novel HaloMS protein–protein interactions using a pull-down assay, we next asked whether these interactions could be supported by another type of analysis, for example, AI predictions based on the amino acid sequence. Accordingly, we used AlphaFold-Multimer (hereafter AFM) to predict and validate protein–protein interactions detected by HaloMS and the pull-down assay. To determine the FDR of the human cultured-cell HaloMS dataset, we measured the AFM sensitivity and background by benchmarking against a set of controls supported by the pull-down assay, PRS, and RRS. This allowed subsequent interpretation of the retest rate of the new dataset by HaloMS, which was also validated by the pull-down assay, in light of these benchmarking data. To measure the sensitivity and background of the AFM assay, we benchmarked it against (i) a set of 48 well-documented interactions from the literature (afmPRS) and (ii) a set of 71 random interactions from the published dataset (afmRRS) [15]. As a new dataset, we tested 32 pairs from the HaloMS dataset from human cultured cells, which were supported by the pull-down assay (Supplementary Table S8). For each dataset, a negative control set with HaloTag-only proteins was also analyzed by AFM to obtain a validity score

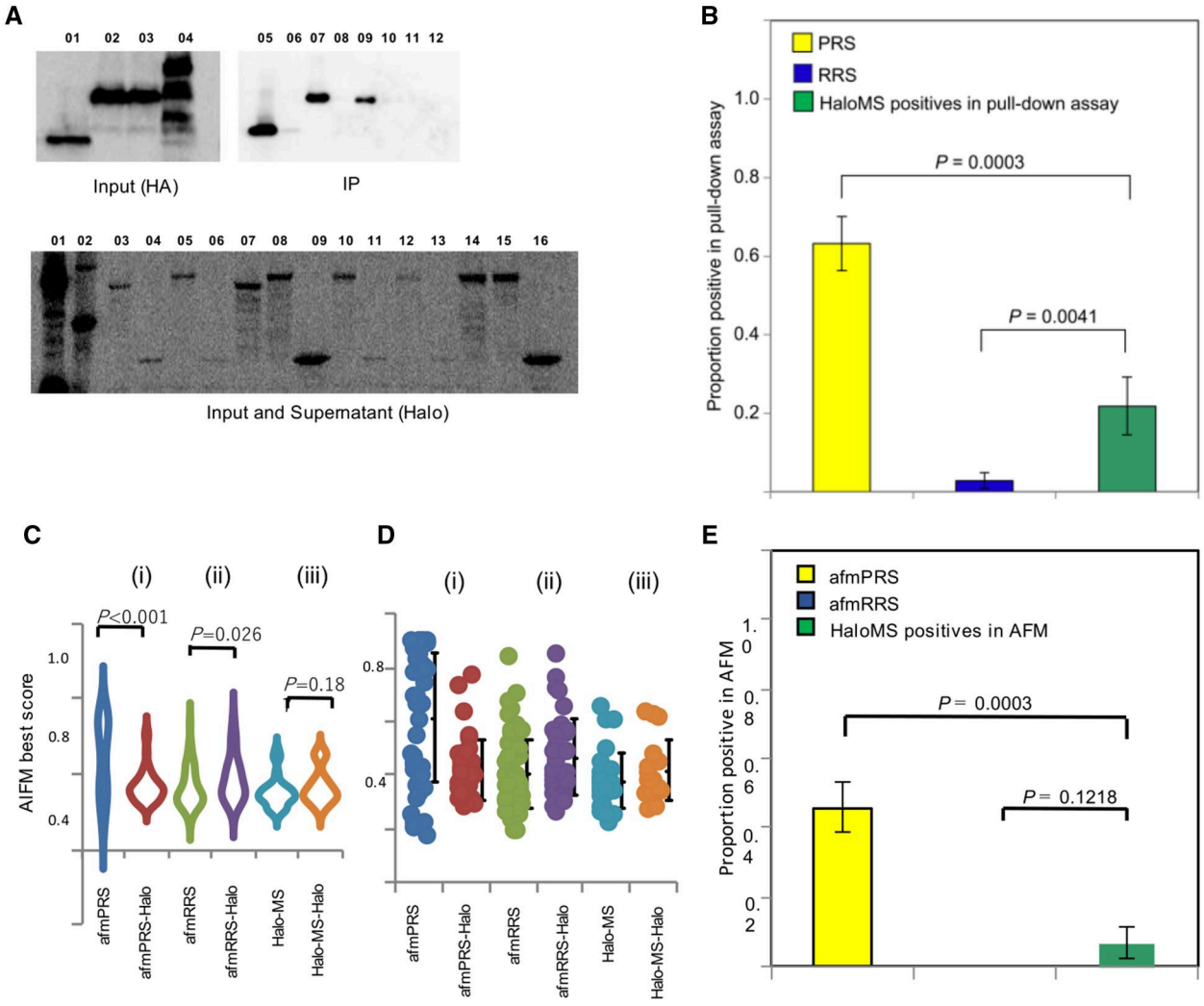

Figure 2. Quality of the HaloMS assay. (A) HaloMS protein–protein interactions are recaptured by pull-down assay. Representative results of the in vitro pull-down experiments for Halo alone, Halo-tagged, and 3×HAtagged proteins; 5% of input is detected using an anti-HA antibody (labeled "Input (HA)," lanes 01–04); copurified Halo-tagged and 3×HA-tagged proteins are detected using an anti-HA antibody (labeled as "IP," lanes 05–12), and 5% of supernatant after binding to HaloLink magnetic beads and input for HaloTaglabeled proteins (labeled "Input and Supernatant (Halo)," lanes 01–16) are tested to determine the relative amounts of Halo-tagged and HA-tagged proteins, as well as the binding efficiency to HaloLink magnetic beads. (B) Proportion of positive scoring pairs by pull-down assay. The histogram shows the proportion of positive scoring pairs from the PRS dataset, the set of randomized sample pairs from reference 15, and the set of sample pairs positive in HaloMS. Error bars represent the SE of the proportion. The proportion of positive PRS pairs (left bar) is significantly higher than that of RRS pairs (middle bar; $P < .0001$, Fisher's exact test). (C) AI-based prediction of protein–protein interactions is detected by HaloMS and pull-down assay. The violin plot distribution of the highest scores is obtained by AlphaFold-Multimer (AFM) across six datasets (Supplementary Table S9). (D) Dot plot distribution showing the average, first-quartile, and third-quartile AFM best scores for six datasets: afmPRS, afmPRS-Halo, afmRRS, afmRRS-Halo, HaloMS, and HaloMS-Halo. (E) AI-based prediction of protein–protein interactions. The proportion of positive scoring pairs from the afmPRS dataset, the set of randomized sample pairs from reference 15, and the set of sample pairs positive in HaloMS are shown. Error bars represent the SE of the proportion.

(Supplementary Table S9). The distribution of AFM best scores from the three datasets validates the characteristics of each (Fig. 2C and D).

