



Data Article

Dataset on the mass spectrometry-based proteomic profiling of mouse embryonic fibroblasts from a wild type and DYT-TOR1A mouse model of dystonia, basally and during stress



Kunal Shroff^a, Zachary F. Caffall^a, Erik J. Soderblom^{c,d}, Greg Waitt^d,
Tricia Ho^d, Nicole Calakos^{a,b,c,e,*}

^a Department of Neurology, Duke University Medical Center, Durham, NC, United States

^b Department of Neurobiology, Duke University Medical Center, Durham, NC, United States

^c Department of Cell Biology, Duke University Medical Center, Durham, NC, United States

^d Proteomics and Metabolomics Core Facility, Duke Center for Genomic and Computational Biology, Durham, NC, United States

^e Duke Institute for Brain Sciences, Duke University, Durham, NC, United States

ARTICLE INFO

Article history:

Received 22 September 2021

Revised 1 November 2021

Accepted 16 November 2021

Available online 20 November 2021

Keywords:

Dystonia

TorsinA

Proteomics

Subcellular fractionation

Thapsigargin

Stress response

Movement disorder

ABSTRACT

Here, we present quantitative subcellular compartment-specific proteomic data from wildtype and DYT-TOR1A heterozygous mouse embryonic fibroblasts (MEFs) basally and following thapsigargin (Tg) treatment [1]. In this experiment, we generated MEFs from wild type (WT) and a heterozygous DYT-TOR1A mouse model of dystonia. Subsequently, these MEF cultures were treated with either 1 μ M Tg or dimethylsulfoxide vehicle (Veh) for six hours. Following treatment, the cells were fractionated into nuclear and cytosolic fractions. Liquid chromatography, tandem mass spectrometry (LC/MS/MS)-based proteomic profiling identified 65,056 unique peptides and 4801 unique proteins across all samples. The data presented here provide subcellular compartment-specific proteomic information within a dystonia model system both basally and under cellular stress.

DOI of original article: [10.1016/j.nbd.2021.105464](https://doi.org/10.1016/j.nbd.2021.105464)

* Corresponding author at: Department of Neurology, Duke University Medical Center, Durham, NC, United States

E-mail address: nicole.calakos@duke.edu (N. Calakos).

<https://doi.org/10.1016/j.dib.2021.107609>

2352-3409/© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>)

These data can inform future experiments focused on studying the function of TorsinA, the protein encoded by TOR1A, and its potential role in nucleocytoplasmic transport and proteostasis. In addition, the information in this article can also inform future mechanistic studies investigating the relationship between DYT-TOR1A dystonia and the cellular stress response to advance understanding of the pathogenesis of dystonia.

© 2021 The Author(s). Published by Elsevier Inc.
This is an open access article under the CC BY license
(<http://creativecommons.org/licenses/by/4.0/>)

