DYT6 mutated THAP1 is a cell type dependent regulator of the SP1 family

Fubo Cheng,[1,2] Wenxu Zheng,[3] Peter Antony Barbuti,[4,5] Paola Bonsi,[6] Change Liu,[7] Nicolas Casadei,[1,8] Giulia Ponterio,[6] Maria Meringolo,[6] Jakob Admard,[1,8] Claire Marie Dording,[4,9] Libo Yu-Taeger,[1,10] Huu Phuc Nguyen,[10] Kathrin Grundmann-Hauser,[1] Thomas Ott,[1] Henry Houlden,[11] Antonio Pisani,[12,13] Rejko Krüger[4,14,15] and Olaf Riess[1,8]

1 Institute of Medical Genetics and Applied Genomics, University of Tuebingen, Tuebingen, Germany
2 Department of Neurology, the First Hospital of Jilin University, Changchun, China
3 Institute for Ophthalmic Research Centre for Ophthalmology, University of Tuebingen, Tuebingen, Germany
4 Transversal Translational Medicine, Luxembourg Institute of Health (LIH), Strassen, Luxembourg
5 Department of Neurology, Columbia University Irving Medical Center, New York, NY, USA
6 Laboratory of Neurophysiology and Plasticity, IRCCS Fondazione Santa Lucia, Rome, Italy
7 Institute of Biology, University of Hohenheim, Garbenstrasse 30, 70599 Stuttgart, Germany
8 NGS Competence Center Tuebingen, Institute of Medical Genetics and Applied Genomics, University of Tuebingen, Tuebingen, Germany
9 Department of Oncology, Luxembourg Institute of Health (LIH), Strassen, Luxembourg
10 Department of Human Genetics, Faculty of Medicine, Ruhr University Bochum, Bochum, Germany
11 Department of Neuromuscular Diseases, UCL Queen Square Institute of Neurology, London, UK
12 Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy
13 IRCCS C. Mondino Foundation, Pavia, Italy
14 Translational Neuroscience, Luxembourg Centre for Systems Biomedicine (LCSB), University of Luxembourg, Belvaux, Luxembourg
15 Parkinson Research Clinic, Centre Hospitalier de Luxembourg (CHL), Luxembourg

Correspondence to: Dr Fubo Cheng
Institute of Medical Genetics and Applied Genomics, University of Tuebingen, Tuebingen, Germany
E-mail: fubo.cheng@med.uni-tuebingen.de

Running title: THAP1 is a regulator of SP1 family

Abstract

DYT6 dystonia is caused by mutations in the transcription factor THAP1. THAP1 knock-out or knock-in mouse models revealed complex gene expression changes which are potentially responsible for the pathogenesis of DYT6 dystonia. However, how THAP1 mutations lead to these gene expression alterations and whether the gene expression changes are also reflected in the brain of THAP1 patients are still unclear. In this study we used epigenetic and transcriptomic approaches combined with multiple model systems (THAP1 patients' frontal cortex, THAP1 patients' iPSCs derived midbrain dopaminergic neurons, THAP1 heterozygous knock-out rat model, and THAP1 heterozygous knock-out SH-SY5Y cell lines) to uncover a novel function of THAP1 and the potential pathogenesis of DYT6 dystonia. We observed that THAP1 targeted only a minority of differentially expressed genes caused by its mutation. THAP1 mutations lead to dysregulation of genes mainly through regulation of SP1 family members, SP1 and SP4, in a cell type dependent manner. Comparing global differentially expressed genes detected in THAP1 patients' iPSCs derived midbrain dopaminergic neurons and THAP1 heterozygous knock-out rat striatum, we observed many common dysregulated genes and 61 of them are involved in dystonic syndromes related pathways, like synaptic transmission, nervous system development, and locomotor behavior. Further behavioral and electrophysiological studies confirmed the involvement of these pathways in THAP1 knock-out rats. Taking together, our study characterized the function of THAP1 and contributes to the understanding of the pathogenesis of primary dystonia in human and rat. As SP1 family members are dysregulated in some neurodegenerative diseases, our data may link THAP1 dystonia to multiple neurological diseases and may thus provide common treatment targets.

Keywords: primary dystonia; THAP1 dystonia; epigenetics; SP1 family; therapeutic targets

Abbreviations: THAP1=Thanatos-associated domain-containing apoptosis-associated protein 1; DEGs = differentially expressed genes; iPSCs = induced pluripotent stem cells; mDA = midbrain dopaminergic neurons; MEME = Multiple Em for Motif Elicitation; ChIP = Chromatin Immunoprecipitation; CNS: Central nervous system; ChIs = striatal cholinergic interneurons

Introduction

THAP1 (Thanatos-associated domain-containing apoptosis-associated protein 1) belongs to the family of THAP proteins, characterized by the presence of an evolutionary conserved protein motif at their amino-terminal, the THAP domain (Thanatos-associated domain).[1] The THAP domain was originally recognized as a motif similar to P element transposase. Further studies proved that the THAP domain of THAP1 is a C2CH containing zinc-dependent DNA-binding domain which may directly bind to sequence-specific genomic regions to regulate gene expression.[2]

Mutations in the *THAP1* gene are responsible for DYT6 dystonia, the second most common subtype of primary dystonia.[3] As THAP1 is designated as transcription factor, the most possible pathogenetic pathway of DYT6 dystonia would be that THAP1 mutations lead to dysregulation of its target genes which are responsible for DYT6 syndromes. Until now, only few target genes were reported. *RRM1* is the first reported target gene of THAP1. Through regulating RRM1 expression THAP1 controls the endothelial cell differentiation.[4] TOR1A is another target gene of THAP1, which may link DYT1 with DYT6.[5,6] Unexpectedly, all DYT6 cellular and animal models did not show any expression changes of TOR1A. Other targets of THAP1, like SOD2, were reported by our previous study.[7] However, none of these target genes could link THAP1 mutation to dystonic syndromes. Evidences from ChIP-seq and RNA-seq analysis showed that mouse THAP1 binds to only small part (around 5%) of differentially expressed genes (DEGs) detected in THAP1$^{-/-}$ mouse embryonic stem cells, which may indicate that THAP1 deletion lead to dysregulation of genes mainly through indirect ways,[8] for example through regulation of other key transcription factors.

Based on mouse models and using of transcriptomic approaches it was show that many DEGs caused by THAP1 mutation/ deletion are involved in a diversity of pathways, some of these are overlapping with other types of primary dystonia, like DYT1 and DYT25.[9,10] For example, genes related to eIF2α signaling and synaptic plasticity pathway found in DYT25 were dysregulated in both THAP1 heterozygous knock-out and THAP1 C54Y heterozygous knock-in mice.[10] In THAP1 central nervous system (CNS) conditional knock-out mice, myelination initiation associated genes, as well as several other dystonia subtype associated genes, were dysregulated.[9,11] A recent study using THAP1 patients' iPSC differentiated into cortical neurons revealed the potential mechanisms related to the disease penetrance.[12] However, the gene expression changes and affected pathways in iPSCs derived midbrain dopaminergic neurons (mDA) of THAP1 patients' as well as other species like rat are far less clear so far. Identifying the key transcription factors which link THAP1 to DEGs in human brain may provide new therapeutic targets though which we could correct the expression of DEGs found in THAP1 dystonia.

In this study, using transcriptomic and epigenetic approaches we identified pathways how THAP1 mutations cause dysregulation of a large number of genes. The core regulators are

mainly SP1 family members, as they were dysregulated in different model systems of THAP1 dystonia. Comparing DEGs of THAP1 iPSCs derived mDA and of THAP1 knock-out rat striatum, we identified a large number of common dysregulated genes related to dystonic syndromes. Further behavioral and electrophysiological studies confirmed these changes. Taking together, our results characterized the mechanisms by which THAP1 mutations lead to dysregulation of genes and revealed the potential pathogenesis of DYT6 dystonia in human and rat, which may provide new potential therapeutic targets.

Material and methods

Cell culture and THAP1 / SP1 knock-out SH-SY5Y cell lines generation

All the cell lines were karyotyped and authenticated before used for experiment (Supplementary data). Mycoplasma contamination was tested regularly every three months. SH-SY5Y (ATCC® CRL-2266™) cells cultured in DMEM medium with 10% FBS and 1% Glutamax at 37 °C in a humidified atmosphere containing 5% CO2.

THAP1 and SP1 knock-out cell lines were generated by CRISPR/Cas9 technique. Guide RNA (gRNA) targeted exon2 of THAP1 gene (TGAACAAATACTGCTATACT) or exon3 of SP1 gene (GTATGTGACCAATGTACCAG) were designed by GPP sgRNA Designer (Broad Institute) and clone into pX459 version 2.0 vector (Addgene Plasmid: 62988). Transfection was done by co-transfect of pX459/gRNA and pBabe empty vector (ratio 10:1) into dopaminergic neuroblastoma cells (SH-SY5Y cells) using lipofectamine 2000. 24 hours after transfection, cells were selected by puromycin (final concentration 2 ug/ ml) for another 48 hours. The surviving cells were sub-cultured in 10 cm dishes until single clone formation. Sanger sequencing and western blot were used to genotype cell clones and verify THAP1 protein expression.

Western blotting

Western blot analysis was performed as described previously using the following antibodies: THAP1 (1:1000, Proteintech), H3K27ac (1:1000, Cell Signaling), H3K4me3 (1:1000, Cell Signaling), SP1 (1: 1000, Cell Signaling), Actin (1:5000, Cell Signaling).[7] Proteins were separated in SDS-PAGE gels and transferred to nitrocellulose membranes. The primary antibodies were incubated at 4°C overnight. After incubated with secondary antibody, the blot membrane was directly detected in Odyssey® Fc Dual-Mode Imaging System and analyzed in Image Studio (LI-COR Biosciences). All western blots were repeated at least three times.

Generation of induced pluripotent stem cells (iPSCs)

Ethical approval for the development of and research pertaining to patient-derived cell lines has been given by the National Committee for Ethics in Research (CNER, Luxembourg). Native fibroblasts were obtained from skin biopsies with the written informed consent of donors. Fibroblasts were maintained and cultured as previously described.[13] In total, four fibroblasts from two male patients with the same *THAP1* mutation (c.197_198 delAG; p.E66fs84X) (age at 45 and 47 years old, respectively) and two age-and gender-matched neurological normal controls were used to generate iPSCs. Molecular karyotyping and identity analysis was performed using HumanOmni 2.5 Exome-8 DNA Analysis BeadChip (Life&Brain GmbH). The iPS cells were cultured as previously published.[13-15]

Directed differentiation of iPSC into midbrain dopaminergic neurons (mDA)

Midbrain dopaminergic (mDA) neurons were differentiated from iPSCs according to the previously reported protocols.[13,16] iPSCs were first derived into neural precursor cells (NPCs). Briefly, N2B27 media (composed of 50% Neurobasal (21103049, Thermo Fisher Scienfic), 50%DMEM/F12 (21331046, Thermo Fisher Scienfic), 0.5% N2 (17502048, Thermo Fisher Scienfic), 1% B27-R.A.(12587-010, Thermo Fisher Scienfic), 1% Glutamax (35050-061, Thermo Fisher Scienfic) and 1% Penicillin/Streptomycin (15140-163, Thermo Fisher Scienfic)) supplemented with 10 µM SB431542 (Sigma, S4317), 1 µM Dorsomorphin (Sigma, P5499), 3 µM CHIR99021 (Axon, 1386), and 500 mM Purmorphamine (Sigma, SML0868) was used as iPSCs culture media for four days led to derivation of neural epithelial structures (NESCs). Afterwards, SB431542 and Dorsomorphin were removed and Ascorbic Acid (150 µM, Sigma, A8960) was added to the culture media. The NESCs were manually picked and enzymatically purified using Accutase® (Life Technologies, A6964) for a minimum of 5 passages before undergoing characterization and being hereafter referred to as NPCs.