In the benchmark experiments for the afmPRS (i) and afmRRS (ii) sets, the reproducibility of afmPRS differed significantly (*t*-test, $P < .001$) from the negative-control pairs; those of afmPRS-Halo and afmRRS were not significantly different, as expected for a negative control; afmRRS-Halo scored higher than afmRRS ($P = .026$). However, the distribution of recaptured HaloMS interactions was statistically indistinguishable from HaloMS-Halo (Fig. 2C), (iii) by AFM ($P = .187$) despite 21.8% reproducibility in the pull-down assay. We averaged the five scores to determine the model confidence obtained by AFM to compute the significance of differences between each protein pair and the HaloTag-only control pairs, as in the pull-down assay (for a "digital" pull-down assay; Supplementary Table S10). In benchmark experiments using (i) and (ii) reference sets, 22 of 48 afmPRS pairs (46%) and 0 of 55 afmRRS pairs (0%) scored significantly higher than the HaloTag-only controls, defining the sensitivity and background of the digital pull-down assay. Of the 30 pairs of HaloMS, the scores for 2 (6%) were significantly positive in the digital pull-down assay, statistically different from the afmPRS result ($P = .0003$, Fisher's exact test; Fig. 2E). This reproduction rate was similar to the previously obtained pull-down result, including all positive pairs (6/30 pairs, 20%, Supplementary Table S10), and was not significantly different from the afmRRS set ($P = .1218$, Fisher's exact test; Fig. 2E). The afmPRS, afmRRS, and HaloMS scores by AFM ranged from 0.500 to 0.652, 0.3061 to 0.3473, and

0.3465 to 0.42369, respectively, estimated using a 95% confidence interval. These results support the findings of the pull-down evaluation, that is the HaloMS data represent complex interactions, not just binary interactions, with few directly interacting pairs.

HaloMS TF interactome network for adipocyte differentiation response

As a proof-of-principle, we used HaloMS to generate an adipocyte interactome dataset. An adipocyte cell stores triglyceride droplets and helps convert lipids into energy. In addition, an adipocyte can secrete hormones and other effector (bio) chemicals that play important roles in metabolism control, including cell turnover [39], adaptation [40–42], and endocrine functions [43, 44]. These adipocytes, which are associated with metabolic diseases through different pathways, can be used for mapping protein interaction networks by applying HaloMS technology; it may be possible to construct a centralized signaling pathway. We set 17 human proteins as query proteins, including 9 TFs (ATF1, ATF2, ATF3, ATF5, GATA2, CEBPA, CEBPD, CREB1, and PPARA) with known roles in adipocytes that mediate differentiation from pre-adipocytes. We also developed an alternative capturing technique to produce query proteins on the 96-well plate; the HaloTag ligand concentrations were lower than those used in the first-generation HaloMS to lower the cost of the HaloTag capturing plate. This also addressed the issue of the original HaloTag plate ceasing to be commercially available. We assessed the protein-capturing efficiency of four query proteins (GATA2, ATF3, CTNNA1, and HaloTag) based on MS intensity. The intensities of the query proteins were 7.78–9.79 ($\log_{10}$ scale), indicating that consistent protein levels were being captured (Supplementary Fig. S3A, red), with the capture levels similar to those of first-generation HaloMS (8.14–9.62, $\log_{10}$ scale MS intensity, Supplementary Fig. S3A, light blue).

Considering that most TFs function as part of a complex [30, 45–47], we expected that the identification of additional proteins that interact with TFs in adipocytes would reveal novel adipocyte differentiation signaling components and connections among the already-known signaling pathways. Furthermore, for several human proteins in a set of 17, including 9 TFs with known roles in adipocytes, some interactions were already identified in the BioGrid database (thebiogrid.org). To address systematically the adipocyte-specific interaction network, we probed four cell types (adipocyte, pre-adipocyte, HeLa, and Human Embryonic Kidney cell as HEK) with the 17 query proteins using HaloMS. Using these cell types, the technique was applicable to heavily used cancerous, multi-generation, passaging cultures, such as HEK and HeLa, as well as to primary cells that presumably reflect physiological conditions, such as adipocytes.

To confirm protein-expression levels in the input sample (the four types of cells), we assessed global protein expression across the cell types using LC–MS/MS after proteins were digested, alkylated, and acidified. Among the 9405 proteins with similar expression ranges, we identified 7258 in adipocytes, 7552 in pre-adipocytes, 8149 in HeLa cells, and 8225 in HEK cells (Supplementary Fig. S3B and C, Supplementary Table S11). The expression level of over 65% of proteins was similar in the different cells (Supplementary Fig. S3D); 6% of proteins fell outside the commonly expressed protein set. The intensities of query proteins immobilized on the 96-well plate and used for the screen were 7.64–9.34 ($\log_{10}$ scale), indicating that consistent protein levels were produced from the 17 ORFs in the assay (Supplementary Fig. S3E, Supplementary Table S12). We scored a positive interaction when the average intensity from multiple datasets was >1 SD above the median of the ratio, a 3.07-fold difference on average, obtained with only the 33-kDa HaloTag proteins as the negative control. Pairs with a negative control of 0 and no ratio available were considered positive at 1 SD from the median of the intensity distribution. According to these criteria, the HaloMS screen of the 17 human proteins in the 4 different cell types produced a TF-HaloMS dataset containing 6512 interactions and 1690 proteins (Fig. 3A and Supplementary Table S13). Over 70% of the interactions were cell type specific, unlike the expression of input proteins. Only 96 of 6512 interactions (1.5%) were commonly detected in all examined cell types. Comparison of the interaction number and protein expression levels revealed that for the adipocyte-specific interactions (1201 pairs), low-expression proteins were involved in the highest number of interactions (488 pairs, 40%; Fig. 3B). Other cell-specific protein–protein interactions tended to be detected in proteins with relatively high expression levels, for example, in HeLa and HEK cells (Fig. 3B). We then determined the overlap of interactions between the TF-HaloMS dataset and literature-curated interactions in BioGrid (thebiogrid.org). The HaloMS screen of the four different cell types recapitulated a statistically significant 443 of 5974 BioGrid interactions in the shared protein space ($P < .00001$, χ^2 test; Supplementary Table S13). The significant overlap suggested that HaloMS using first- and second-generation 96-well plates was an approach that is statistically reliable, not only to interactions described in the literature but also to novel protein–protein interactions.

Biological validity of the TF-HaloMS dataset

We next obtained evidence of the biological validity of the TF-HaloMS dataset and new insights into signal transduction during adipocyte differentiation. To this end, we validated several of the biologically interesting novel interactions using AI. Specifically, we predicted protein–protein interactions using ColabFold Multimer (CFM), offering an accelerated prediction of protein structures and complexes with a 40- to 60-fold faster search than AFM [11]. To estimate the FDR of the HaloMS adipocyte dataset and pre-adipocyte-specific interactions, we determined the sensitivity and background of CFM using the PRS dataset and compared the findings to the AFM results (Fig. 3C and D). Due to limited machine resources, we used one CFM score to analyze significant differences for 49 well-documented interactions from the literature: (cfPRS) [16], HaloTag-only negative controls (cfPRS-Halo), and the adipocyte- and pre-adipocyte-specific interaction dataset from HaloMS (Fig. 3E–F and Supplementary Table S14). The distribution of the dataset CFM scores validated the traits of each dataset. In the benchmark experiments for the datasets for: (i) afmPRS and the negative control (afmPRS-Halo) and (ii) cfPRS and the negative control (cfPRS-Halo), the reproducibility of cfPRS, as well as the afmPRS, was significantly different (t-test, $P < .001$) from that of negative-control pairs (t-test, $P < .001$; Fig. 3C). Of the 1201 and 1110 interactions selected based on all adipocyte- or all pre-adipocyte-specific interactions in the TF-HaloMS dataset, 1093 and 1028, respectively, were validated using CFM (Supplementary Table S14). The TF-HaloMS dataset evaluated by CFM included 86 known adipocyte-specific pairs and 71 known pre-adipocyte-specific pairs, and the distribution of the CFM scores for these and novel interactions were not significantly different ($P = .027$; Fig. 3E, Supplementary Table S14). This overlap between known and novel interactions suggested that the TF-HaloMS dataset of adipocyte- and pre-adipocyte-specific interactions are a reliable protein–protein interaction dataset.