Specifications Table

Subject	Neuroscience: Cellular and Molecular
Specific subject area	Molecular pathogenesis of DYT-TOR1A dystonia
Type of data	Mass spectrometry data Proteomic data Tables
How data were acquired	Quantitative peptide data were acquired via Quantitative LC/MS/MS using a nanoAcquity UPLC system coupled to a Thermo Orbitrap Fusion Lumos high resolution accurate mass tandem mass spectrometer via a nanoelectrospray ionization source. Peptide data was imported and accurate mass/retention time aligned via Proteome Discoverer 2.2 (Thermo Scientific Inc.) using the Minora Feature Detector algorithm to determine area-under-the-curve label-free quantitative proteomic data.
Data format	Raw mass spectrometry data is presented in the MassIVE data repository. A conversion key between the machine ID used to label the raw data and the sample information used to label the data in further analyses is presented in Supplementary Table 1. Raw data on peptide and protein quantification under the various experimental conditions is presented in Supplementary Tables 2 and 3.
Parameters for data collection	Unbiased label-free quantitative proteomic data was collected from three wildtype and three DYT-TOR1A heterozygous primary mouse embryonic fibroblast (MEF) cell lines. The cells were then either stressed with 1 μ M thapsigargin (Tg) for 6 hours or a volumetric equivalent of DMSO vehicle (Veh) before being lysed and fractionated into a nuclear and cytosolic fraction.
Description of data collection	DYT-TOR1A mice were bred with C57BL/6 mice to produce three wildtype and three heterozygous DYT-TOR1A mouse embryos. Primary MEF cultures were isolated from these embryos, treated with either Tg or Veh, and fractionated into nuclear and cytosolic fractions. Every fraction was then trypsinized to enable unique peptide identification via LC/MS/MS. Peptide spectrum matches were determined using the Mascot algorithm against a <i>Mus musculus</i> database and AUCs of each peptide precursor were measured using the Minora algorithm within Protein Discoverer 2.2.
Data source location	Duke University Durham, NC United States
Data accessibility	MassIVE Repository Dataset Identifier: MSV000087774 Direct FTP link to data: ftp://massive.ucsd.edu/MSV000087774/
Related research article	K. Shroff, Z.F. Caffall, N. Calakos, (2021). DYT-TOR1A Subcellular Proteomics Reveals Selective Vulnerability of the Nuclear Proteome to Cell Stress, Neurobiol. Dis. https://doi.org/10.1016/j.NBD.2021.105464

Value of the Data

- Our data provide information about subcellular compartment-specific proteome changes in DYT-TOR1A dystonia cell models. Because the TorsinA protein traffics between two subcellular compartments, an analysis separating the cytosolic and nuclear compartments is useful to identify disruptions that are specific to a compartment and could be missed by evaluating unfractionated cells or tissue.
- These data can be beneficial to researchers interested in: cell biology of the AAA⁺-ATPase TorsinA, biology of cell stress responses, and understanding how the DYT-TOR1A mutation alters the cellular proteome to gain insights to dystonia pathogenesis.
- These data can be further used to develop experiments aimed at understanding the function of TorsinA and investigating its potential role in regulating nucleocytoplasmic transport and/or proteostasis.
- These data can be used to identify novel proteins that may be mediators of TOR1A's pathogenic effects particularly when expressed in combination with exogenous cellular stress.

1. Data Description

We generated quantitative proteomic profiles of the nuclear and cytosolic subcellular fractions of wildtype (WT) and DYT-TOR1A heterozygous mouse embryonic fibroblasts (MEFs) treated with either DMSO vehicle (Veh) or 1 μ M thapsigargin (Tg) for six hours. Fig. 1 shows a schematic of the workflow used to prepare the experimental samples for proteomic analysis.

The proteomic profiles for each sample are presented within the supplementary data files. Through the MassIVE repository link, we provide the unprocessed mass spectrometry data. The data is organized into separate .RAW files where each file represents data from a unique sample. Supplementary Table 1 provides information about the corresponding sample associated with each machine ID key which was used to label the .RAW files found in the MassIVE repository.

In Supplementary Table 2, we provide raw quantitative abundances for all 65,056 unique peptides identified by our proteomics methodology. These abundances are presented for each sample within the experiment. The peptides are grouped within the table based on their alignment with proteins found within the Swissprot *Mus musculus* protein sequence database. In Supplementary Table 3, we present the raw and normalized quantitative abundances for all 4801 unique proteins which had alignments with at least one unique peptide within our peptide dataset. Quantitative abundances for all these proteins are presented for each sample within the experiment, as well as quality control samples. For each protein, separate coefficients of variance are calculated for every experimental condition (i.e. every combination of genotype, treatment, and subcellular fraction). These calculations are based on the normalized quantitative protein abundances from each biological replicate.