The NPCs were directly differentiated in N2B27 media containing 1 µM Purmorphamine, 200 µM Ascorbic Acid, and 100 ng/mL FGF8b (Peprotech, 100-25). After 8 days of differentiation, the culture medium was change to N2B27 supplemented with 0.5 µM Purmorphamine and 200µM Ascorbic Acid for another two days. From 10 days of directed differentiation until the designated experimental timepoint (d45 onwards) the following substances were added to N2B27 media, BDNF (10 ng/Ml; Peprotech, 450-02), GDNF (10 ng/mL; Peprotech, 450-10), TGFβ3 (1 ng/mL; Peprotech, 100-21), cAMP (500µM; Applichem, A0455) and Ascorbic Acid (200µM).

ChIP-seq and ChIP-qPCR analyses

Chromatin immuneprecipitation (ChIP) sample preparation and next generation sequencing data analysis were done according to the ChIP-seq guidelines of ENCODE as previously described.[7,17] Briefly, cells were crosslinked by 1% formaldehyde and sonicated in RIPA buffer to 200-500 bp. Approximately 10 ng of immunoprecipitated DNA was processed using the NEB

Ultra II DNA Library Preparation Kit for Illumina (NEB, E7103L). Next generation sequencing was done on Illumina NextSeq500 / NovaSeq 6000 system using the 75bp / 100bp single-end high output sequencing kit. Briefly, ChIP-seq raw reads were aligned to the hg19 genome assembly using BWA (v0.7.17). ChIP-seq peaks calling were analyzed using MACS2 (v2.1.1.20160309).[18] Transcription binding motif was predicted by Multiple Em for Motif Elicitation (MEME).[19] ChIP-seq target genes were analysis using RStudio (verion 4.0.3). The target genes were defined as the ChIP-seq binding signals on either the gene body or maximum 10 kb upstream of TSS sits. At least two biological ChIP-seq replicates were performed for each group.

Quantitative analysis of ChIP results was made by ChIP-qPCR. The ChIP eluted DNA was used for quantitative real-time qPCR analysis. Each experiment was repeated at least three times. Primers used for ChIP-qPCR were listed in supplementary table 1.

Generation of a THAP1 knock-out rat model

All animal experiments were approved by the local Ethics Review Committees (Regierungspräsidium Tübingen) (Animal research number: HG6/14) and performed in accordance with the Animal Welfare Committee guidelines of University Hospital Tuebingen, Tuebingen. To prevent environmental bias, animals with different genotypes were co-housed. THAP1 knock-out rat model was generated in Sprague Dawley (SD) rats by using CRISPR/Cas9 system as previous reported.[20] Guide RNA targeted exon 1 of rat THAP1 gene (gRNA: GTGCAGTCCTGTTCCGCCTA) was designed using public software (crispr.mit.edu) and subcloned into pX330 vector. SgRNAs were in vitro transcribed using MEGAshortscript T7 kit (Life Technologies) according to the manufacturer's instruction.

The Cas9 mRNA was in vitro transcript using PmeI-digested Cas9 expression JDS246 plasmid (Addgene plasmid # 43861) and the mMESSAGE mMACHINE T7 ULTRA Transcription Kit (Life Technologies) according to the manufacturer's instruction. The sgRNA and Cas9 RNA were used to generate knock-out rats by microinjection into the pronucleus of rat zygotes. Knock-out rats were generated and kept on SD background. Genotypes of offspring were determined by PCR and Sanger sequencing using DNA extracted from ear biopsies and specific primers (5'- CTGGGAACGTTACCTTCGGC-3' and 5'- CGGAGCACGCCAGTTACCTA-3').

RNA-seq analysis

Using of human brain tissue and iPSCs derived mDA neurons for RNA-seq was approved by ethic committee of University Hospital Tuebingen, Tuebingen, Germany (Project number: 565/2019BO2). Using rat striatum tissues (3 months old rat) for RNA-seq analysis was approved by the local Ethics Review Committees (Regierungspräsidium Tübingen) (Animal research number: HG6/14). Total RNA was isolated using RNeasy Mini Kit (QIAGEN) or RNeasy Lipid

Tissue Mini Kit (QIAGEN) following manufacturer's protocol. A total of 100ng total RNA was subjected to polyA enrichment and cDNA libraries preparation using NEBNext Ultra II Directional RNA library Preparation Kit for Illumina (NEB, #E7760L). Libraries were sequenced as paired-end reads on a NovaSeq6000 (Illumina) with a depth of at least 20 million reads each. Reads were aligned using HISAT2 (v2.2.1) to the Ensembl Homo Sapiens GRCh37 (hg19) or Ensembl Rat RGSC6.0 (rn6). Differential expression was determined in each of the pair-wise comparisons between experimental and control groups using DESeq (V1.22.1). Batch effects were analyzed using principal component analysis (PCA).

At least four male rats (3-month-old) of each group ($THAP1^{+/-}$ group and wild-type control group) were used. For iPSCs derived mDA neurons, three time's independent differentiation were done by using two THAP1 iPSCs clones and two control iPSCs clones. In total 12 independent iPSCs derived mDA neuron samples (6 sample of $THAP1^{+/-}$ neurons and 6 samples of control iPSCs derived mDA neurons) were used for RNA-seq. For human frontal cortex RNA-seq analysis, frozen post-mortem frontal cortex tissue were obtain from two THAP1 patients (with THAP1 p.S21F and p.R29Q mutation, respectively) and three neurological normal control individuals (Supplementary table 2). All the data related to these two THAP1 patients were published previously.[21] Sex and age matched three control samples were obtained from Brain bank Tuebingen. RNA-seq analysis was done twice in each individual.

Behavioral studies

All behavioral studies were performed by a trained observer blind to the genetic status. Rats of different genotype were housed together to avoid biasing based on environment. All animal studies were approved by the local Ethics Review Committees (Regierungspräsidium Tübingen) (Animal research number: HG6/14). Only male rats were included in all behaviour studies.

Tail suspension test

Tail suspension test was performed as described previously to observe the limb clasping behavior.[22] Rats were hung by their tail for at least 40 second to observe forelimb and hind limb clasping. Rats were periodically tested by the presence of abnormal limb clasping every month between 3 months and 1 year of age (n=14 rats in each group).

Rotarod test

The rotarod test was performed as described previously.[22] Three training sessions were followed by two test sessions at the age of 3-month-old. Rats received 12 training sessions in 3 consecutive days. Each training session lasted 2 min. Two tests were performed directly after the training days. The latency to fall was recorded and the mean of two tests was taken for the analysis. If the rat did not fall within 5 min, the test would be ended and recorded as 5 min.

Locomotor movement assessment

The locomoter activity measurement was conducted with the PhenoMaster homecage apparatus (TSE systems) as described previously.[22] Rats' spontaneous horizontal and vertical activities were measured continuously for 22 hours and analyzed with custom-made software. The activity detection is achieved using infrared sensors arranged in horizontal (x,y) level for ambulatory activity and vertical (z) level for rearing. A cohort of 10 males of Thap1 $^{+/-}$ rats and 10 wild-type male littermates were analyzed. The test was performed at the age of 3, 6, 9, and 12 moths.

Fine motor skill tests

Fine motor skills were measured using skilled reaching test according to our previously published protocol.[23] A cohort of 8 male Thap1$^{+/-}$ rats and 8 wild-type male littermates at the age of 5-month-old were analyzed for skilled reaching test. The reaching behavior was analyzed by measuring total success, first attempt success, and total number of attempts.

Single pellet reaching boxes were made of clear Plexiglas (35 cm wide, 35 cm long and 35 cm high). Five days before the training session, the rats were habituated daily into the reaching cage with food pellets on the shelf for 15 min. When 80% out of 100 reach attempts were made with one limb during this single session, that limb was identified as the preferred one.

In the subsequent training sessions, food pellets were located contralateral to this limb at a distance of 1.5 cm from the slit openings. At this stage, individual food pellets were placed into the indentation contralateral to the preferred paw. During the subsequent 11 days, rats continued to receive daily 15 min training sessions, consisting of discrete trials (10 pellets for warm up and 20 pellets for scoring). Rats were given one test session per day until they reached a stable performance. The success rate of the rats' reaching attempts was used as outcome measure. Reaches were considered successful if the rat had grabbed the pellet on the first reach attempt and managed to retrieve and eat it without dropping it into the cage bedding. All other behaviors were considered failure. Success rates were calculated as the number of successful reaches out of the final 20 pellets offered at each session.

Reaching behaviour was analyzed by measuring: (1) Total success % = (number of pellets obtained / 20) × 100%; (2) Success on first % = (number of pellets obtained on first advance/20) × 100%; (3) Total number of attempts = Total number of attempts to reach 20 pellets.[24]

Electrophysiological analysis

The experiments were approved by the ethics committee of IRCCS Fondazione Santa Lucia, and authorized by the Italian Ministry of Health (authorization nr. 223–2017-PR). Procedures were conducted in accordance with the Italian law (D.Lgs 26/2014), and the European Union Directive 2010/63/EU. Rats were sacrificed by decapitation under halothane anaesthesia, the brains were immediately removed from the skulls and cut with a vibratome (Leica

Microsystems) in Krebs' solution (artificial cerebrospinal fluid, aCSF), in mM: NaCl (126), KCl (2.5), MgCl2 (1.3), NaH2PO4 (1.2), CaCl2 (2.4), glucose (10), NaHCO3 (18), bubbled with 95% O2 and 5% CO2. Coronal corticostriatal slices (200 µm thick) recovered in oxygenated aCSF for about 30–60 min, and then were transferred into a recording chamber, continuously superfused with oxygenated aCSF.

Recordings were made with AxoPatch 200B and Multiclamp 700B amplifiers, coupled to pClamp 10 software (Molecular Devices). Borosilicate glass pipettes (resistance 2.5–5 MΩ) were pulled on P-97 pullers (Sutter Instruments). For patch-clamp recordings, neurons were visualized using differential interference contrast and infrared optics, and a monochrome CCD camera. For whole-cell recordings of striatal cholinergic interneurons (ChIs), electrodes were filled with an intracellular solution containing the following (in mM): K^+-gluconate (125), NaCl (10), CaCl2 (1.0), MgCl2 (2.0), 1,2bis (2-aminophenoxy) ethane-N,N,N,N-tetra-acetic acid (BAPTA) (0.1), HEPES (10), GTP (0.3) Mg -ATP (2.0); pH 7.3. Membrane currents were continuously monitored and access resistance, measured in voltage-clamp, was within the 5–30 MΩ range. Basic electrophysiological properties of ChIs, such as resting membrane potential (RMP), input resistance (IR), rheobase and current-voltage relationship, were evaluated in the current-clamp configuration. Drugs were purchased from Tocris Cookson and Sigma Aldrich and diluted to the final concentration into aCSF immediately before slice perfusion.

Statistical analyses

The statistical analysis of RNA-seq and ChIP-seq including sample size and replications, were described in each method section, respectively. Sample size was determined based on previous experiences. Two-tailed unpaired t-test was performed when two groups were compared. Statistical analyses were performed using the Prism software 8.0 (GraphPad Software, La Jolla, CA).). Significant differences were considered when *$p \leq 0.05$; **$p \leq 0.01$, ***$p \leq 0.001$, and ****$p \leq 0.0001$.

Data availability

Some ChIP-seq datasets were obtained from NCBI GEO, using the following GEO Series accession numbers: K562 THAP1 ChIP-seq (GSM803408), hESC SP1 ChIP-seq (GSM803377), hESC SP2 ChIP-seq (GSM1010776), and hESC SP4 ChIP-seq (GSM1010743).

RNA-seq and ChIP-seq datasets generated in this study are available through the Gene Expression Omnibus (GEO) via accession GSE141278 and GSE184961. All other data are available from the corresponding author upon reasonable request. Supplementary data were uploaded to Mendeley Data (https://data.mendeley.com/datasets/b3sh9b3th6/1)

Results

THAP1 targets a minority of DEGs found in THAP1$^{+/-}$ SH-SY5Y cell lines

In order to characterize the target genes of THAP1 and differentially expressed genes (DEGs) caused by its mutations, we first generated THAP1 knock-out SH-SY5Y cell lines using CRISPR/Cas9 technique. Two independent cell lines carrying two different mutations in exon2 of the *THAP1* gene were generated (with *THAP1* c.149insG/ p.Y50X and c.150insA/ p.Y50X mutations), respectively (Supplementary figure 1A). Western blot analysis showed significantly decreased expression of total THAP1 protein in both cell lines (Figure 1A). Using RNA-seq analysis, we detected 279 differentially expressed genes (DEGs), including 138 down-regulated and 141 up-regulated genes, in THAP1$^{+/-}$ cell lines compared to control cell lines (FDR < 0.01) (Figure 1B, Supplementary data). DAVID Gene ontology analysis showed "neurotransmitter secretion" as the only enriched pathway associated with down-regulated genes, while "extracellular matrix organization" is the most significantly affected pathway associated with up-regulated genes (Figure 1C, Supplementary figure 1B).