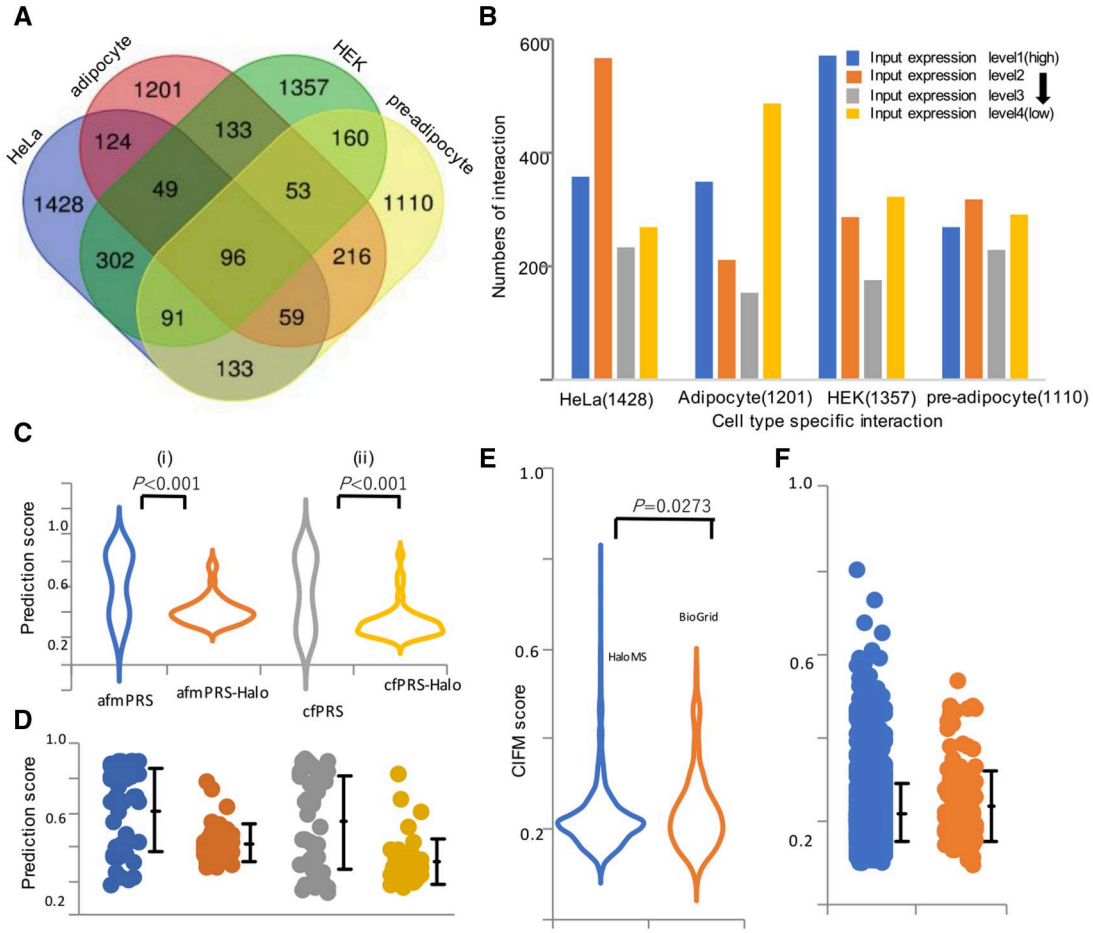

Figure 3. Interactome profiling in different cell types. (A) Venn diagram of the interactome status of four cell types. Over 70% of the interactions are cell-type specific, unlike the input protein expression. Venn diagram generated using software available at https://bioinformatics.psb.ugent.be/webtools/Venn/. See also Supplementary Table S13. (B) Comparison of interactions and protein expression levels. The highest number of interactions for low-expression proteins is detected among adipocyte-specific interactions. (C) Violin plot distribution of AFM and CF scores in the PRS dataset showing the validity of the interaction assessment (Supplementary Table S14). (D) Dot plot distribution of the average, first-quartile, and third-quartile AFM and CF scores for the datasets. (E) Violin plot distribution of CF scores for the HaloMS dataset validating the interaction assessment of each dataset of adipocyte- and preadipocyte-specific interactions. (F) Dot plot distribution of the average, first-quartile, and third-quartile CF scores for each dataset. In the HaloMS experiments for (iii) novel (left) and known (right) interactions of the adipocyte–pre-adipocyte-specific interaction dataset, the distribution of the CF score overlaps in the novel versus known-pair comparison. See also Supplementary Table S14.

HaloMS-based exploration of network communities in adipocytes

During protein network analysis, it is important to identify densely interacting components that function together, that is the intra-network community [48]. Accordingly, we analyzed the pathways activated in the four different cell types. More specifically, we used average-linkage clustering to identify intra-network communities in adipocytes and investigated their biological relevance. Pathway analysis for the cell type-specific pairs of proteins (Supplementary Fig. S4 and Supplementary Table S13) identified 83 communities containing more than four proteins in the HaloMS dataset as significantly cell type-associated proteins (Supplementary Fig. S4, Supplementary Tables S15 and S16). The adipocyte-specific interactions included mitochondria-associated neurodegenerative disease pathways (including Alzheimer's, Parkinson's, and Huntington's diseases; Supplementary Figs S4 and S5A and Supplementary Table S17) [49]. The activation of these pathways implied that the thermogenesis subnetwork in adipocytes was activated, indicating the biological relevance of the TF-HaloMS dataset. The specific enrichment of the PPAR signaling pathway [50], nonalcoholic fatty liver disease [49], and fluid shear stress and atherosclerosis [51, 52] in adipocytes further supported the validity of the interactions detected in the current study (Supplementary Fig. S5B–D, Supplementary Tables S18–S20). The pre-adipocyte-specific interactions, including a pathway for adrenergic signaling in cardiomyocytes (Supplementary Fig. S6A and Supplementary Table S21) [53], thyroid hormone synthesis (Supplementary Fig. S6B and Supplementary Table S22) [54, 55], AMPK signaling (Supplementary Fig. S6C and Supplementary Table S23) [46], platelet activation (Supplementary Fig. S6D and Supplementary Table S24) [56], peroxisomes (Supplementary Fig. S6E and Supplementary Table S25) [57, 58], and N-glycan biosynthesis (Supplementary Fig. S6F and Supplementary Table S26) [59], supported the correct profiling of interactions by HaloMS.