2. Experimental Design, Materials and Methods

2.1. Animals

DYT-TOR1A knock-in mice (*Tor1a* ^{Δ GAG/+}) (courtesy of Dr. W. Dauer, University of Texas-Southwestern) [2] on C57BL/6 background were bred in standard housing conditions with food and water provided ad libitum. All procedures were approved by the Duke University Institutional Animal Care and Use Committee (IACUC).

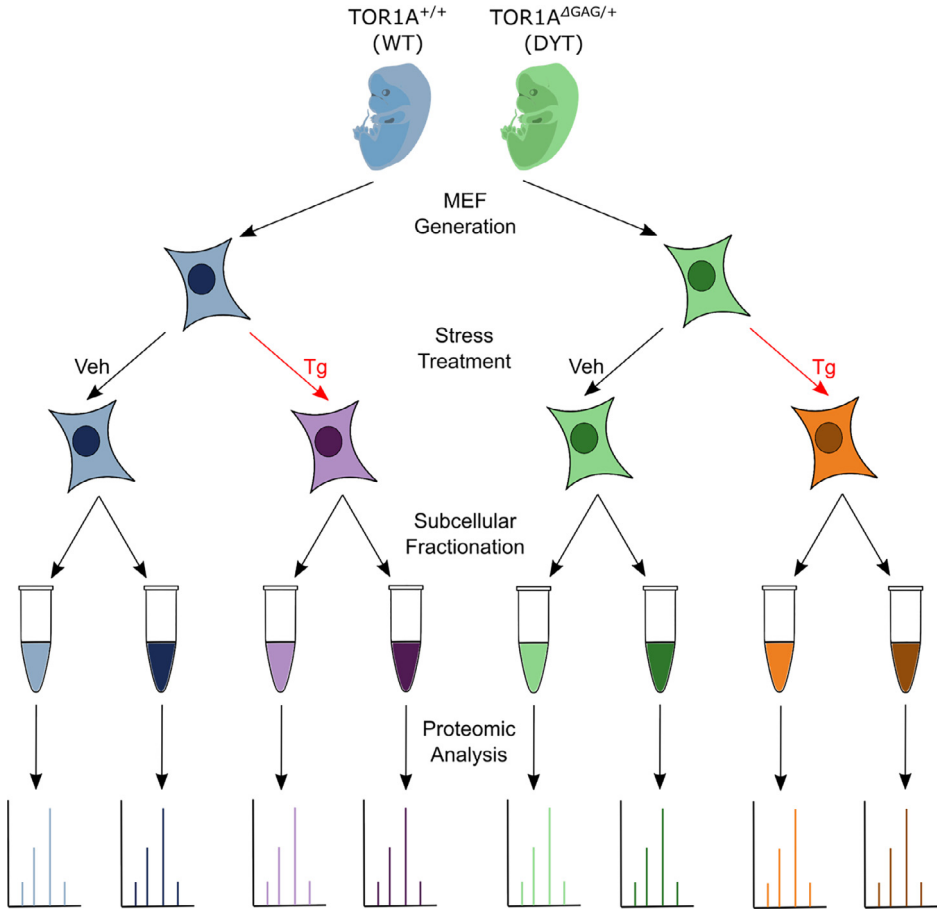


Fig. 1. Experimental workflow for proteomic samples. MEFs were generated from heterozygous DYT-TOR1A (*TOR1A^{ΔGAG/+}*) mouse embryos and WT littermates which were subsequently treated with either 1μM Tg for 6 hours or Veh. Treated MEFs were then fractionated into nuclear and cytosolic fractions and each fraction was subjected to quantitative differential proteomic analysis.

2.2. Mouse embryonic fibroblast (MEF) extraction, isolation, and immortalization

Female wildtype C57BL/6 mice were crossed with male heterozygous DYT-TOR1A knock-in mice (*Tor1a^{ΔGAG/+}*) on a C57BL/6 background to produce DYT-TOR1A model MEF and WT control MEF cell lines. MEF extraction occurred with minor modifications from the protocol as described in [3]. Three DYT-TOR1A MEF lines and three WT MEF lines were produced from a single litter using a single animal for each of the different lines.