Next, we performed THAP1 ChIP-seq using wild-type SH-SY5Y cells to examine the overlap between DEGs found in THAP1$^{+/-}$ cell lines and direct targets of THAP1. ChIP-seq analysis revealed that THAP1 mainly binds to gene promoter regions (near transcription start sites, TSS) and its peaks are surrounded by both H3K27ac and H3K4me3 peaks (Figure 1D and E). MEME motif discovery analysis revealed two potential binding motifs of THAP1 (Figure 1F). Target gene analyses showed that THAP1 binds to a total of 2960 genes (Figure 1G) (THAP1 binds either to the gene body or maximum 10 kb upstream of TSS of these genes). However, comparing THAP1 target genes with DEGs caused by mutant THAP1, we observed only a few overlapping genes (7.5%, 21/279 genes) (Figure 1G). These data indicate that THAP1 mutations lead to dysregulation of genes mainly through indirect pathways.

Motif discovery analysis linked SP1 with the alterations of histone modifications in THAP1$^{+/-}$ SH-SY5Y cell lines

Next, we sought to analyze how THAP1 mutation leads to dysregulation of its indirect targets. We first performed H3K27ac and H3K4me3 ChIP-seq to show global changes of gene promoters and *cis*-regulatory elements (CRE) activities after heterozygous knock-out of THAP1, as H3K27ac and H3K4me3 are highly correlated with gene expression levels and H3K27ac is also an effective marker to determine active enhancers.[25,26] Differential binding analysis revealed a total of 1921 genomic regions that changed the H3K27ac ChIP-seq signals (diff-H3K27ac) in THAP1$^{+/-}$ cell lines compared to control cell lines, including 333 regions with decreased signal and 1588 regions with increased signal (Figure 2A and 2B, Supplementary figure 2A). Differential binding analysis of H3K4me3 ChIP-seq revealed that a total of 2762 regions changed their ChIP-seq signal (diff-H3K4me3 regions) in THAP1$^{+/-}$ cell lines compared to

control cell lines. Most of these regions (2746/2762) showed increased signals, while only few of them (16/2762) showed decreased signals (Figure 2A and 2B, Supplementary figure 2B).

Next we sought to know which transcription factor / histone modification enzyme affects the histone modifications in THAP1$^{+/-}$ SH-SY5Y cells. We first discovered the common DNA motifs of the regions with changed histone modifications and then matched these common motifs to the motif of known transcription factors. As most of diff-H3K27ac and diff-H3K4me3 regions showed increased ChIP-seq signals, we first investigated whether the protein level of H3K27ac and H3K4me3 are globally changed. Unexpectedly, we did not observe significant changes of H3K27ac and H3K4me3 proteins (Figure 2C). In order to exclude ChIP-seq bias due to slightly increased H3K27ac and H3K4me3 protein levels, we used only the top regions with highest fold change of ChIP-seq signals for binding motif discovery analyses (top n = 606 diff-H3K27ac regions and top n = 476 diff-H3K4me3 regions). MEME motif discovery analysis revealed three motifs matched to top diff-H3K27ac regions and another three motifs matched to top diff-H3K4me3 regions (Figure 2D). Some transcription factors binding motif could match to these motifs significantly (Table 1). For example, ZNF384 and SP1 matched to motifs found in both diff-H3K27ac and diff-H3K4me3 (Figure 2E and Table 1). Interestingly, SP1 is a target gene of THAP1 and is up-regulated in THAP1$^{+/-}$ SH-SY5Y cells. ZNF384 also matched to motif 1 of diff-H3K27ac and motif 1 of diff-H3K4me3, but THAP1 does not bind to the ZNF384 promoter or gene body and its expression remains normal in THAP1$^{+/-}$ SH-SY5Y cells. Other SP1 family members, like SP2 and SP3, also showed similarity to the discovered motifs of diff-H3K4me3, but their mRNA levels remained unchanged in THAP1$^{+/-}$ SH-SY5Y cells (Table 1, Supplementary data). These data indicate that SP1 may be responsible for the changes of histone modifications in THAP1+/- SH-SY5Y cell lines.

SP1 up-regulation is mainly responsible for DEGs in THAP1$^{+/-}$ SH-SY5Y cell lines

Epigenetic analysis indicates that SP1 may be responsible for the H3K27ac and H3K4me3 modification alterations in THAP1$^{+/-}$ SH-SY5Y cells. Next, we performed MEME motif discovery analysis to predict common upstream motifs of DEGs found in THAP1$^{+/-}$ SH-SY5Y cells by using the promoter regions of these DEGs as input. As expected, SP1 binding motif matched to one of the predicted up-stream motifs (Supplementary figure 2C and 2D), which may suggest that up-regulation of SP1 is responsible for DEGs found in THAP1$^{+/-}$ SH-SY5Y cells. In order to further substantiate this indication, we first checked the protein level of SP1 and observed significantly increased protein levels in THAP1$^{+/-}$ SH-SY5Y cells (Figure 3A). RNA-seq analysis also showed increased SP1 mRNA levels in THAP1$^{+/-}$ SH-SY5Y cells (Figure 3B). THAP1 ChIP-seq showed strong binding signal to the *SP1* promoter and a weak binding signal to the potential enhancer in intron 3 of the *SP1* gene (Figure 3B). H3K27ac and H3K4me3 ChIP-seq showed increased signals on *SP1* promoter and potential enhancer regions in THAP1$^{+/-}$ cell lines compared to control cell lines (Figure 3B).

Next, we performed SP1 ChIP-seq to confirm whether SP1 binds to the DEGs observed in THAP1$^{+/-}$ cell lines. ChIP-seq analysis revealed that SP1 binding sites mainly enriched around TSS sites and overlapped with H3K27ac and H3K4me3 peaks (Figure 3C and 3D). MEME motif discovery analysis showed that the SH-SY5Y SP1 ChIP-seq binding motif is similar to the motif found in databases (Figure 2E and Figure 3E). The THAP1 and SP1 binding regions overlapped with half of diff-H3K27ac and diff-H3K4me3 regions (Figure 3F). Interestingly, SP1 has more peaks which overlapped with diff-H3K27ac and diff-H3K4me3 regions compared to THAP1 (Figure 3F). These data indicate that SP1 but not THAP1 itself may be mainly responsible for the change of H3K27ac and H3K4me3 modifications in THAP1$^{+/-}$ cell lines. Similarly, nearly half of DEGs (119 out of 279 genes) found in THAP1$^{+/-}$ cell lines are the target genes of SP1 (Figure 3G). Furthermore, heterozygous knock-out of THAP1 (THAP1$^{+/-}$) could strengthen the SP1 binding activities and alter the histone modifications of SP1 binding sites (Figure 3H, Supplementary figure 3C). Further evidence from DEGs analysis of THAP1$^{+/-}$ / SP1$^{-/-}$ cell lines also supports our hypothesis. We observed a large number of DEGs (1101 down-regulated genes and 1111 up-regulated genes, FDR < 0.01) in THAP1$^{+/-}$ / SP1$^{-/-}$ SH-SY5Y cells compared to wild-type SH-SY5Y cells (Figure 3I and 3J, Supplementary figure 3A and 3B), while more than half of the dysregulated genes (66 out of 138 down-regulated genes and 78 out of 141 up-regulated genes) found in THAP1$^{+/-}$ SH-SY5Y cells were either normalized or reversed their expression changes after further knock-out SP1 (Figure 3J and 3K, Supplementary figure 3). There are still some genes which do not change their expression after knock-out SP1, some of them might belong to direct or indirect targets of THAP1 outside of SP1. Nevertheless, all these data indicate that THAP1 up-regulates expression of SP1 which in turn is a key factor for most of the dysregulated genes found in THAP1$^{+/-}$ SH-SY5Y cell lines.

THAP1 mutation causes dysregulation of genes related to dystonic phenotypes in human and rat

Next, we investigated what genes are dysregulated in iPSC derived mDA neurons of THAP1 patients and in striatum tissues of THAP1 knock-out rat, as both mDA neurons and striatum are key cells / brain regions responsible for primary dystonic symptoms. For this purpose we first generated iPSC cells using fibroblasts collected from THAP1 patients and then differentiated them into mDA neurons (Supplementary figure 4). Unbiased RNA-seq analysis was performed to detect global gene expression changes. In THAP1 iPSCs derived mDA neurons we observed 2288 down-regulated and 3092 up-regulated genes (FDR<0.01, Log_2(Fold change)>1) compared to control iPSCs derived mDA neurons (Figure 4A, Supplementary 6A, Supplementary data). Gene ontology analysis showed that the most significantly affected pathway is “chemical synaptic transmission” in the group of down-regulated genes, while “SRP-dependent co-translational protein targeting to membrane” is the most significantly affected pathway of the up-regulated genes (Supplementary figure 6B).

Next, we generated a Thap1 knock-out rat model using CRISPR/Cas9 technique (Supplementary figure 5A). We observed decreased THAP1 protein levels in the striatum of THAP1$^{+/-}$ rat

(Supplementary figure 5B). Cross-breeding and embryo analysis indicated that homozygous knock-out THAP1 lead to embryonic lethality similar to THAP1 knock-out and knock-in mouse models,[27] as we did not obtain any $THAP1^{-/-}$ rat by cross-breeding $THAP1^{+/-}$ male rats with $THAP1^{+/-}$ female rats, and early lethal embryos were observed at day E12 (12 days after fertilization) (Supplementary figure 5C and 5D). We next performed unbiased RNA-seq analysis and selected the top 1000 up- or down-regulated DEGs for further analyses (including 1254 down-regulated genes and 1153 up-regulated genes, with FDR<0.01, fold change>20%) (Figure 4B, Supplementary figure 6C, Supplementary data). Gene ontology analysis revealed that "chemical synaptic transmission" is the most significantly affected pathway of down-regulated genes, while "SRP-dependent co-translational protein targeting to membrane" is the most significantly affected pathway of up-regulated genes, which are similar to the results of gene ontology analysis of iPSC derived mDA neurons DEGs (Supplementary figure 6D).

When comparing the DEGs of THAP1 patients' iPSCs derived mDA neurons and DEGs of $THAP1^{+/-}$ rat striatum, we observed many common dysregulated genes (Figure 4C). Gene ontology analysis showed that the most significantly affected pathway of common up-regulated genes is "SRP-dependent co-translational protein targeting to membrane", which is the same as the pathway found in both THAP1 iPSCs derived mDA neurons and $THAP1^{+/-}$ rat striatum up-regulated genes, respectively (Supplementary figure 6). Gene ontology analysis of common down-regulated genes showed that "chemical synaptic transmission" and "regulation of synaptic transmission, glutamatergic" ("Synaptic transmission" in short for both pathways) are the top two significantly affected pathways (Figure 4D). In total 33 common down-regulated genes are involved in these two pathways (Figure 4E). Other pathways potentially related to pathogenesis of primary dystonia, like "nervous system development" and "locomotory behavior" pathways, were also identified and 28 down-regulated DEGs were involved in (Figure 4E). These pathways could reflect the clinical syndromes of primary dystonia.

The fold-changes of dysregulated genes related to dystonic syndromes in THAP1+/- rat striatum are lower than in human iPSCs derived mDA neurons (Figure 4E). As THAP1 has autoregulatory function,[28] this observation may be due to the fact that THAP1 has different autoregulatory activities in different species or in different cell types. We thus checked the level of wild-type THAP1 allele in iPSCs derived mDA neurons and in rat striatum, as it can reflect the autoregulatory activity of THAP1. Unexpectedly, we observed strong autoregulatory activity of THAP1 in iPSCs derived mDA neurons and SH-SY5Y cells but not in rat striatum, the amount of wild-type THAP1 allele in THAP1 iPSCs derived mDA neurons reached to in average 83% compared to control iPSCs derived mDA neurons (Figure 4F). In SH-SY5Y cells this ratio reaches to 65% (Supplementary figure 6E). However, the amount of wild-type THAP1 allele in $THAP1^{+/-}$ rat striatum is only 53% of all THAP1 in wild-type rat striatum (Figure 4F).