The relationship between the interacting proteins is a critical question in HaloMS, which was devised specifically to detect protein complexes. Accordingly, we clustered groups of proteins interacting with the 17 protein baits obtained specifically from adipocytes by HaloMS using UMAP, dendrograms, and tree-view clustering according to structure and prediction scores by CFM (Supplementary Fig. S7). In the UMAP dataset of GATA2, a key regulator of adipocyte biogenesis, several same-family proteins

(e.g. G3BP1 and G3BP2), same-pathway proteins (e.g. SF3B3, POLR3A, and IARS1 for transcription and translation, and DHX15, DHX9, RTCB, SNRPA1, and UPF1 for RNA splicing), and subunit proteins of the same complexes (e.g. SSRP1 and SUPT16H for the FACT complex) were relatively closely clustered, as anticipated (Supplementary Fig. S8). The GTP-binding protein regulator DRG1, which had the highest CFM score (0.596 out of the maximum score of 1.0) among adipocyte-specific protein–protein interactions with GATA2, was located close to DPYSL2, which plays a role in neuronal development and polarity, both in the dendrogram as well as UMAP analysis (Supplementary Figs S7–S9 and Supplementary Table S27). RNA-binding proteins LUC7L2 and RALY were also located in close proximity to UMAP (Supplementary Fig. S8). Furthermore, CUL3, a member of the proteasomal degradation pathway, was positioned close to DRG1, also known as NEDD3, in both dendrogram and UMAP analysis; it has been reported to interact with the cullin-associated protein, CAND2 (Supplementary Figs S7–S9 and Supplementary Table S27) [60–62]. They, DRG1, NEDD3, and CAND2, may be involved in the cullin-RING ubiquitin ligase network as a complex [63].

Hence, the clustering analyses based on AI inference of HaloMS data may contribute to the knowledge of how multiple TF complexes form to control adipocyte biogenesis. They may also help to illuminate additional hidden features of the protein network that are regulated by previously unknown pathways.

HaloMS reveals novel interactions of heterogeneous nuclear ribonucleoproteins

The TF-HaloMS dataset includes protein interactions among TFs from several signaling pathways that are specific to the differentiated adipocyte (Supplementary Figs. S4, S9A, and Supplementary Table S28). Interconnections between the long-noncoding RNA (lnRNA) and heterogeneous nuclear ribonucleoproteins (hnRNPs) in adipocyte biology have been documented [64]. More than 100 brown-adipocyte lnRNAs, including lncBATE1, were identified as requirements for the establishment and maintenance of brown-adipocyte identity. In addition, they were found to interact with hnRNPU, which is also required for brown adipogenesis [65, 66]. In the TF-HaloMS dataset generated in the current study, several key regulators of adipocyte biogenesis, namely, ATF2, ATF5, CREB1, PPARA, GATA2, CEBPA, and CEBPD, interact with HNRNP-family proteins, HNRNPL and HNRNPU (Supplementary Fig. S4 and Supplementary Table S28). These interactions were independently detected specifically in the adipocyte, except for the PPARA–HNRNPL interaction, which was inferred in prostate cancer cells by the affinity-capture-RNA method [67]. HNRNPU interacts with the lnRNA Blnc1, which has been proposed to upregulate brown and beige fat thermogenesis and white fat expansion by interacting with the transcription regulators EBF2 and Zbtb7b and thus to inhibit adipose tissue inflammation [64, 66, 68, 69]. Our findings suggest a new network and role for the HNRNP family, especially HNRNPU, interacting with key proteins as a direct regulator of adipocyte biogenesis and maintenance, permitting crosstalk between lnRNA and cell differentiation. HNRNPU interactions with key proteins, such as TFs, might decide the cell fate by controlling access to the genome to specific lnRNA [65, 68, 70]. This interaction sharing by multiple hnRNPs with a key regulator may provide a route to facilitate the processing of mRNA for cell differentiation signaling.

HaloMS reveals novel interactions of ATP-dependent RNA helicase DEAD box proteins (DDXs)

In the obtained TF-HaloMS network, we identified a subnetwork of ATP-dependent RNA helicases, the DDXs. These proteins are involved in RNA metabolism, from transcription to degradation, and are important players in gene expression [71]. HaloMS revealed novel networks involving DDXs, relative to the previously reported cellular process map [71], extending the known network by the interactions of at least six key regulators, GATA2, ATF3, CEBPD, CREB1, CEBPA, and PPARA (Supplementary Fig. S9B and Supplementary Table S29). As detected by HaloMS, these six key regulators of adipocyte biogenesis interact with DDX17, DDX39B, DDX46, and DDX5 in an adipocyte-specific manner. Consistent with previous observations [72–78], the HaloMS network analysis also revealed molecular links between DDXs (DDX1, DDX41, DDX39B, and DDX17) and key regulators of carcinoma biogenesis (CASP3, CTNNA1, and FOS), although the molecular mechanism of these connections remains unclear. Based on the TF-HaloMS analysis, we propose the existence of cross-regulation of adipocyte differentiation and carcinoma cell turnover signaling via protein–protein interactions between the key regulators and DDXs. Indeed, RNA helicases from the DEAD box family, DDX5 and DDX17, cooperate with hnRNPH/F splicing factors to define cell specificity [79]. The HaloMS interaction analysis also points to a possible connection between DDX17, HNRNPs (H1/H2/U), and key TFs of adipocyte biogenesis (CEBPA, CEBPD, and ATF3; Supplementary Fig. S9C and Supplementary Table S30). As DDX17 functions as a coregulator of master transcriptional regulators of cell differentiation [79, 80], these members might comprise the core network of adipocyte differentiation cooperating with lnRNA.

HaloMS reveals novel interactions of the spliceosome in adipocyte

The TF-HaloMS pathway analysis of adipocyte- and pre-adipocyte-specific pairs of proteins (Supplementary Fig. S4 and Supplementary Table S13) identified major spliceosome-associated proteins in both cases (Supplementary Table S16). Comparing the respective cell-specific protein networks, the adipocyte-specific network (72 edges) had approximately 7-fold more interactions than the pre-adipocyte-specific network (10 edges; Supplementary Fig. S9D and Supplementary Table S31). In the adipocyte-specific network, SRSF3, SNRPD2, SRSF1, DDX46, SRSF7, U2AF1, PUF60, and HNRNPU shared four or more interactions each with adipocyte regulators. SRSF3 protein is a regulator of brown-adipocyte formation [81] and is involved in hepatocyte maturation and metabolism [82]. SRSF1 generates the PPAR gamma isoform and is involved in pre-adipocyte differentiation [83]. Additionally, the early pre-spliceosomal complex members SRSF7, FUS, DDX46, DDX5, HNRNPA3, and HNRNPU [84] were detected in the adipocyte-specific network, suggesting that the spliceosome plays an important role in adipocyte differentiation in complex with key adipocyte regulatory factors [85].

HaloMS reveals novel interactions of diabetic cardiomyopathy network

In the TF-HaloMS dataset of adipocyte- and pre-adipocyte-specific interactions (Fig. 3A and Supplementary Table S13), we identified a subnetwork of proteins involved in the diabetic cardiomyopathy pathway (Supplementary Fig. S10A and Supplementary Table S32). A considerable relationship between adipose tissue and cardiovascular disease has been reported [86,

87], and inadequate adipose tissue function is reportedly responsible for cardiovascular events and heart failure in individuals with obesity, metabolic syndrome, and type-2 diabetes [88]. The HaloMS dataset expanded the known diabetic cardiomyopathy network, relative to interactions previously reported in the literature (Supplementary Fig. S10A and Supplementary Table S32), by at least 39 protein interactions [89]. In this subnetwork, COL3A1 interacts with seven adipocyte regulatory molecules, namely, ATF2, ATF3, ATF5, CEBPA, CREB1, GATA2, and PPARA. According to a previous study, COL3A1 (collagen III protein) is required for adipogenesis with the adhesion G-protein-coupled receptor Gpr56 [90], and the protein also induces the progression of ischemic heart failure [91]. Based on the TF-HaloMS network data, we hypothesize that diabetic cardiomyopathy and adipocyte differentiation occur via protein–protein interactions among three collagen family proteins, eight mitochondrial proteins, and nine master regulators of adipocyte (Supplementary Fig. S10A and Supplementary Table S32); 30% of this diabetic cardiomyopathy subnetwork overlapped with the thermogenesis subnetwork (Supplementary Fig. S10B and C, Supplementary Table S32) composed of mitochondrial proteins. These findings point to a novel potential connection between diabetes-related signaling pathways, such as the PPAR signaling pathway, and energy homeostasis [92].