At approximately 14 days post-coitum, the pregnant dam was euthanized using isoflurane followed by decapitation. The uterine horns were dissected out and rinsed in 70% (v/v) ethanol followed by PBS (Gibco, Invitrogen) before being placed into a Petri dish containing PBS (Gibco, Invitrogen). Each individual embryonic sac was isolated from the uterine horns and placenta, and then placed into a separate Petri dish containing PBS. Within each separate dish, the embryo was dissected out of the embryonic sac.

Following dissection from the embryonic sac, the head and red organs of each embryo were surgically excised. The remaining embryonic tissue was placed into a clean Petri dish where it

was minced with a sterile razor blade for 1 min. 1 mL of 0.05% trypsin/EDTA (Gibco, Invitrogen) was added to each dish containing embryonic tissue. The mixture was transferred into a 15 mL Falcon tube and incubated at 37°C for 10 min. Following incubation, MEFs were dissociated via pipetting. The incubation and dissociation steps were repeated twice more to further dissociate the MEFs from the extracellular matrix of the embryonic tissue. Following the last dissociation step, 2 mL of fetal bovine serum-containing media ("MEF media" described in Cell Culture section below) was added to each tube to inactivate the trypsin. The tubes were subsequently centrifuged at 500 x g for 5 min. The supernatant was removed, and the cell pellet was resuspended in 10 mL of warm MEF media.

This solution was then plated on TC dishes coated in 1% Matrigel (Corning). The MEFs were genotyped after two passages and subsequently immortalized via transfection by a plasmid encoding the SV40 T antigen as described in [4]. Immortalized cell lines were used within 5 passages.

2.3. Genotyping

All genotyping was conducted as previously described in [2].

2.4. Cell culture

MEFs were grown in MEF media prepared with 500 mL of DMEM, high glucose, pyruvate (Thermo Fisher, #11995), 50 mL of Fetal Bovine Serum, 5 mL of Antibiotic-Antimycotic (Gibco, Invitrogen), 5 mL of 200 mM L-Glutamine (Gibco, Invitrogen), 5 mL of MEM Non-Essential Amino Acids Solution (Gibco, Invitrogen), and 500 µL of 2-Mercaptoethanol (Sigma-Aldrich). MEFs were grown in incubators at 37°C/5% CO₂.

2.5. Thapsigargin treatment and subcellular fractionation

Three separate experiments were performed exposing MEFs to the cell stressor Tg. During each experiment, a pair of MEF lines (1 WT and 1 DYT-TOR1A) were grown within 10 cm dishes to 90% confluence. Treatment consisted of exchanging the MEF media within a dish with MEF media treated with either 1 µM Tg dissolved in DMSO or an equivalent volume of DMSO (Veh). Following the treatment, the dishes were placed back in incubators at 37°C/5% CO₂ for six hours. After six hours of treatment, the MEFs were trypsinized to enable dissociation from the 10 cm treatment dishes. The dissociated MEFs were placed in 15 mL Falcon tubes prior to being centrifuged to produce a MEF pellet.

The MEFs were subsequently fractionated into nuclear and cytosolic fractions via a protocol carried out as described in [5]. Briefly, the MEF pellet was resuspended in a mild detergent solution consisting of 0.01% NP40 dissolved in PBS. This solution was incubated at room temperature for 3 min. Following incubation, the resuspended solution was centrifuged to collect the supernatant (cytosolic fraction). Further solubilization was done in the same buffer alongside sonication to penetrate the double bilayer membranous nuclear compartment. A second centrifugation was performed to collect the supernatant (nuclear fraction).

2.6. Quantitative LC/MS/MS and proteomic analysis

Twenty-four samples in total were submitted to the Duke Proteomics and Metabolomics Shared Resource (two subcellular fractions [Nuc, Cyto] from each of the six MEF lines treated

with either Tg or Veh). All of the samples were analyzed within a single set of liquid chromatography with tandem mass spectroscopy (LC/MS/MS) data acquisitions, irrespective of the treatments occurring as part of separate cell culture experiments. These acquisitions were conducted using a label-free quantification approach.