THAP1$^{+/-}$ rats present behavior abnormalities and dopaminergic dysfunction

THAP1$^{+/-}$ rats present behavior abnormalities and dopaminergic dysfunction

As RNA-seq analysis revealed that many genes related to dystonic syndromes were dysregulated in THAP1$^{+/-}$ rats, we further performed electrophysiological and behavioral studies to prove the changes of dystonic related syndromes in THAP1$^{+/-}$ rats. The THAP1$^{+/-}$ rats did not exhibit any spontaneous abnormal motor behaviors. We first performed a series of behavior studies to characterize the behavioral abnormalities in THAP1$^{+/-}$ rats. Tail suspension test showed two THAP1 $^{+/-}$ rats present hind limb clasping or hind limb clasping like behavior in a cohort of 14 THAP1$^{+/-}$ rats (14.3%) (Figure 5A and 5B). One rat presented hind limb clasping like behavior at the age of 11 months, while another rat present typical hind limb clasping behavior at the age of nine months (Figure 5A and 5B). Rota-Rod tests did not show gross motor performance difference between THAP1$^{+/-}$ rat and wild-type rats (Supplementary figure 7). Next, we performed fine motor skill tests and observed lower performance scores in a skilled reaching task in THAP1$^{+/-}$ rats when compared to wild-type rats (Figure 5C). At training day 10 and 11, both, total success rate and success rate on first attempt of THAP1$^{+/-}$ rats are lower than in wild-type control rats (Figure 5C). The total attempt times for reaching the same amount of pellets are more in THAP1$^{+/-}$ rats than in wild-type rats (Figure 5C). Further phenomaster behavioral analysis showed decreased locomotor activities of THAP1$^{+/-}$ rats in night phase compared to wild-type rats at the age of 9 months (Figure 5D). We also observed hyperactivities in the light phase, as the overall distance and Z-scale distance in light phase were increased at the age of 9 months and 12 months in THAP1$^{+/-}$ rats compared to wild-type control rats (Figure 5E), which may indicate sleeping problems of THAP1$^{+/-}$ rats.

Next, we performed electrophysiological studies to characterize the striatal cholinergic interneurons (ChIs). The voltage responses to hyperpolarizing current steps are normal in THAP1$^{+/-}$ rats compared to wild-type rats (Figure 6A). The ChIs membrane properties also did not show abnormalities in THAP1$^{+/-}$ rats when compared to wild-type rats (Figure 6B). We further performed pharmacological studies to analyze the responses to different transmitter receptor agonists on ChIs firing frequency. The response of THAP1$^{+/-}$ ChIs to the muscarinic agonist oxotremorine was similar to the wild-type ChIs, both types of ChIs showed inhibitory responses (Figure 6C). Additionally, both THAP1$^{+/-}$ and wild-type ChIs showed a similar increased firing frequency when perfusing the striatal slices with the β-adrenoceptor agonist isoproterenol (Figure 6C). We then tested the responses of ChIs to dopaminergic agents. In wild-type rats, bath-application of 3 μM amphetamine, a dopamine releaser, did not significantly alter the spontaneous firing activity of ChIs, while increasing the concentration to 30 μM enhanced the action potential frequency of the recorded neurons (Figure 6D and 6E). In THAP1$^{+/-}$ rats, ChIs showed a similar excitatory response to 30 μM amphetamine, but 3 μM amphetamine caused a significant reduction of the ChIs firing frequency (p<0.05) (Figure 6D and 6E). Interestingly, this abnormal inhibitory response to a low amphetamine concentration could be abolished by the dopamine receptor 2 (DRD2) antagonist sulpiride (Figure 6E). As these changes

are related to DRD2 function, we thus checked the protein level of DRD2, but we did not observe DRD2 protein level changes in THAP1$^{+/-}$ rat striatum (Supplementary figure 5B). RNA-seq also did not show expression change of DRD2 in THAP1$^{+/-}$ rat striatum (Supplementary data). However, THAP1$^{+/-}$ ChIs showed an abnormal excitatory response to the DRD2 agonist quinpirole, in contrast to the slight inhibitory effect observed in wild-type ChIs (p<0.05) (Figure 6F). Overall, all these data support the hypothesis of a DRD2-dependent dopaminergic dysfunction in THAP1$^{+/-}$ rat striatum.

Cell type dependent expression changes of SP1 and SP4 may contribute to dystonic syndromes

In SH-SY5Y cells we observed that mutations of THAP1 lead to dysregulation of genes mainly through up-regulation of the transcription factor SP1. Thus, we sought to know whether THAP1 mutation leads to dysregulation of genes related to dystonic syndrome in THAP1 patients' iPSCs derived mDA neurons and THAP1$^{+/-}$ rat striatum are also due to up-regulation of SP1. We first checked the expression of SP1 in iPSCs derived mDA neurons and rat striatum in RNA-seq data and we observed significant increased expression of SP1 in THAP1 patients' iPSCs derived mDA neurons, but not in THAP1$^{+/-}$ rat striatum (Figure 7A). However, we observed significantly down-regulation of another SP1 family member, SP4, in both THAP1 patients' iPSCs derived mDA neurons and THAP1$^{+/-}$ rat striatum (Figure 7A). Next, we wanted to know whether SP4 is also the target gene of THAP1. After analyzing SH-SY5Y cells THAP1 ChIP-seq data and K562 cells THAP1 ChIP-seq data, we observed that THAP1 has binding activities to multiple SP1 family members, including SP1, SP2, and SP4 (Figure 7B). Comparing the expression of SP1, SP2, and SP4 in frontal cortex tissue of THAP1 patients, THAP1 patients' iPSCs derived mDA neurons, THAP1$^{+/-}$ SH-SY5Y cells, and THAP1$^{+/-}$ rat striatum, we observed cell / tissue types dependent expression changes of these genes. SP1 and SP2 are predominately up-regulated in all these cells / tissues except in THAP1 patients' frontal cortex (SP1 is significantly down-regulated in THAP1 patients' frontal cortex), while SP4 is predominately down-regulated in all these cell types / tissues except in THAP1$^{+/-}$ SH-SY5Y cells (SP4 expression remains normal in THAP1$^{+/-}$ SH-SY5Y cells) (Figure 7A).

After analyzing the binding motifs and binding profiles of three SP1 family members, we found that they shared very similar binding motifs (Figure 7C), but both ChIP-seq signal intensity and ChIP-seq binding profiles analyses revealed that SP4 showed more similarities to SP1 compared to SP2 (Supplementary figure 8A, 8C, and 8D). SP4 also shared more binding sites with SP1 than with SP2 (Supplementary figure 8B). As SP4 may repress the transcriptional function of SP1 by competing with its binding sites,[29] down-regulation of SP4 may somehow enhance the effect of up-regulation of SP1 in regulating expression of some genes.

Comparing the common dysregulated genes related to dystonic syndromes detected in iPSCs derived mDA neurons and rat striatum to the target genes of SP1 identified by ChIP-seq analysis, we observed a large number of overlapping genes (31 of 61 genes) (Figure 7D). Interestingly, the

expression fold-changes of those genes in THAP1 patients' iPSCs derived mDA neurons are larger than in THAP1$^{+/-}$ SH-SY5Y cells and in THAP1$^{+/-}$ rat striatum (Figure 7D), which is similar to the expression fold-changes of SP1 and SP4, as higher fold changes of SP1 and SP4 expression were observed in THAP1 patients' iPSCs derived mDA neurons than in other cells / tissues (Figure 7A). However, we observed opposite expression change of SP1 in THAP1 patients' frontal cortex compared to other cells/ tissues (SP1 expression is decreased in THAP1 patients' frontal cortex but increased in THAP1 patients' iPSCs derived mDA neuron and in THAP1$^{+/-}$ SH-SY5Y cells) (Figure 7A). Similarly, we observed opposite expression changes of some SP1 target genes in THAP1 patients' frontal cortex compared to other cells / tissues (Figure 7D). In global gene expression analyses of THAP1 patients' frontal cortex and THAP1 patients' iPSCs derived mDA neurons, we also observed a large number of genes (n=388) to be oppositely dysregulated (Supplementary figure 8E). As the brain striatum and mDA neurons rather than frontal cortex are key regions which are responsible for dystonia syndromes, all these data further support that cell type-dependent dysregulation of SP1 family members caused by THAP1 mutation may contribute to cell type dependent expression changes of genes related to dystonic syndromes.

Discussion

In this study, we found that THAP1 targets only 7.5% of dysregulated genes detected in THAP1$^{+/-}$ SH-SY5Y cells, which indicates that THAP1 mutations lead to dysregulation of genes mainly through indirect pathways. With the help of epigenetic approaches (H3K27ac and H3K4me3 ChIP-seq) and complementary transcriptome analysis of different model systems we showed that THAP1 mutation leads to gene expression changes mainly through regulating transcription factor SP1 family members, including SP1 and SP4, in a cell type dependent manner. The dysregulated genes detected in both human THAP1 patients' iPSCs derived mDA neurons and rat THAP1 knock-out rat striatum are enriched in pathways directly related to dystonic syndromes. Our data characterized the function of THAP1 and revealed the putative pathogenic mechanism of THAP1 mutations in both human and rat.

Several lines of evidence showed that SP1 family members are associated to other neurological diseases, like Parkinson's disease, Alzheimer diseases, and Huntington disease. For example, in MPTP injection induced Parkinson's disease model, SP1 expression was increased and knockdown of SP1 could rescue MPTP-induced α-synuclein accumulation, inflammation and apoptosis *in vitro*.[30] Similarly, SP1 was up-regulated in Huntington disease mouse brain, in frontal cortex of Alzheimer diseases patients, and in Alzheimer diseases transgenic mouse brain.[31-33] SP1 is ubiquitously expressed in almost all cell types, but the reasons that the effect of differential SP1 expression was observed mostly in neurodegenerative diseases are not very clear so far. This might partially due to the critical role of SP1 in neurons. For example, SP1 is important for neuronal dendrite outgrowth and synapse formation, and knock-out Sp1 leads to

reduced number of neurons in the cortex and hippocampus.[34] Another possibility might be that some cell-type dependent co-factor of SP1 or different SP1 modifications, like phosphorylation, affect its cell type dependent functions.[35]

As SP1 is dysregulated in some neurological diseases, manipulating its expression was proposed as a potential therapeutic target for these diseases. For example in Huntington disease cell model, RNAi knock-down of SP1 showed neuroprotective effect.[31] Using microRNA-375 to inhibit the function of SP1 could also ameliorate the damage of dopaminergic neurons and reduce oxidative stress and inflammation in Parkinson's disease rat model.[36] As SP1 also plays a critical role in the growth and metastasis of many tumor types,[37] some antitumor antibiotics (like mithramycin) which could compete the binding sites of SP1 were used to treat some neurological diseases experimentally. For example, mithramycin treatment could increase the survival rate, improve motor performance and attenuate dopaminergic neurotoxicity in Huntington disease mice.[38] Mithramycin treatment can also increase the lifespan of a fly model of Huntington disease.[39] Tolfenamic acid is a nonsteroidal anti-inflammatory drug (NSAID) and can induce the degradation of SP1 protein. Alzheimer diseases mice treated with Tolfenamic showed disruption of Alzheimer diseases pathologic processes, like reduction of APP and decreasing of both soluble and insoluble Aβ1–40 and Aβ1–42 levels.[40] As SP1 is clearly up-regulated in THAP1 patients' iPSCs derived mDA neurons and THAP1$^{+/-}$ SH-SY5Y cells, and knock-out of SP1 in THAP1$^{+/-}$ SH-SY5Y cells can rescue its gene expression changes, repressing the function of SP1 may be used as potential therapeutic target of primary dystonia, at least of DYT6 dystonia.