HaloMS reveals links between endoplasmic reticulum (ER)-phagy, neurodegenerative disease, and adipocytes

Among interactions of a group of proteins not mapped by the aforementioned pathway analysis in the TF-HaloMS dataset ($P > .001$, Supplementary Fig. S4 and Supplementary Table S16), we found a novel subnetwork of proteins involved in ER-phagy, neurodegenerative disease [93], ER-phagy, neurodegenerative disease, and myelodysplastic and master regulators in adipocytes (Supplementary Fig. S10D and Supplementary Table S33). The relationships between adipose tissue, diabetes, and ER-phagy have been previously explored [94, 95]. UFL1, which ubiquitinates UFM1, interacts with DDRGK1 and is involved in ER-phagy and also in the development of diabetes [96, 97]. This subnetwork includes DNAJC13 (Parkinson's disease) [98], MAP7D1 (autism) [99], EIF2B4 (leukoencephalopathy with vanishing white matter) [100, 101], TREX1 (Aicardi–Goutières syndrome, systemic lupus erythematosus, familial chilblain lupus, Cree encephalitis, cryofibrinogenemia, and retinal vasculopathy with cerebral leukodystrophy) [102], and LUC7L2 (myeloid malignancies) [103, 104]. The related genes are involved in neurodegenerative and myelodysplastic diseases, providing new insights as novel network members of the adipocyte master regulator.

Overall, the protein–protein interactions described in the case studies above generated a wide range of hypotheses on the mutual modulation of adipocyte-specific signaling pathways. These and more can be derived from the HaloMS dataset, which demonstrates the power of HaloMS to uncover unexpected network complexity even in characterized signaling systems.

Discussion

We have described the development of HaloMS, a HaloTag-based APMS methodology, and its application to map interactions for human TFs involved in adipocyte regulation. Validation in a benchmarked interaction assay demonstrated that the quality of the TF-HaloMS dataset is comparable to that of published interaction datasets. The quality of the data was further confirmed by the validation of a subset of interactions, including benchmarked interaction, using AI. Validation in a pull-down experiment supported the notion that the HaloMS positives were due to direct protein–protein interactions rather than indirect protein–protein interactions. Significant overlap with previously described interactions in a public database of human proteins indicated that HaloMS provided a high-quality protein–protein interaction dataset.

Despite the statistically reliable overlap of HaloMS with the *Arabidopsis* dataset, the HaloMS overlap was moderate, indicating that the HaloMS dataset reflects the findings of the Y2H assay, which is the main data source in the Arabidopsis dataset [18, 105]; this may suggest that fewer data may be obtained by APMS interaction assays in plants. Further, the moderate finding of reproducibility using AI approaches suggests that the interactions detected by HaloMS are novel and scarce in the AI training dataset. The recently released AlphaFold 3 has been suggested to have better performance than AlphaFold 2 Multimer, but the code is not publicly available and its verification will take sufficient time [106]. We are planning to improve these prediction and evaluation steps using the newest deep learning models, AlphaFold 3 future.

Protein arrays, which use the same protein-capturing method as HaloMS, are based on *in vitro* synthesis of target proteins and do not consider their expression levels in the cell, thereby resulting in a constant copy number of target proteins. This may explain why target proteins involved in "transcription," such as TFs with a low expression level in the cell, are more likely to interact with TFs as query proteins. However, HaloMS uses cellular protein extracts as the source of targets, which may better reflect the dynamic range of proteins expressed in the cell.

HaloMS offers several advantages over conventional protein interaction assays, such as Y2H, protein arrays, and one-by-one APMS. The high-throughput synthesis and capture of bait proteins *in situ* eliminates the need for individual in-cell protein expression and purification. Additionally, because native samples are used as targets, rather than proteins expressed *in vitro* or in yeast, complexes composed of cofactors other than proteins, such as TF–RNA-binding protein–lnRNA complexes, can be detected. In addition to cultured cells, HaloMS can readily be applied to other samples, particularly human organs or tissues for which query transformations are difficult, and to diseased specimens for which intact conditions are examined. We envision that proteome-scale HaloMS could be adapted for other sample targets, such as brown-, white-, and beige-adipocyte, and non-transformable diabetic specimens. However, the HaloMS technique still requires a control experimental group to validate the reliability and stability of the results in advance when it is used for the assay for a variety of biological samples such as organs, tissues, and synthesized proteins even if the control experiments using Arabidopsis eliminate some issues which are necessary to consider in the reliability and stability of the HaloMS methods. Biologically, the TF-HaloMS dataset confirms known interactions, provides biochemical support for genetic and molecular biological observations, and illuminates previously unknown mutual regulation of the adipocyte signaling pathways.

HaloMS has a few potential limitations. Notably, we identified a large number of cell-type-specific interaction partners for the 17 query proteins. An important reason for the large number of interactions could be the selection of bait proteins that are critical components of some of the most vital adipocyte regulation processes. By screening TFs of adipocyte central signaling

pathways against different cell types, it is possible that we inadvertently identified cell-type-specific connected TF-related proteins and family members, such as hnRNPs and DDXs. These differentiated adipocyte-specific TFs and TF-related proteins have intrinsically disordered regions, and their network members may aggregate to transcriptionally regulate adipocyte differentiation [107–111]. HaloMS functions by immobilizing query proteins to a solid support, which may sterically limit the accessibility of the bait proteins to target proteins from the cell lysate. Thus, once interactions are identified despite the anchoring proteins, these interactions can be validated by the second assay, including AI technology. The new technology described in this study has the potential to yield significant results in mapping transcriptional protein complexes in patients; however, its detection is currently limited by the minimal amount of protein input necessary to detect protein–protein interactions. This study involves a sensitivity limitation with respect to the examination of these transcriptional protein complexes in rare/limited cell populations that may not be abundant in healthy donors or patients. In such cases, ultra-sensitive mass spectrometry methods have been developed that can identify trace samples and should be used for relevant analyses [112].

Acknowledgements

We would like to thank Editage (www.editage.jp) for English language editing.

Author contributions

Junshi Yazaki (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Funding acquisition [equal], Investigation [equal], Methodology [equal], Project administration [equal], Resources [equal], Software [equal], Supervision [equal], Validation [equal], Visualization [equal], Writing—original draft [equal], and Writing—review & editing [equal]), Takashi Yamanashi (Data curation [equal], Formal analysis [equal], Investigation [equal], Methodology [equal], Resources [equal], Software [equal], Visualization [equal], and Writing—original draft [equal]), Shino Nemoto (Investigation [equal], Resources [equal], Supervision [equal], and Writing—original draft [equal]), Atsuo Kobayashi (Investigation [equal], Methodology [equal], Resources [equal], and Validation [equal]), Yongwoon Han (Data curation [equal] and Investigation [equal]), Tomoko Hasegawa (Investigation [equal] and Resources [equal]), Akira Iwase (Investigation [equal] and Resources [equal]), Masaki Ishikawa (Investigation [equal]), Ryo Konno (Investigation [equal]), Koshi Imami (Supervision [equal]), Yusuke Kawashima (Conceptualization [equal], Data curation [equal], Formal analysis [equal], Funding acquisition [equal], Investigation [equal], Methodology [equal], Resources [equal], and Supervision [equal]), and Jun Seita (Conceptualization [equal], Funding acquisition [equal], Methodology [equal], Project administration [equal], Resources [equal], and Supervision [equal]).