Fractions were first normalized to 10 μg based on BCA results. Fractions were then spiked with undigested casein at a total of 20, 30, or 40 fmol/ μg , reduced with 10 mM dithiothreitol for 30 min at 80°C, and alkylated with 20 mM iodoacetamide for 30 min at room temperature. Next, they were supplemented with a final concentration of 1.2% phosphoric acid and 741 μL of S-Trap (ProtiFi) binding buffer (90% MeOH/100mM TEAB). Proteins were trapped on the S-Trap, digested using 20 ng/ μL sequencing grade trypsin (Promega) for 1 hour at 47°C, and eluted using 50 mM TEAB, followed by 0.2% FA, and lastly using 50% ACN/0.2% FA. All fractions were then lyophilized to dryness and resuspended in 20 μL 1% TFA/2% acetonitrile containing 12.5 fmol/ μL yeast alcohol dehydrogenase (ADH_YEAST). Three QC Pools were created: 1) 3 μL from each of the nuclear fractions (NucPool), 2) 3 μL from each of the cytosolic fractions (CytoPool) 3) 3 μL from each of all of the fractions, both nuclear and cytosolic (TotalPool). All QC Pools were run periodically randomly interspersed throughout the test fractions.

Quantitative LC/MS/MS was performed on 2 μL of each fraction, using a nanoAcquity UPLC system (Waters Corp.) coupled to a Thermo Orbitrap Fusion Lumos high resolution accurate mass tandem mass spectrometer (Thermo) via a nanoelectrospray ionization source. Briefly, the fraction was first trapped on a Symmetry C18 20 mm \times 180 μm trapping column (5 $\mu\text{L}/\text{min}$ at 99.9/0.1 v/v water/acetonitrile). Analytical separation was then performed using a 1.8 μm Acquity HSS T3 C18 75 μm \times 250 mm column (Waters Corp.) with a 90-min linear gradient of 5 to 30% acetonitrile with 0.1% formic acid at a flow rate of 400 nl/min with a column temperature of 55°C. Data collection on the Fusion Lumos mass spectrometer was performed in a data-dependent acquisition (DDA) mode of acquisition with a $r = 120,000$ (@ m/z 200) full MS scan from m/z 375 – 1500 with a target AGC value of $2\text{e}5$ ions.

MS/MS scans were acquired at Rapid scan rate (Ion Trap) with an AGC target of $5\text{e}3$ ions and a max injection time of 100 ms. The total cycle time for MS and MS/MS scans was 2 s. A 20s dynamic exclusion was employed to increase depth of coverage. The total analysis cycle time for each fraction injection was approximately 2 hours.

Following 35 total UPLC-MS/MS analyses (excluding conditioning runs, but including 3 replicate QC Pool (TotalPool), 4 replicate nuclear (NucPool) and 4 replicate cytosolic Pool (CytoPool) injections), data was imported into Proteome Discoverer 2.2 (Thermo Scientific Inc.). Analyses were aligned based on the accurate mass and retention time of detected ions (“features”) using Minora Feature Detector algorithm in Proteome Discoverer. Protein levels are based on the relative peptide abundance measures which were calculated by area-under-the-curve (AUC) of the selected ion chromatograms of the aligned features across all runs. Levels are reported in arbitrary units (a.u.).