In this study, we observed consistent down-regulation of SP4 in THAP1 patients' iPSCs derived mDA neuron, THAP1 patients' frontal cortex tissue, and THAP1$^{+/-}$ rat striatum. Similar reduced expression of SP4 was also found in a THAP1 knock-out mouse model.[10] SP4 is a member of SP1 family and shares many similarities with SP1. In Alzheimer diseases and patients with frontotemporal lobar degeneration, the expression of SP4 was increased,[41, 42] which is opposite of the results found in THAP1 patients. These data, consistent with the absence of gross neurodegeneration,[21] also indicate the pathogenesis of primary dystonia is different from neurodegenerative diseases. However, decreased expression of SP4 may contribute to the pathogenesis of primary dystonia as well, as SP4 complete knock-out mice presented neurodevelopment problems.[43] Most interestingly, SP4 deficient mice showed abnormal sensorimotor symptom,[44] a common symptom found in dystonia patients.[45, 46]

Few studies reported gene expression changes in THAP1 knock-in / knock-out or conditional knock-out mouse models,[9-11] but the gene expression changes in THAP1 knock-out rat models or in human DYT6/THAP1 patients brain are unknown. In this study, we generated patients-derived iPSCs and differentiated them into midbrain dopaminergic neurons and also collected frontal cortex tissues from DYT6/THAP1 patients to analyze global gene expression changes caused by THAP1 mutations in human brain. In our THAP iPSC-derived mDA neurons, we selected two patients carrying the same *THAP1* deletion mutation (c.197_198 delAG) which produces truncated THAP1 protein. Although many different THAP1 mutations distributed over the entire

gene have been reported, nearly one-third of them are nonsense mutations which may produce truncated THAP1 protein or be decayed due to nonsense-mediated RNA decay. As the nuclear localization signal domain is damaged, these truncated proteins are not able to import into nuclei and probably have lost function in regulating transcription. As no gain-of-function mutation was reported so far, our THAP1 truncated mutation does represent most of the loss-of-function mutations, like nonsense mutations or mutations which cause decreased protein stability.[7]

We chose to differentiate iPSCs into dopaminergic neurons because the dopaminergic dysfunction was reported by many studies as one of the critical issues for the pathogenesis of primary dystonia. Striatum dopaminergic system is also the key node which is responsible for the pathogenesis of primary dystonia.[47] We therefore compared the gene expression changes of these two tissues / cells to find out common gene expression changes. These genes may highly associate to the pathogenesis of DYT6 dystonia in human and rat, which in turn help us to understand the roles / changes of dopaminergic system in DYT6 dystonia. The gene expression changes in frontal cortex of DYT6 patients were also analyzed by using RNA-seq analysis, in order to show the expression changes of DYT6 associated critical genes in this region, although frontal cortex may not be the key node responsible for DYT6 dystonia. Comparing the RNA-seq data from different tissues / cell types did also help us to analyze the cell / tissue type dependent function of THAP1. Further studies focusing on gene expression changes in THAP1 patients' brain basal ganglia regions or THAP1 iPSCs derived GABAergic medium spiny neurons may be still needed to further characterize the pathogenesis of DYT6 dystonia.

Using different model systems, we found that the changes of SP1 family members are cell type dependent (Figure 7A). THAP1 is, although ubiquitously expressed, present in almost all cell types, but its function might vary between them. For example, in mESC cells, THAP1 is critical for mESC survival and differentiation.[8] Mutations in THAP1 cause neurological disorders (DYT6 dystonia), which may also indicate the critical role of THAP1 in nervous system. However, the underlying cell type dependent function of THAP1 is not clear so far. Different modifications or cell type dependent interacting co-factors can all affect function of THAP1 in different cell types, just like the cell type dependent function of transcription factor SP1.[35] Further studies should be performed to define the cell type dependent role of THAP1.

Unbiased RNA-seq analysis showed a large number of genes to be dysregulated in THAP1 iPSCs derived mDA neurons and in rat striatum. Although both THAP1 target gene SP1 and THAP1 itself could regulate the cell proliferation, but cell proliferation is not among the highly significant pathways in the gene ontology analysis of our RNA-seq data (Supplementary figure 6D). This observation may again support the cell type dependent function of THAP1. Another possibility could be due to the fact that the cell proliferation of striatal neurons or iPSCs differentiated neurons are not activated completely. In striatum of our THAP1 heterozygous knock-out rat, the most significant affected pathways among the down-regulated genes are chemical synaptic transmission and nervous system development (Supplementary figure 6B). In THAP1 heterozygous knock-out and THAP1 C54Y knock-in mouse striatum, the most

significantly affected pathways are EIF2 signaling pathway and signaling by Rho family GTPases pathways, respectively,[10] while in THAP1 CNS conditional knock-out mice the nervous system development is the most significant affected pathway.[11] However, in another THAP1 CNS conditional knock-out mice myelination delaying was observed as the top affected pathway. [9] These observations may be due to the different species, the genetic background of the mice, or due to different THAP1 mutations that were used in the different studies.

Further electrophysiological characterization of the THAP1$^{+/-}$ rat model supported striatal dopaminergic dysfunction in THAP1 dystonia. Indeed, we observed an enhanced sensitivity of THAP1$^{+/-}$ striatal ChIs to low concentration of amphetamine, causing a significant inhibitory response, which is not observed in wild-type neurons. Low concentrations of amphetamine are expected to release a limited amount of dopamine, driving preferentially DRD2 activation. In line with this prediction, the enhanced inhibitory response of THAP1$^{+/-}$ ChIs to 3 µM amphetamine was prevented by a DRD2 antagonist. Multiple dystonia rodent models show an altered dopamine neurotransmission. In particular, TOR1A models show an aberrant excitatory response of striatal ChIs to DRD2 activation, instead of the canonical inhibitory effect observed in wild-type neurons,[48] that has been attributed to an altered intracellular signaling pathway.[49,50] Of note, our electrophysiological analysis shows a similar paradoxical excitatory response to DRD2 activation by quinpirole in the THAP1$^{+/-}$ rat. The apparent discrepancy between the excitatory responses to quinpirole in ChIs of THAP1$^{+/-}$ rats with respect to the enhanced inhibition caused by amphetamine might reside in the different agonist activating the DRD2, either a biased agonist or low concentrations of endogenous dopamine, respectively. Indeed, DRD2 receptors can act on different intracellular signaling pathways and their agonists are preferentially biased towards some of them. Notably, a similar abnormal excitatory effect of quinpirole has been recently reported in a THAP1$^{C54Y/-}$ mouse model.[51] Furthermore, significant reductions in caudate and putamen D2 receptor availability were also observed in both DYT1 and DYT6 patients.[52] Overall, our data further support the hypothesis of a central role of striatal dopaminergic neurotransmission in dystonia pathophysiology.

Similar to most of the primary dystonia rodent models, we did not observe clear spontaneous dystonic symptoms in our newly generated THAP1 heterozygous knock-out rat model. The typical dystonic behavior hind limb clasping occurred only in low percentage of rats (2/14) at the age of one year. One possibility of lower occurrence of hind limb clasping might be due to the disease penetrance of THAP1 dystonia, as the reported penetrance of THAP1 dystonia is about 60%.[53] Another possibility could be that the fold-changes of dysregulated genes related to dystonic syndromes in THAP1$^{+/-}$ rat striatum are lower than in human iPSCs derived mDA neurons (Figure 4E), which may lead to weaker phenotype in THAP1$^{+/-}$ rats compared to THAP1 patients. The strong autoregulation of THAP1 in human iPSC-derived mDA neurons may also support the critical function of THAP1 in human neurons compared to in rat striatal neurons (Figure 4F). Subsequently, mild loss of THAP1 protein may cause more severe consequences in human brain than in rat brain.

Further locomotor behavior analyses revealed decreased locomotor activity at dark phase. Interestingly, light phase hyperactivity was observed in THAP1$^{+/-}$ rats, which may indicate the sleeping problem of THAP1$^{+/-}$ rats. Sleep impairment is also a common non-motor symptom observed in many dystonia patients.[54] Humans depend heavily on dexterous skill which requires precisely coordinated multi-joint and digit movement. Functional abnormalities like dystonia may affect this skill.[55] We thus performed single pellet reaching test to evaluate it. THAP1$^{+/-}$ rats showed deficiency in fine motor skill learning and coordination compared to wild-type rats, which further proved dystonic related symptom in the THAP1$^{+/-}$ rat model.

Taking together, our data showed that THAP1 mutations lead to dysregulation of genes mainly through regulating SP1 family members in human and rat brain. RNA-seq revealed the gene expression changes and affected pathways in THAP1 patients' brain as well as in THAP1 knock-out rat brain. Further behavioral and electrophysiological studies confirmed the changes of dystonic symptoms in our THAP1 knock-out model. In this study we characterized the function of THAP1 and pathological mechanism of DYT6/THAP1 dystonia in depth, which may provide new insight into therapeutic strategies of primary dystonia, like through correction the expression changes of SP1 family members.

Acknowledgements

We would like to thank the DYT6 patients for donating their skin tissue for this research. We want to thank the Biobank Hertie Institute for Clinical Brain Research, University of Tuebingen, Germany for providing control fibroblasts for this study. We also want to thank NICHD Brain and Tissue Bank for Developmental Disorders, Baltimore, MD, USA and Harvard Brain Tissue Resource Center (HBTRC), McLean Hospital, Belmont, MA, USA for providing DYT6 brain tissues for this study. We acknowledge computing support by Galaxy project (https://usegalaxy.org).

Funding agencies

The research was partially supported by the Fortüne junior grant, University of Tuebingen (FC: 2407-0-0) and the Deutsche Forschungsgemeinschaft (DFG) (OR: RI 682/19-1 AOBJ663994). NGS sequencing methods were performed with the support of the DFG-funded NGS Competence Center Tübingen (INST 37/1049-1). Work of PB and RK is supported by an Excellence Award for Research (PEARL) from the Fonds National de Recherche Luxembourg (to RK; FNR/P13/6682797).

Competing financial interests:

The authors declare no competing financial interests.

Supplementary material

Supplementary material is available at *Brain* online.

Reference

1. Roussigne M, Kossida S, Lavigne AC, et al. The THAP domain: a novel protein motif with similarity to the DNA-binding domain of P element transposase. Trends Biochem Sci. 2003;28(2):66-9.

2. Clouaire T, Roussigne M, Ecochard V, Mathe C, Amalric F, Girard JP. The THAP domain of THAP1 is a large C2CH module with zinc-dependent sequence-specific DNA-binding activity. Proc Natl Acad Sci U S A. 2005;102(19):6907-12.

3. Fuchs T, Gavarini S, Saunders-Pullman R, et al. Mutations in the THAP1 gene are responsible for DYT6 primary torsion dystonia. Nat Genet. 2009;41(3):286-8.

4. Cayrol C, Lacroix C, Mathe C, et al. The THAP-zinc finger protein THAP1 regulates endothelial cell proliferation through modulation of pRB/E2F cell-cycle target genes. Blood. 2007;109(2):584-94.

5. Kaiser FJ, Osmanoric A, Rakovic A, et al. The dystonia gene DYT1 is repressed by the transcription factor THAP1 (DYT6). Ann Neurol. 2010;68(4):554-9.

6. Gavarini S, Cayrol C, Fuchs T, et al. Direct interaction between causative genes of DYT1 and DYT6 primary dystonia. Ann Neurol. 2010;68(4):549-53.

7. Cheng F, Walter M, Wassouf Z, et al. Unraveling Molecular Mechanisms of THAP1 Missense Mutations in DYT6 Dystonia. J Mol Neurosci. 2020;70(7):999-1008.

8. Aguilo F, Zakirova Z, Nolan K, et al. THAP1: Role in Mouse Embryonic Stem Cell Survival and Differentiation. Stem Cell Reports. 2017;9(1):92-107.

9. Yellajoshyula D, Liang CC, Pappas SS, et al. The DYT6 Dystonia Protein THAP1 Regulates Myelination within the Oligodendrocyte Lineage. Dev Cell. 2017;42(1):52-67.e4.

10. Zakirova Z, Fanutza T, Bonet J, et al. Mutations in THAP1/DYT6 reveal that diverse dystonia genes disrupt similar neuronal pathways and functions. PLoS Genet. 2018;14(1):e1007169.