Supplementary data

Supplementary data are available at *Biology Methods and Protocols* online.

Funding

This study was supported by The Ministry of Education, Science, Sports, and Culture, Grant-in-Aid for Scientific Research (C), 2019–2021 [19K0574 to J.Y.]. Takeda Science Foundation, Visionary Research Grant to J.Y. Japan Diabetes Foundation, Novo Nordisk Pharma Ltd. research grant 2022 to J.Y.

Data availability

Plasmids generated in this study are available upon request. These plasmid vectors below are available from the Arabidopsis Biological Resource Center (ABRC; https://abrc.osu.edu/). pIX-Halo: ccdb destination vector: https://www.arabidopsis.org/servlets/TairObject?type=vector&id=1001200298. pIX-3×HA: ccdB destination vector: https://www.arabidopsis.org/servlets/TairObject?type=vector&id=1001200524. MS data generated in this study were deposited in JPOST under the accession number PXD041085 (https://repository.jpostdb.org/). The code used in the current study has been deposited in GitHub under the link; https://github.com/seitalab/HaloMS and is publicly available. Additional information needed to reanalyze the data reported in this work is available from the corresponding author upon request.

References

1. Wild S, Roglic G, Green A, Sicree R, King H. Global burden of diabetes, 1995–2025: prevalence, numerical estimates, and projections. *Diabetes care* 2004; **27**: 1047–1053.
2. Inagaki T, Sakai J, Kajimura S. Transcriptional and epigenetic control of brown and beige adipose cell fate and function. *Nat Rev Mol Cell Biol* 2016;**17**:480–95.
3. Ishigaki K, Akiyama M, Kanai M *et al.* Large-scale genome-wide association study in a Japanese population identifies novel susceptibility loci across different diseases. *Nat Genet* 2020; **52**:669–79.
4. Kokaji T, Hatano A, Ito Y *et al.* Transomics analysis reveals allosteric and gene regulation axes for altered hepatic glucose-responsive metabolism in obesity. *Sci Signal* 2020;**13**:eaaz1236.
5. Siersbæk R, Nielsen R, Mandrup S. Transcriptional networks and chromatin remodeling controlling adipogenesis. *Trends Endocrinol Metab* 2012;**23**:56–64.
6. Siersbæk R, Rabiee A, Nielsen R *et al.* Transcription factor cooperativity in early adipogenic hotspots and super-enhancers. *Cell Rep* 2014;**7**:1443–55.
7. Rolland T, Taşan M, Charloteaux B *et al.* A proteome-scale map of the human interactome network. *Cell* 2014;**159**:1212–26.
8. Jumper J, Evans R, Pritzel A *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* 2021;**596**:583–9.
9. Tunyasuvunakool K, Adler J, Wu Z *et al.* Highly accurate protein structure prediction for the human proteome. *Nature* 2021;**596**:590–6.
10. Rohl CA, Strauss CE, Misura KM, Baker D. Protein structure prediction using Rosetta. *Methods Enzymol* 2004;**383**:66–93.
11. Mirdita M, Schütze K, Moriwaki Y *et al.* ColabFold: making protein folding accessible to all. *Nat Methods* 2022;**19**:679–82.
12. Pei J, Zhang J, Cong Q. Human mitochondrial protein complexes revealed by large-scale coevolution analysis and deep learning-based structure modeling. *Bioinformatics* 2022; **38**:4301–11.
13. Akdel M, Pires DEV, Pardo EP *et al.* A structural biology community assessment of AlphaFold2 applications. *Nat Struct Mol Biol* 2022;**29**:1056–67.
14. Evans R, O'Neill M, Pritzel A *et al.* Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021. https://doi.org/10.1101/2021.10.04.463034.