Mascot Distiller and Mascot Server (v 2.5, Matrix Sciences) were utilized to produce fragment ion spectra from the MS/MS data and to perform database searches. The MS/MS data was searched against the SwissProt *M. musculus* database (downloaded in Apr 2017) to identify relevant *M. musculus* proteins. The MS/MS data was also searched against an equal number of reversed-sequence “decoys” for false discovery rate determination. Database search parameters included fixed modification on Cys (carbamidomethyl), variable modifications on Meth (oxidation) and Asn and Gln (deamidation), full trypsin enzyme rules, and mass tolerances of 5ppm precursor ions and 0.8da product ions. Peptide Validator and Protein FDR Validator nodes in Proteome Discoverer were used to annotate the data at a maximum 1% protein false discovery rate. Sample loading data normalization was achieved by normalization of the raw total peptide intensity sum across all proteins from each sample.

Following sample loading normalization, imputation of missing values was performed in the following manner. If less than half of the values were missing across all samples, values were imputed with an intensity derived from a normal distribution defined by measured values within the same intensity range (20 bins). If greater than half values were missing for a peptide across all samples and a peptide intensity was $>5\text{e}6$, then it was concluded that peptide was

misaligned and its measured intensity was set to 0. All remaining missing values are imputed with the lowest 5% of all detected values.

Ethics Statement

All animal procedures were approved by the Duke University Institutional Animal Care and Use Committee (IACUC). All animal procedures performed comply with ARRIVE guidelines and were carried out in accordance with the National Institutes of Health guide for care and the use of Laboratory animals, as defined by NIH Publications No. 8023 (revised 1978).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships which have or could be perceived to have influenced the work reported in this article.

CRediT Author Statement

Kunal Shroff: Conceptualization, Investigation, Writing – original draft, Writing – review & editing; **Zachary F. Caffall:** Supervision, Writing – review & editing; **Erik J. Soderblom:** Data curation, Supervision, Writing – review & editing; **Greg Waitt:** Investigation, Data curation, Formal analysis; **Tricia Ho:** Investigation, Data curation; **Nicole Calakos:** Conceptualization, Funding acquisition, Supervision, Writing – review & editing.

Acknowledgments

The authors wish to acknowledge critical expertise and suggestions provided by Shataakshi Dube, William Dauer, Connor King, and Miranda Shipman. K.S. thanks the members of his undergraduate thesis committee, Ron Grunwald and Matthew Oliver, for continued support and guidance. We would also like to acknowledge the following funding sources for this project: Huang Undergraduate Summer Research Fellowship (K.S.), Duke Health Scholar Award (N.C.), Cure Dystonia Now (N.C.), Dystonia Medical Research Foundation (N.C.), and Tyler's Hope for a Dystonia Cure (N.C.).

Supplementary Materials

Supplementary material associated with this article can be found in the online version at doi:[10.1016/j.dib.2021.107609](https://doi.org/10.1016/j.dib.2021.107609).

References

- [1] K. Shroff, Z.F. Caffall, N. Calakos, DYT-TOR1A subcellular proteomics reveals selective vulnerability of the nuclear proteome to cell stress, *Neurobiol. Dis.* (2021), doi:[10.1016/j.nbd.2021.105464](https://doi.org/10.1016/j.nbd.2021.105464).
- [2] R.E. Goodchild, C.E. Kim, W.T. Dauer, Loss of the dystonia-associated protein torsinA selectively disrupts the neuronal nuclear envelope, *Neuron* 48 (2005) 923–932, doi:[10.1016/j.neuron.2005.11.010](https://doi.org/10.1016/j.neuron.2005.11.010).
- [3] J. Jozefczuk, K. Drews, J. Adjaye, Preparation of mouse embryonic fibroblast cells suitable for culturing human embryonic and induced pluripotent stem cells, *J. Vis. Exp.* (2012), doi:[10.3791/3854](https://doi.org/10.3791/3854).
- [4] Harding, H., 2003. Immortalization of MEF with SV40 T antigen, Internet.
- [5] K. Suzuki, P. Bose, R.Y. Leong-Quong, D.J. Fujita, K. Riabowol, REAP: A two minute cell fractionation method, *BMC Res. Notes.* (2010), doi:[10.1186/1756-0500-3-294](https://doi.org/10.1186/1756-0500-3-294).