11. Frederick NM, Shah PV, Didonna A, Langley MR, Kanthasamy AG, Opal P. Loss of the dystonia gene Thap1 leads to transcriptional deficits that converge on common pathogenic pathways in dystonic syndromes. Hum Mol Genet. 2019;28(8):1343-1356.

12. Baumann H, Ott F, Weber J, et al. Linking Penetrance and Transcription in DYT-THAP1: Insights From a Human iPSC-Derived Cortical Model. Mov Disord. 2021;36(6):1381-1391.

13. Schöndorf DC, Aureli M, McAllister FE, et al. iPSC-derived neurons from GBA1-associated Parkinson's disease patients show autophagic defects and impaired calcium homeostasis. Nat Commun. 2014;5:4028.

14. Reinhardt P, Schmid B, Burbulla LF, et al. Genetic correction of a LRRK2 mutation in human iPSCs links parkinsonian neurodegeneration to ERK-dependent changes in gene expression. Cell Stem Cell. 2013;12(3):354-67.

15. Barbuti P, Santos B, Dording C, et al. Generation of two iPS cell lines (HIHDNDi001-A and HIHDNDi001-B) from a Parkinson's disease patient carrying the heterozygous p.A30P mutation in SNCA. Stem Cell Res. 2020;48:101951.

16. Barbuti P, Antony P, Santos B, et al. Using High-Content Screening to Generate Single-Cell Gene-Corrected Patient-Derived iPS Clones Reveals Excess Alpha-Synuclein with Familial Parkinson's Disease Point Mutation A30P. Cells. 2020;9(9):2065.

17. Landt SG, Marinov GK, Kundaje A, et al. ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res. 2012;22(9):1813-31.

18. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol. 2008;9(9):R137.

19. Machanick P, Bailey TL. MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics. 2011; 27(12): 1696–1697.

20. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8(11):2281-2308.

21. Paudel R, Li A, Hardy J, Bhatia KP, Houlden H, Holton J. DYT6 Dystonia: A Neuropathological Study. Neurodegener Dis. 2016;16(3-4):273-8.

22. Grundmann K, Glöckle N, Martella G, et al. Generation of a novel rodent model for DYT1 dystonia. Neurobiol Dis. 2012;47(1):61-74.

23. Manfré G, Clemensson EKH, Kyriakou EI, et al. The BACHD Rat Model of Huntington Disease Shows Specific Deficits in a Test Battery of Motor Function. Front Behav Neurosci. 2017;11:218.

24. Qian Y, Lei G, Castellanos FX, Forssberg H, Heijtz RD. Deficits in fine motor skills in a genetic animal model of ADHD. Behav Brain Funct. 2010;6:51.

25. Karlić R, Chung HR, Lasserre J, Vlahovicek K, Vingron M. Histone modification levels are predictive for gene expression. Proc Natl Acad Sci U S A. 2010;107(7):2926-31.

26. Creyghton MP, Cheng AW, Welstead GG, et al. Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A. 2010;107(50):21931-6.

27. Ruiz M, Perez-Garcia G, Ortiz-Virumbrales M, et al. Abnormalities of motor function, transcription and cerebellar structure in mouse models of THAP1 dystonia. Hum Mol Genet. 2015;24(25):7159-70.

28. Erogullari A, Hollstein R, Seibler P, et al. THAP1, the gene mutated in DYT6 dystonia, autoregulates its own expression. Biochim Biophys Acta. 2014;1839(11):1196-204.

29. Kwon HS, Kim MS, Edenberg HJ, Hur MW. Sp3 and Sp4 can repress transcription by competing with Sp1 for the core cis-elements on the human ADH5/FDH minimal promoter. J Biol Chem. 1999;274(1):20-8.

30. Wang R, Yang Y, Wang H, He Y, Li C. MiR-29c protects against inflammation and apoptosis in Parkinson's disease model in vivo and in vitro by targeting SP1. Clin Exp Pharmacol Physiol. 2020;47(3):372-382.

31. Qiu Z, Norflus F, Singh B, et al. Sp1 is up-regulated in cellular and transgenic models of Huntington disease, and its reduction is neuroprotective. J Biol Chem. 2006;281(24):16672-80.

32. Santpere G, Nieto M, Puig B, Ferrer I. Abnormal Sp1 transcription factor expression in Alzheimer disease and tauopathies. Neurosci Lett. 2006;397(1-2):30-4.

33. Citron BA, Dennis JS, Zeitlin RS, Echeverria V. Transcription factor Sp1 dysregulation in Alzheimer's disease. J Neurosci Res. 2008;86(11):2499-504.

34. Hung CY, Hsu TI, Chuang JY, Su TP, Chang WC, Hung JJ. Sp1 in Astrocyte Is Important for Neurite Outgrowth and Synaptogenesis. Mol Neurobiol. 2020;57(1):261-277.

35. O'Connor L, Gilmour J, Bonifer C. The Role of the Ubiquitously Expressed Transcription Factor Sp1 in Tissue-specific Transcriptional Regulation and in Disease. Yale J Biol Med. 2016;89(4):513-525.

36. Cai LJ, Tu L, Li T, et al. Up-regulation of microRNA-375 ameliorates the damage of dopaminergic neurons, reduces oxidative stress and inflammation in Parkinson's disease by inhibiting SP1. Aging (Albany NY). 2020;12(1):672-689.

37. Safe S, Abdelrahim M. Sp transcription factor family and its role in cancer. Eur J Cancer. 2005;41(16):2438-48.

38. Ferrante RJ, Ryu H, Kubilus JK, et al. Chemotherapy for the brain: the antitumor antibiotic mithramycin prolongs survival in a mouse model of Huntington's disease. J Neurosci. 2004;24(46):10335-42.

39. Sleiman SF, Langley BC, Basso M, et al. Mithramycin is a gene-selective Sp1 inhibitor that identifies a biological intersection between cancer and neurodegeneration. J Neurosci. 2011;31(18):6858-70.

40. Subaiea GM, Adwan LI, Ahmed AH, Stevens KE, Zawia NH. Short-term treatment with tolfenamic acid improves cognitive functions in Alzheimer's disease mice. Neurobiol Aging. 2013;34(10):2421-30.

41. Boutillier S, Lannes B, Buée L, et al. Sp3 and Sp4 Transcription Factor Levels Are Increased in Brains of Patients with Alzheimer’s Disease. Neurodegener Dis. 2007;4(6):413-23.

42. Villa C, Ghezzi L, Fenoglio C, et al. Genetics and expression analysis of the specificity protein 4 gene (SP4) in patients with Alzheimer's disease and frontotemporal lobar degeneration. J Alzheimers Dis. 2012;31(3):537-42.

43. Zhou X, Qyang Y, Kelsoe JR, Masliah E, Geyer MA. Impaired postnatal development of hippocampal dentate gyrus in Sp4 null mutant mice. Genes Brain Behav. 2007;6(3):269-76.

44. Zhou X, Long JM, Geyer MA, et al. Reduced expression of the Sp4 gene in mice causes deficits in sensorimotor gating and memory associated with hippocampal vacuolization. Mol Psychiatry. 2005;10(4):393-406.

45. Tinazzi M, Rosso T, Fiaschi A. Role of the somatosensory system in primary dystonia. Mov Disord. 2003;18(6):605-22.

46. Conte A, Belvisi D, De Bartolo MI, et al. Abnormal sensory gating in patients with different types of focal dystonias. Mov Disord. 2018;33(12):1910-1917.

47. Goodchild RE, Grundmann K, Pisani A. New genetic insights highlight 'old' ideas on motor dysfunction in dystonia. Trends Neurosci. 2013; 36(12): 717-725.

48. Downs AM, Roman KM, Campbell SA, Pisani A, Hess EJ, Bonsi P. The neurobiological basis for novel experimental therapeutics in dystonia. Neurobiol Dis. 2019;130:104526.

49. Scarduzio M, Zimmerman CN, Jaunarajs KL, Wang Q, Standaert DG, McMahon LL. Strength of cholinergic tone dictates the polarity of dopamine D2 receptor modulation of striatal cholinergic interneuron excitability in DYT1 dystonia. Exp Neurol. 2017;295:162–175.

50. Bonsi P, Ponterio G, Vanni V, et al. RGS9-2 rescues dopamine D2 receptor levels and signaling in DYT1 dystonia mouse models. EMBO Mol Med. 2019;11(1):e9283.

51. Jaunarajs KLE, Scarduzio M, Ehrlich ME, McMahon LL, Standaert DG. Diverse Mechanisms Lead to Common Dysfunction of Striatal Cholinergic Interneurons in Distinct Genetic Mouse Models of Dystonia. J Neurosci. 2019;39(36):7195-7205.

52. Carbon M, Niethammer M, Peng S, et al. Abnormal striatal and thalamic dopamine neurotransmission: Genotype-related features of dystonia. Neurology. 2009;72(24):2097-103.

53. Saunders-Pullman R, Raymond D, Senthil G, et al. Narrowing the DYT6 dystonia region and evidence for locus heterogeneity in the Amish-Mennonites. Am J Med Genet A. 2007;143A(18):2098-2105.

54. Kuyper DJ, Parra V, Aerts S, Okun MS, Kluger BM. Nonmotor manifestations of dystonia: a systematic review. Mov Disord. 2011;26(7):1206-17.

55. Bova A, Kernodle K, Mulligan K, Leventhal D. Automated Rat Single-Pellet Reaching with 3-Dimensional Reconstruction of Paw and Digit Trajectories. J Vis Exp. 2019;(149):10.3791/59979.

Figure Legends

Figure 1. RNA-seq and ChIP-seq revealed THAP1 targets only a minority of DEGs cause by its mutation. (A) Western blot analysis showed that protein levels of THAP1 in two THAP1$^{+/-}$ SH-SY5Y cell clones (THAP1$^{+/-}$ C1 and C2) were decreased significantly. (****: p<0.0001, Comparison by unpaired two-tailed *t* test) (B) Volcano plot shows the differentially expressed genes (DEGs) in THAP$^{+/-}$ cell lines identified by RNA-seq analysis. The green dots indicate the down-regulated genes and red dots show the up-regulated genes. (C) Gene-ontology analysis shows the only significantly affected pathways of down-regulated genes is neurotransmitter secretion (green color). Red color showed five top pathways annotated by DAVID bioinformatics tool of up-regulated genes. (D) Heatmap of two replicates of THAP1 ChIP-seq showed THAP1 binding regions are enriched around transcription start sites (TSS). (E) ChIP-seq binding profile analysis showed that THAP1 binding sites are enriched near TSS and overlapping to H3K27ac and H3K4me3 ChIP-seq peaks. (F) MEME motif discovery analysis found two motifs that THAP1 potentially binds to. (G) THAP1 targets to only 14 down-regulated and 7 up-regulated DEGs found in THAP1$^{+/-}$ cell lines.

Figure 2. **Epigenetic analysis linked SP1 with the alterations of H3K27ac and H3K4me3 modifications in THAP1$^{+/-}$ SH-SY5Y cell lines.** (A) ChIP-seq binding profile analysis of H3K27ac (upper) and H3K4me3 (lower) modifications in control and THAP1$^{+/-}$ cell lines. Both H3K27ac and H3K4m3 ChIP-seq showed increased binding profile in THAP1$^{+/-}$ cell lines (THAP1 C1 and C2). (B) Volcano plots showed that 1921 regions (diff_H3K27ac) changed the H3K27ac ChIP-seq signals (left panel, red and purple dots), and 2762 regions (diff_H3K4me3) changed the H3K4me3 ChIP-seq signals (right panel, red and purple dots), respectively. Top fold changed regions (purple dots) (Top n=606 diff-H3K27ac regions and top n=476 diff-H3K4me3 regions) were selected to motif discovery analysis. (C) Western blot analysis showed that the protein level of H3K27ac and H3K4me3 did not change significantly in THAP1$^{+/-}$ cell lines. (D) Top 3 motifs were discovered by using top diff-H3K27ac and diff-H3K4me3 regions (left: top three motifs of diff-H3K27ac; right: top 3 motifs of diff-H3K4me3). (E) MEME motif discovery revealed that the SP1 binding motif significantly matched to motifs 2 of diff-H3K27ac and motif 2 and 3 of diff-H3K4me3 (below each motif indicated the results statistical analysis).