15. Harvey I, Boudreau A, Stephens JM. Adipose tissue in health and disease. *Open Biol* 2020;**10**:200291.
16. Yazaki J, Galli M, Kim AY *et al.* Mapping transcription factor interactome networks using HaloTag protein arrays. *Proc Natl Acad Sci USA* 2016;**113**:E4238–47.
17. Yazaki J, Galli M, Kim AY, Ecker JR. Profiling interactome networks with the HaloTag-NAPPA in situ protein array. *Curr Protoc Plant Biol* 2018;**3**:e20071.
18. Yazaki J, Kawashima Y, Ogawa T *et al.* HaloTag-based conjugation of proteins to barcoding-oligonucleotides. *Nucleic Acids Res* 2020;**48**:e8.
19. Yazaki J. Novel protein-oligonucleotide conjugation method involving a high-affinity capture HaloTag. *Bio Protoc* 2020; **10**:e3759.
20. Kadota Y, Macho AP, Zipfel C. Immunoprecipitation of plasma membrane receptor-like kinases for identification of phosphorylation sites and associated proteins. *Methods Mol Biol* 2016;**1363**:133–44.
21. Amodei D, Egertson J, MacLean BX *et al.* Improving precursor selectivity in data-independent acquisition using overlapping windows. *J Am Soc Mass Spectrom* 2019;**30**:669–84.
22. Kawashima Y, Watanabe E, Umeyama T *et al.* Optimization of data-independent acquisition mass spectrometry for deep and highly sensitive proteomic analysis. *Int J Mol Sci* 2019;**20**:5932.
23. Kawashima Y, Nagai H, Konno R *et al.* Single-shot 10K proteome approach: over 10,000 protein identifications by data-independent acquisition-based single-shot proteomics with ion mobility spectrometry. *J Proteome Res* 2022;**21**:1418–27.
24. Searle BC, Swearingen KE, Barnes CA *et al.* Generating high quality libraries for DIA MS with empirically corrected peptide predictions. *Nat Commun* 2020;**11**:1548.
25. Gessulat S, Schmidt T, Zolg DP *et al.* Prosit: proteome-wide prediction of peptide tandem mass spectra by deep learning. *Nat Methods* 2019;**16**:509–18.
26. Searle BC, Pino LK, Egertson JD *et al.* Chromatogram libraries improve peptide detection and quantification by data independent acquisition mass spectrometry. *Nat Commun* 2018;**9**:5128.
27. UniProt Consortium. UniProt: the universal protein KnowledgeBase in 2021. *Nucleic Acids Res* 2021;**49**:D480–89.
28. Lin Z, Akin H, Rao R *et al.* Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science* 2023;**379**:1123–30.
29. McInnes L, Healy J, Saul N *et al.* UMAP: uniform manifold approximation and projection for dimension reduction. *J Open Source Softw* 2018;**3**:861.
30. Li X, Wang W, Wang J *et al.* Proteomic analyses reveal distinct chromatin-associated and soluble transcription factor complexes. *Mol Syst Biol* 2015;**11**:775.
31. Wolf FA, Angerer P, Theis FJ. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* 2018;**19**:15.
32. Virtanen P, Gommers R, Oliphant TE *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods* 2020;**17**:352–272.
33. Yu H, Braun P, Yildirim MA *et al.* High-quality binary protein interaction map of the yeast interactome network. *Science* 2008;**322**:104–10.
34. Arabidopsis Interactome Mapping Consortium. Evidence for network evolution in an Arabidopsis interactome map. *Science* 2011;**333**:601–7.
35. Braun P, Tasan M, Dreze M *et al.* An experimentally derived confidence score for binary protein–protein interactions. *Nat Methods* 2009;**6**:91–7.
36. Venkatesan K, Rual J-F, Vazquez A *et al.* An empirical framework for binary interactome mapping. *Nat Methods* 2009; **6**:83–90.
37. Braun P. Interactome mapping for analysis of complex phenotypes: insights from benchmarking binary interaction assays. *Proteomics* 2012;**12**:1499–518.
38. Cusick ME, Yu H, Smolyar A *et al.* Literature-curated protein interaction datasets. *Nat Methods* 2009;**6**:39–46.
39. Ding H, Zheng S, Garcia-Ruiz D *et al.* Fasting induces a subcutaneous-to-visceral fat switch mediated by microRNA-149-3p and suppression of PRDM16. *Nat Commun* 2016;**7**:11533.
40. Blüher M. Adipose tissue dysfunction in obesity. *Exp Clin Endocrinol Diabetes* 2009;**117**:241–50.
41. Ye RZ, Richard G, Gévry N *et al.* Fat cell size: measurement methods, pathophysiological origins, and relationships with metabolic dysregulations. *Endocr Rev* 2022;**43**:35–60.
42. Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature* 2006; **444**:840–6.
43. Nelson LR, Bulun SE. Estrogen production and action. *J Am Acad Dermatol* 2001;**45**:S116–24.
44. Klok MD, Jakobsdottir S, Drent ML. The role of leptin and ghrelin in the regulation of food intake and body weight in humans: a review. *Obes Rev* 2007;**8**:21–34.
45. Brivanlou AH, Darnell JE Jr. Signal transduction and the control of gene expression. *Science* 2002;**295**:813–8.
46. Fontaine F, Overman J, François M. Pharmacological manipulation of transcription factor protein–protein interactions: opportunities and obstacles. *Cell Regen* 2015;**4**:2.
47. Rivera-Reyes R, Kleppa MJ, Kispert A. Proteomic analysis identifies transcriptional cofactors and homeobox transcription factors as TBX18 binding proteins. *PLoS One* 2018;**13**:e0200964.
48. Fortunato S. Community detection in graphs. *Phys Rep* 2010; **486**:75–174.
49. de A Boleti AP, de O Cardoso PH, Frihling BEF, Silva *et al.* Adipose tissue, systematic inflammation, and neurodegenerative diseases. *Neural Regen Res* 2023;**18**:38–46.
50. Ahmadian M, Suh JM, Hah N *et al.* PPARγ signaling and metabolism: the good, the bad and the future. *Nat Med* 2013;**19**:557–66.
51. Choi J, Lee SY, Yoo YM, Kim CH. Maturation of adipocytes is suppressed by fluid shear stress. *Cell Biochem Biophys* 2017; **75**:87–94.
52. Fantuzzi G, Mazzone T. Adipose tissue and atherosclerosis: exploring the connection. *Arterioscler Thromb Vasc Biol* 2007; **27**:996–1003.
53. Collins S. β-adrenoceptor signaling networks in adipocytes for recruiting stored fat and energy expenditure. *Front Endocrin* 2011;**2**:102.
54. Volke L, Krause K. Effect of thyroid hormones on adipose tissue flexibility. *Eur Thyroid J* 2021;**10**:1–9.
55. Obregon MJ. Adipose tissues and thyroid hormones. *Front Physiol* 2014;**5**:479.
56. Vauclard A, Bellio M, Valet C *et al.* Obesity: effects on bone marrow homeostasis and platelet activation. *Thromb Res* 2023; **231**:195–205.
57. Park H, He A, Tan M *et al.* Peroxisome-derived lipids regulate adipose thermogenesis by mediating cold-induced mitochondrial fission. *J Clin Invest* 2019;**129**:694–711.
58. Liu J, Lu W, Shi B *et al.* Peroxisomal regulation of redox homeostasis and adipocyte metabolism. *Redox Biol* 2019;**24**:101167.
59. Wongtrakul-Kish K, Herbert BR, Packer NH. Bisecting GlcNAc protein N-glycosylation is characteristic of human adipogenesis. *J Proteome Res* 2021;**20**:1313–27.

60. Kouranti I, Abdel Khalek W, Mazurkiewicz S *et al.* Cullin 3 Exon 9 deletion in familial hyperkalemic hypertension impairs Cullin3-ring-E3 ligase (CRL3) dynamic regulation and cycling. *Int J Mol Sci* 2022;**23**:5151.
61. Chatrathi HE, Collins JC, Wolfe LA *et al.* Novel CUL3 variant causing familial hyperkalemic hypertension impairs regulation and function of ubiquitin ligase activity. *Hypertension* 2022; **79**:60–75.
62. Wan C, Borgeson B, Phanse S *et al.* Panorama of ancient metazoan macromolecular complexes. *Nature* 2015;**525**:339–44.
63. Bennett EJ, Rush J, Gygi SP, Harper JW. Dynamics of cullin-RING ubiquitin ligase network revealed by systematic quantitative proteomics. *Cell* 2010;**143**:951–65.
64. Sun L, Lin JD. Function and mechanism of long noncoding RNAs in adipocyte biology. *Diabetes* 2019;**68**:887–96.
65. Alvarez-Dominguez JR, Bai Z, Xu D *et al.* De novo reconstruction of adipose tissue transcriptomes reveals long non-coding RNA regulators of brown adipocyte development. *Cell Metab* 2015;**21**:764–76.
66. Mi L, Zhao XY, Li S *et al.* Conserved function of the long noncoding RNA Blnc1 in brown adipocyte differentiation. *Mol Metab* 2017;**6**:101–10.
67. Fei T, Chen Y, Xiao T *et al.* Genome-wide CRISPR screen identifies HNRNPL as a prostate cancer dependency regulating RNA splicing. *Proc Natl Acad Sci USA* 2017; **114**:E5207–15.
68. Zhao XY, Li S, Wang GX *et al.* A long noncoding RNA transcriptional regulatory circuit drives thermogenic adipocyte differentiation. *Mol Cell* 2014;**55**:372–82.
69. Li S, Mi L, Yu L *et al.* Zbtb7b engages the long noncoding RNA Blnc1 to drive brown and beige fat development and thermogenesis. *Proc Natl Acad Sci USA* 2017;**114**:E7111–20.
70. Hasegawa Y, Brockdorff N, Kawano S *et al.* The matrix protein hnRNP U is required for chromosomal localization of Xist RNA. *Dev Cell* 2010;**19**:469–76.
71. Linder P, Jankowsky E. From unwinding to clamping—the DEAD box RNA helicase family. *Nat Rev Mol Cell Biol* 2011; **12**:505–16.
72. Heidelberger JB, Voigt A, Borisova ME *et al.* Proteomic profiling of VCP substrates links VCP to K6-linked ubiquitylation and c-Myc function. *EMBO Rep* 2018;**19**:e44754.
73. Wang L, Chen C, Song Z *et al.* EZH2 depletion potentiates MYC degradation inhibiting neuroblastoma and small cell carcinoma tumor formation. *Nat Commun* 2022;**13**:12.
74. Kalkat M, Resetca D, Lourenco C *et al.* MYC protein interactome profiling reveals functionally distinct regions that cooperate to drive tumorigenesis. *Mol Cell* 2018;**72**:836–48.e7.
75. Koch HB, Zhang R, Verdoodt B, 3rd *et al.* Large-scale identification of c-MYC-associated proteins using a combined TAP/MudPIT approach. *Cell Cycle* 2007;**6**:205–17.
76. Shin S, Janknecht R. Concerted activation of the Mdm2 promoter by p72 RNA helicase and the coactivators p300 and P/CAF. *J Cell Biochem* 2007;**101**:1252–65.
77. Mahlokozera T, Patel B, Chen H *et al.* Competitive binding of E3 ligases TRIM26 and WWP2 controls SOX2 in glioblastoma. *Nat Commun* 2021;**12**:6321.
78. Fang X, Yoon JG, Li L *et al.* Landscape of the SOX2 protein–protein interactome. *Proteomics* 2011;**11**:921–34.
79. Dardenne E, Polay Espinoza M, Fattet L *et al.* RNA helicases DDX5 and DDX17 dynamically orchestrate transcription, miRNA, and splicing programs in cell differentiation. *Cell Rep* 2014;**7**:1900–13.
80. Xu K, Sun S, Yan M *et al.* DDX5 and DDX17-multifaceted proteins in the regulation of tumorigenesis and tumor progression. *Front Oncol* 2022;**12**:943032.
81. Peng HY, Liang YC, Tan TH *et al.* RBM4a-SRSF3-MAP4K4 splicing cascade constitutes a molecular mechanism for regulating brown adipogenesis. *Int J Mol Sci* 2018;**19**:2646.
82. Sen S, Langiewicz M, Jumaa H, Webster NJ. Deletion of serine/arginine-rich splicing factor 3 in hepatocytes predisposes to hepatocellular carcinoma in mice. *Hepatology* 2015;**61**:171–83.
83. Aprile M, Cataldi S, Ambrosio MR *et al.* PPARγΔ5, a naturally occurring dominant-negative splice isoform, impairs PPARgamma function and adipocyte differentiation. *Cell Rep* 2018;**25**: 1577–92.e6.
84. Hartmuth K, Urlaub H, Vornlocher H-P *et al.* Protein composition of human prespliceosomes isolated by a tobramycin affinity-selection method. *Proc Natl Acad Sci USA* 2002; **99**:16719–24.
85. Mugabo Y, Sadeghi M, Fang NN *et al.* Elucidation of the 14–3-3ζ interactome reveals critical roles of RNA-splicing factors during adipogenesis. *J Biol Chem* 2018;**293**:6736–50.
86. Berezin AE, Berezin AA, Lichtenauer M. Emerging role of adipocyte dysfunction in inducing heart failure among obese patients with prediabetes and known diabetes mellitus. *Front Cardiovasc Med* 2020;**7**:583175.
87. Noyes AM, Dua K, Devadoss R *et al.* Cardiac adipose tissue and its relationship to diabetes mellitus and cardiovascular disease. *World J Diabetes* 2014;**5**:868–76.
88. Longo M, Zatterale F, Naderi J *et al.* Adipose tissue dysfunction as determinant of obesity-associated metabolic complications. *Int J Mol Sci* 2019;**20**:2358.
89. Lau E, Kluger H, Varsano T *et al.* PKCepsilon promotes oncogenic functions of ATF2 in the nucleus while blocking its apoptotic function at mitochondria. *Cell* 2012;**148**:543–55.
90. Al Hasan M, Martin PE, Shu X *et al.* Type III collagen is required for adipogenesis and actin stress fibre formation in 3T3-L1 preadipocytes. *Biomolecules* 2021;**11**:156.
91. Chen X, Zhang J, Li T *et al.* COL3A1 induces ischemic heart failure by activating AGE/RAGE pathway. *Signa Vitae* 2022;**18**:45–52.
92. Patti ME, Corvera S. The role of mitochondria in the pathogenesis of type 2 diabetes. *Endocr Rev* 2010;**31**:364–95.
93. Hill MA, Sykes AM, Mellick GD. ER-phagy in neurodegeneration. *J Neurosci Res* 2023;**101**:1611–23.
94. Demirtas L, Guclu A, Erdur FM *et al.* Apoptosis, autophagy & endoplasmic reticulum stress in diabetes mellitus. *Indian J Med Res* 2016;**144**:515–24.
95. Quan W, Lim YM, Lee MS. Role of autophagy in diabetes and endoplasmic reticulum stress of pancreatic β-cells. *Exp Mol Med* 2012;**44**:81–8.
96. Jiang Q, Wang Y, Xiang M *et al.* UFL1, a UFMylation E3 ligase, plays a crucial role in multiple cellular stress responses. *Front Endocrinol* 2023;**14**:1123124.
97. Mochida K, Nakatogawa H. ER-phagy: selective autophagy of the endoplasmic reticulum. *EMBO Rep* 2022;**23**:e55192.
98. Yoshida S, Hasegawa T, Suzuki M *et al.* Parkinson's disease-linked DNAJC13 mutation aggravates alpha-synuclein-induced neurotoxicity through perturbation of endosomal trafficking. *Hum Mol Genet* 2018;**27**:823–36.
99. Koizumi H, Fujioka H, Togashi K *et al.* DCLK1 phosphorylates the microtubule-associated protein MAP7D1 to promote axon elongation in cortical neurons. *Dev Neurobiol* 2017;**77**:493–510.
100. Leegwater PA, Vermeulen G, Könst AA *et al.* Subunits of the translation initiation factor eIF2B are mutant in leukoencephalopathy with vanishing white matter. *Nat Genet* 2001;**29**:383–8.

101. Hanson FM, Hodgson RE, de Oliveira MIR *et al.* Regulation and function of eIF2B in neurological and metabolic disorders. *Biosci Rep* 2022;**42**:BSR20211699.
102. Hosseini SA, Labilloy A. *Genetics, TREX1 Mutations*. Treasure Island, FL: StatPearls Publishing, 2023.
103. Singh H, Lane AA, Correll M *et al.* Putative RNA-splicing gene LUC7L2 on 7q34 represents a candidate gene in pathogenesis of myeloid malignancies. *Blood Cancer J* 2013; **3**:e117.
104. Visconte V, Nakashima MO, Rogers HJ. Mutations in splicing factor genes in myeloid malignancies: significance and impact on clinical features. *Cancers* 2019;**11**:1844.
105. Vidal M, Cusick ME, Barabási AL. Interactome networks and human disease. *Cell* 2011;**144**:986–98.
106. Abramson J, Adler J, Dunger J *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024.
107. Sabari BR, Dall'Agnese A, Boija A *et al.* Coactivator condensation at super-enhancers links phase separation and gene control. *Science* 2018;**361**:eaar3958.
108. Alberti S, Hyman AA. Biomolecular condensates at the nexus of cellular stress, protein aggregation disease and ageing. *Nat Rev Mol Cell Biol* 2021;**22**:196–213.
109. Yamazaki T, Nakagawa S, Hirose T. Architectural RNAs for membraneless nuclear body formation. *Cold Spring Harb Symp Quant Biol* 2019;**84**:227–37.
110. Shin Y, Brangwynne CP. Liquid phase condensation in cell physiology and disease. *Science* 2017;**357**:eaaf4382.
111. Perdikari TM, Murthy AC, Ryan VH *et al.* SARS-CoV-2 nucleocapsid protein phase-separates with RNA and with human hnRNPs. *EMBO J* 2020;**39**:e106478.
112. Konno R, Ishikawa M, Nakajima D *et al.* Universal pretreatment development for low-input proteomics using lauryl maltose neopentyl glycol. *Mol Cell Proteomics* 2024;**23**(4):100745.

Biology Methods and Protocols, 2024, 00, 1–14
https://doi.org/10.1093/biomethods/bpae039
Methods Article