Figure 3. THAP1 and its target SP1 are mainly responsible for dysregulation of DEGs in THAP1$^{+/-}$ SH-SY5Y cell lines. (A) Western blot and quantification analysis showed that the protein level of SP1 is increased significantly in THAP1$^{+/-}$ cell lines (****: p<0.0001, Comparison by unpaired two-tailed *t* test). (B) ChIP-seq analysis revealed that THAP1 binds to *SP1* promoter and a potential enhancers in intron 3 to regulate its expression (lane 1). Both H3K27ac (lane 2 and 3) and H3K4me3 (lane 4 and 5) ChIP-seq indicate that the promoter and enhancer activities are increased after heterozygous knock-out of THAP1 (THAP1$^{+/-}$). RNA-seq (land 6 and 7) demonstrates that the RNA level of SP1 is increased in THAP1$^{+/-}$ cell lines. (C) Heatmaps of two biological replicates of SP1 ChIP-seq show that SP1 is mainly enriched around TSS. (D) ChIP-seq binding profile analysis revealed that SP1 enriched on TSS sites and

overlapped to the peaks of H3K27ac and H3K4me3. (E) A typical SP1 binding motif was discovered by MEME motif discovery analysis. (F) Venn diagram showed that THAP1 and SP1 binding regions are overlapped with a large amount of diff-H3K27ac and diff-H3K4me3 regions, while SP1 has more overlapped regions compared to THAP1. (G) Venn diagram showed that almost half of DEGs (119 / 279) found in $THAP1^{+/-}$ SH-SY5Y cell lines are the target genes of SP1. (H) SP1 ChIP-qPCR analysis showed that heterozygous knock-out of THAP1 ($THAP1^{+/-}$) leads to increased SP1 binding activities on some genes (EMILIN1, EPAS, NFATC4, and SYT7). (I) Western blot showed that no SP1 protein was detected in $THAP1^{+/-}$ / $SP1^{-/-}$ SH-SY5Y cells. (J) Venn diagram showed overlaps of DEGs in $THAP1^{+/-}$ SH-SY5Y cells and DEGs in $THAP1^{+/-}$ / $SP1^{-/-}$ SH-SY5Y cells, more than half of DEGs found in $THAP1^{+/-}$ SH-SY5Y cells (66 / 138 down-regulated genes and 78 / 141 up-regulated genes) were not in the DEGs list of $THAP1^{+/-}$ / $SP1^{-/-}$ SH-SY5Y cells. (K) Heatmap showed gene expression changes in $THAP1^{+/-}$ and $THAP1^{+/-}$ / $SP1^{-/-}$ SH-SY5Y cell lines. (Red line below heatmap indicates some genes which were normalized or reversed their expression change in $THAP1^{+/-}$ / $SP1^{-/-}$ SH-SY5Y cells compared to $THAP1^{+/-}$ cells. (*: $p<0.05$, **: $p<0.01$, Comparison by unpaired two-tailed *t* test)

Figure 4. Gene expression analysis of THAP1 patients' iPSCs derived mDA neurons and $THAP1^{+/-}$ rat striatum. (A) Volcano plot showed the dysregulated genes found in mDA neurons of THAP1 patients. In total, 2288 down-regulated genes and 3092 up-regulated genes were observed by RNA-seq analysis (FDR<0.01, $\log_2$(Fold change) >1). (B) Volcano plot showed the dysregulated genes observed in $THAP1^{+/-}$ rat striatum tissues. In total, 1254 down-regulated genes and 1153 up-regulated genes were observed by RNA-seq analysis (FDR<0.01, fold change > 20%). (C) Many common dysregulated genes were observed in THAP1 iPSCs derived mDA neurons and in $Thap1^{+/-}$ rat striatum. (D) Gene-ontology analysis showed the top affected pathways of common dysregulated genes of THAP1 iPSCs derived mDA neuron and $THAP1^{+/-}$ rat striatum. The "chemical synaptic transmission" and "regulation of synaptic transmission, glutamatergic" ("Synaptic transmission" in short for both pathways) are top two significantly affected pathways of common down-regulated genes. Other pathways related to pathogenesis of primary dystonia, like "nervous system development" and "locomotory behavior" pathways, were also identified. (E) Heatmap plot shows the expression change of 61 common dysregulated genes in $THAP1^{+/-}$ iPSCs derived mDA neurons (HiPSC-mDA) and in $THAP1^{+/-}$ rat striatum (RStriatum) associated to pathways of "Synaptic transmission", "Nervous system development", and "Locomotory behavior". (F) The relative expression level of wild-type THAP1 allele in different THAP1 mutant cells/ tissues. The wild-type THAP1 allele was calculated by multiplying the amount of total THAP1 (copies per million reads) with the ratio of wild-type allele (divide the reads number of wild-type THAP1 to the total reads of THAP1 which located on the position of THAP1 mutation, wt/total%). (*: $p<0.05$, ***: $p<0.001$, Comparison by unpaired two-tailed *t* test)

Figure 5. Behavioral studies to characterize $THAP1^{+/-}$ rats. (A) Tail suspension test showed two of 14 rats presented hind limb clasping ($THAP1^{+/-}$ No. 2) or hind limb clasping like behavior

(THAP1$^{+/-}$ No. 1) compare to wild-type rats. (B) Curve plot showed the time of hind limb clasping behavior occurrence of THAP1$^{+/-}$ rats. (C) Fine motor skill test showed that the THAP1$^{+/-}$ rats presented lower scores of both total success rate and success rate on first attempt compared to wild-type control rats (left and middle). The total attempt times to reach the same amount of pellets (20 pellets) are more in THAP1$^{+/-}$ rat compared to wild-type rats (right) (n=8 rats in each group). (D) Phenomaster behavior analysis showed slightly decreased dark phase activities of THAP$^{+/-}$ rats compared to wild-type rats at 9-month old, statistical analysis revealed significant differences of overall distance (left) and Z-scale distance (right) at 9-month old. (n=10 rats in each group) (E) THAP1$^{+/-}$ rats present increased light phase activities compared to wild-type rats. Both overall distance (left) and Z-scale distance (right) activities are increased in THAP1$^{+/-}$ rat compared to wild-type control rat at 9-month and 12-month old. (n=10 rats in each group)

Figure 6. Electrophysiological characterization of THAP1$^{+/-}$ striatal cholinergic interneurons (ChIs). (A) Representative whole-cell recordings of wild-type and THAP1$^{+/-}$ ChIs showed similar voltage responses to hyperpolarizing current steps. (B) ChIs membrane properties analysis showed that no significant differences were observed between different genotypes (Wild-type n=25, and THAP1$^{+/-}$ n=34, $p > 0.05$). (C) Summary plot of the effects of the muscarinic agonist, oxotremorine (10 µM) (Wild-type n= 6, THAP1$^{+/-}$ n=4) and β-adrenergic agonist, isoproterenol (30 µM) (Wild-type n= 6, THAP1$^{+/-}$ n=7) on the firing frequency of ChIs from Wild-type and THAP1$^{+/-}$ rats. No significant differences were observed between different genotypes ($p>0.05$). (D) Representative patch-clamp recordings of the response to amphetamine of spontaneously firing wild-type and THAP1$^{+/-}$ ChIs. Bath-application of 3 µM amphetamine does not alter the firing frequency of the recorded wild-type ChI (lane 1), whereas 30 µM amphetamine increases the frequency of spontaneous action potentials (lane 2). In a THAP1$^{+/-}$ ChI, 3 µM amphetamine induces a net decrease of the firing frequency (lane 3), whereas 30 µM amphetamine induces an increase of firing frequency similar to the response observed in the Wild-type ChI (lane 4). (E) Summary plot of the dose-dependent effects of amphetamine on the firing frequency of ChIs from Wild-type and THAP1$^{+/-}$ rats. Statistical analysis showed a significant difference of the effect of 3 µM amphetamine between Wild-type and THAP1$^{+/-}$ ChIs (Wild-type n=5, THAP1$^{+/-}$ n=7) ($p<0.05$, Comparison by unpaired *t* test). The inhibitory effect of amphetamine in THAP1$^{+/-}$ ChIs was prevented by pre-incubation with a dopamine D2 receptor antagonist sulpiride (3 µM) (Wild-type n=4, THAP1$^{+/-}$ n=5). The excitatory effect of 30 µM amphetamine was not significantly different between Wild-type and THAP1$^{+/-}$ ChIs (Wild-type n=5, THAP1$^{+/-}$ n=5). (F) Plots reporting the effect of the DRD2 agonist quinpirole (10 µM) on the firing frequency of ChIs from Wild-type and THAP1$^{+/-}$ rats. In wild-type neurons bath-application of quinpirole decreases the firing frequency ($p<0.05$, Comparison by paired *t* test) (Wild-type n= 11), whereas in THAP1$^{+/-}$ ChIs the DRD2 agonist (quinpirole) increases the firing frequency ($p<0.05$, Comparison by paired *t* test) (THAP1$^{+/-}$ n= 8). (aCSF: artificial cerebrospinal fluid)

Figure 7. Cell type dependent expression changes of SP1 family members caused by THAP1 mutations. (A) Heatmap plot showed the expression of SP1, SP2, and SP4 in THAP1 patients' frontal cortex (HFrontCortex), THAP1 patients' iPSCs derived midbrain dopamine neurons (HiPSC-mDA), THAP1$^{+/-}$ rat striatum (RStriatum), and SH-SY5Y cells. SP1 and SP2 are predominately up-regulated in all these cells / tissues except in HFrontCortex, while SP4 is predominately down-regulated in all these cell types / tissues except in SH-SY5Y cells. (B) ChIP-seq signal coverage analysis showed that THAP1 binds to the promoter of SP1, SP2, and SP4 in both SH-SY5Y and K562 cells. (C) The binding motifs of SP1, SP2, and SP4 are very similar. (D) Heatmap plot presented expression changes of 31 overlapping genes of SP1 targets and common dysregulated gene in HiPSC-mDA and RStriatum related dystonic syndromes. The expression fold-changes of these genes are stronger in iPSC derived mDA neurons compared to other cells / tissues. Some genes in human THAP1 frontal cortex (HFrontCortex) were oppositely dysregulated compared to THAP1 HiPSC-mDA or THAP1 rat striatum, like GRM3, EFNA5, FUT9, GNAO1, and NEGR1.

Table 1 Statistical results of discovered top diff-H3K27ac and diff-H3K4me3 motifs and matched transcription factors

Regions	Top motifs	E values	Matched sites (n=total)	Matched transcription factors (TF)
Diff - H3K27ac	Motif 1	5.40×10^{-47}	94 (606)	ZNF384
	Motif 2	1.10×10^{-35}	132 (606)	SP1, ZNF263, Zfx
	Motif 3	4.50×10^{-26}	27 (606)	NA
Diff - H3K4me3	Motif 1	1.90×10^{-43}	53 (476)	ZNF384, Mtf1
	Motif 2	5.50×10^{-30}	161 (476)	SP2, SP1, KLF5, Zfp281
	Motif 3	7.80×10^{-8}	69 (476)	SP1, SP2, Zfp281, Egr1, EGR3, SP3, KLF5

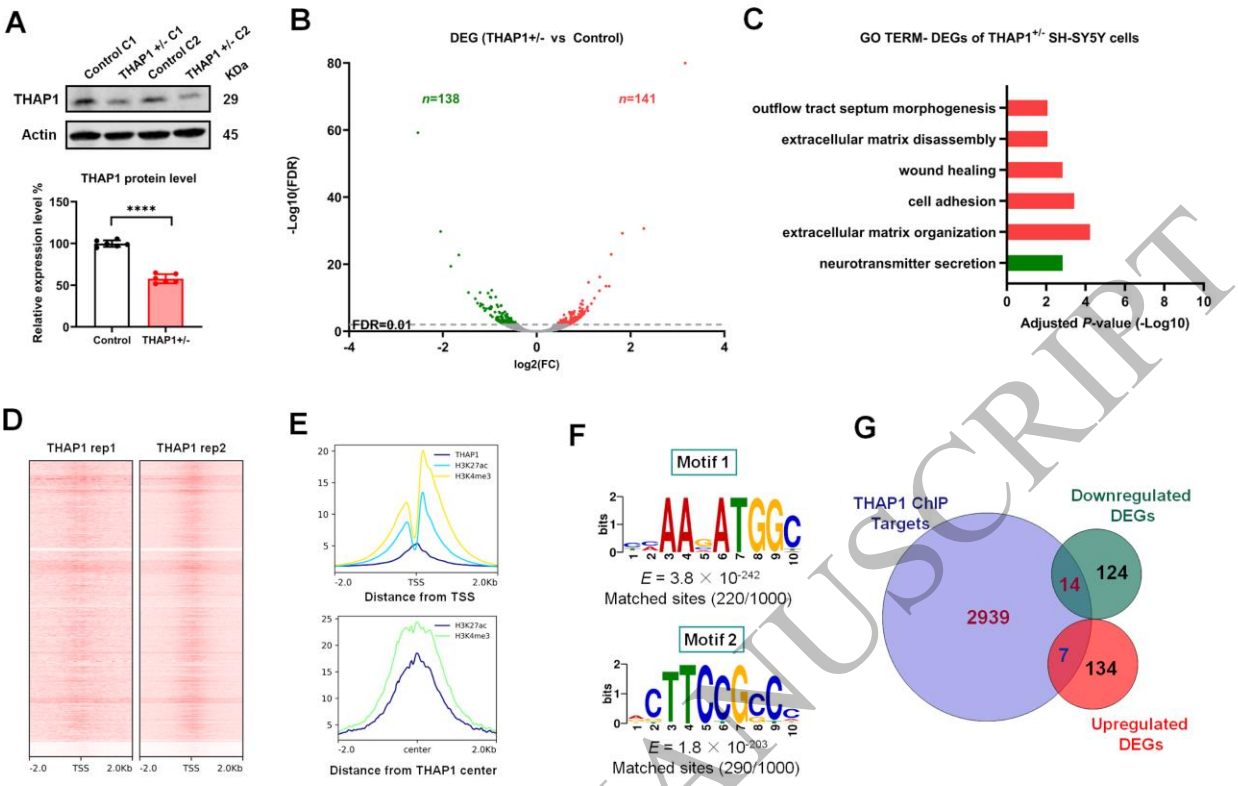

Figure 1

165x112 mm (.17 x DPI)

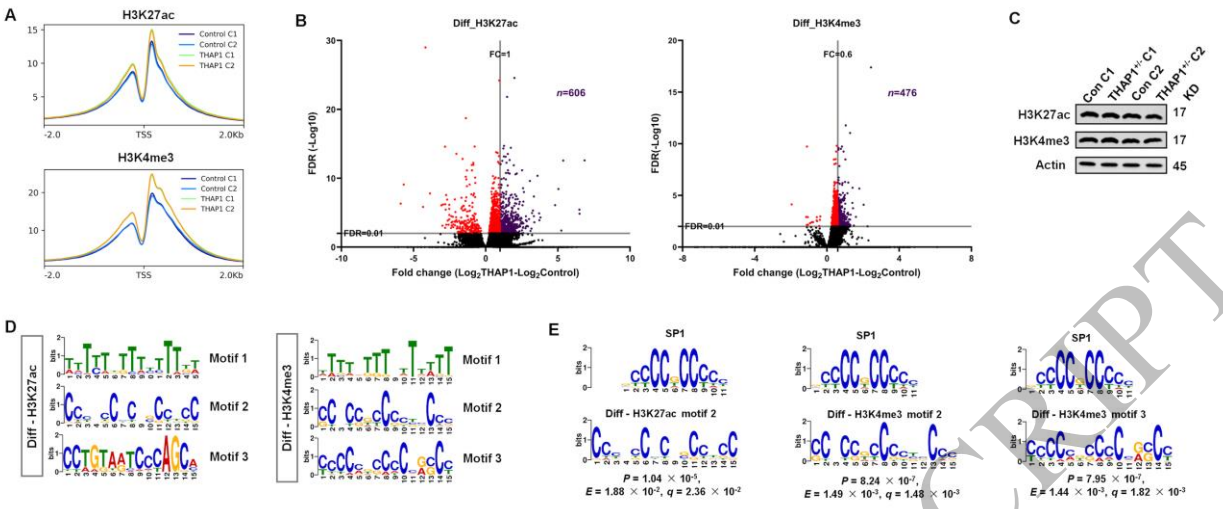

Figure 2
165x70 mm (.17 x DPI)

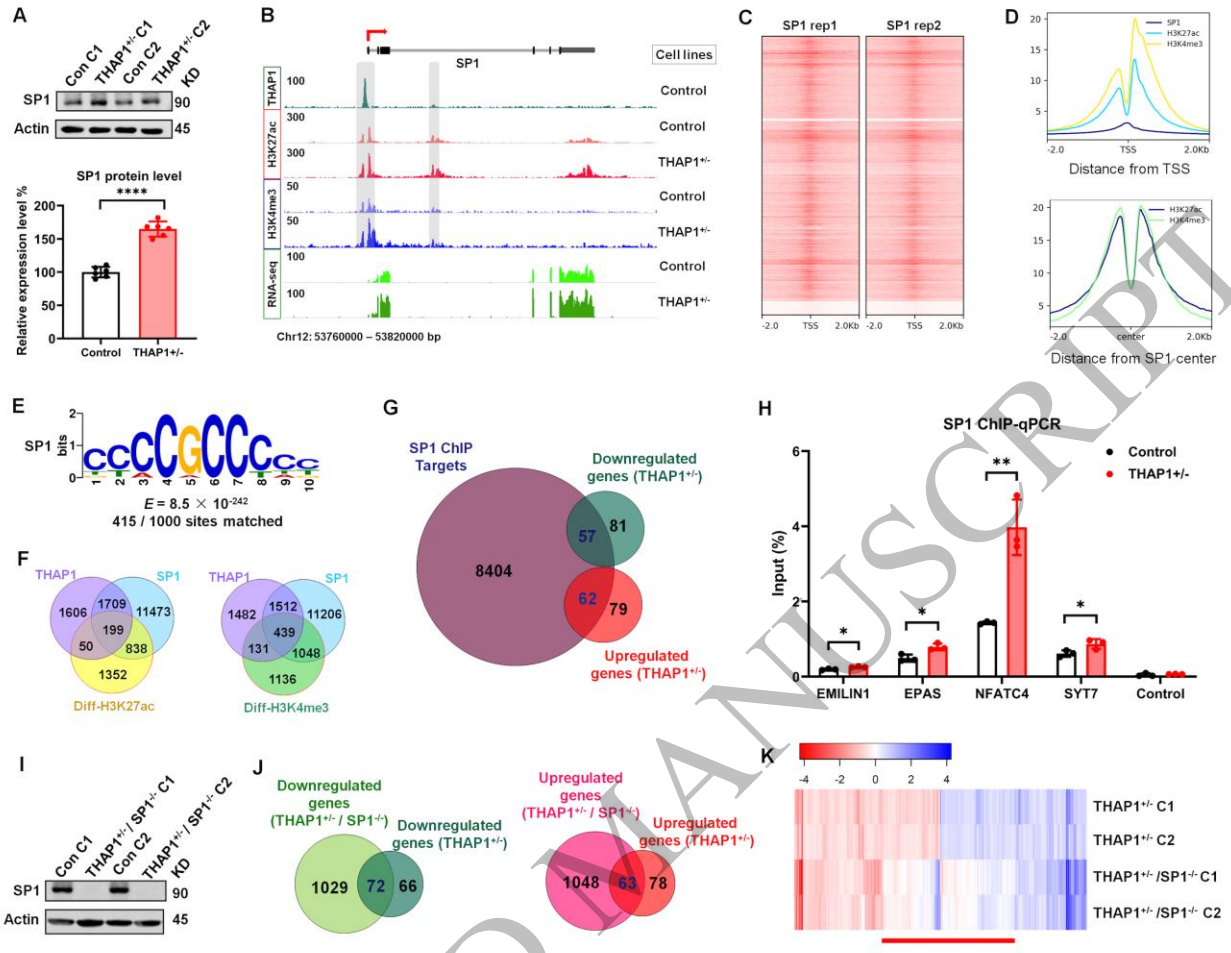

Figure 3
165x130 mm (.17 x DPI)

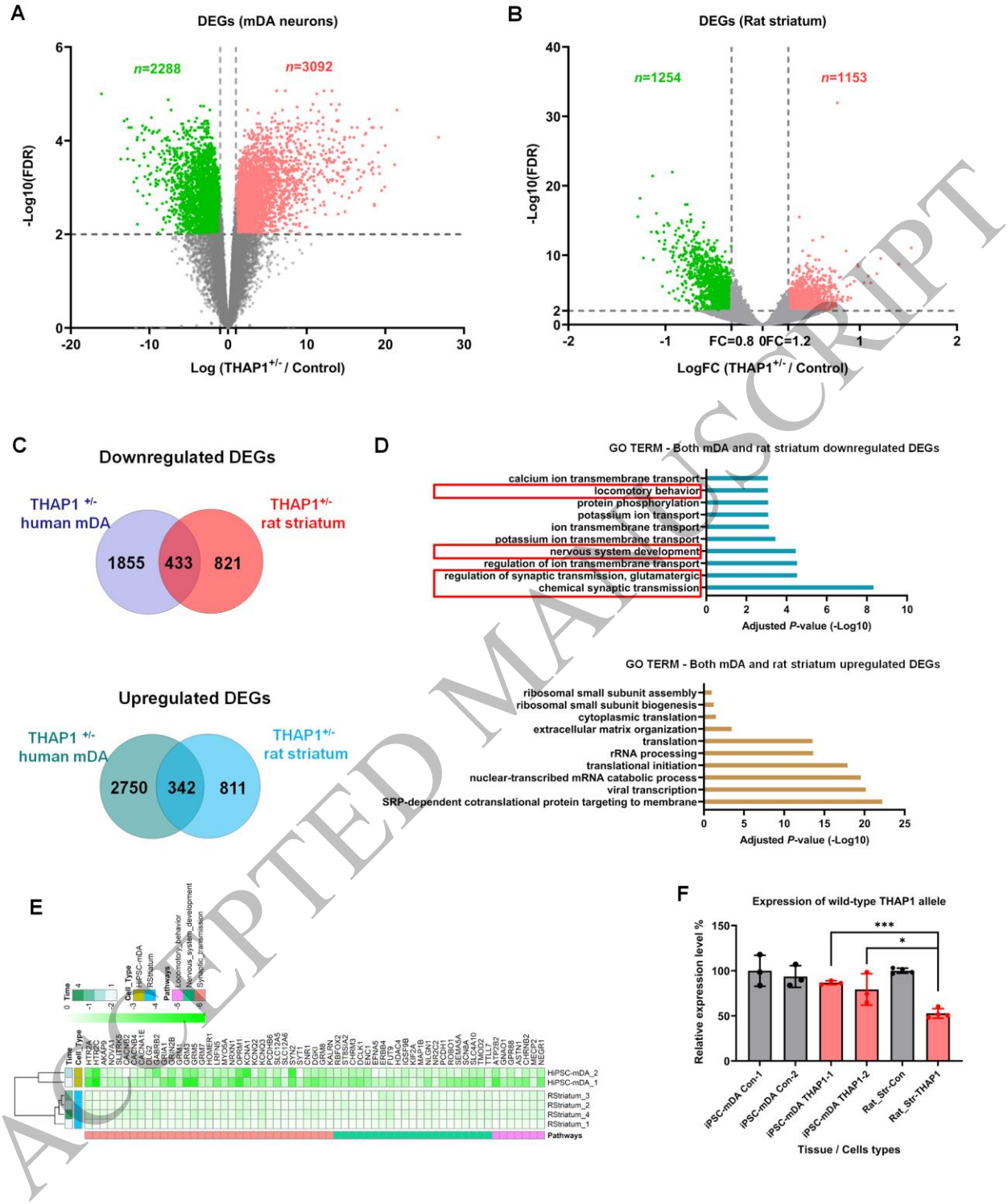

Figure 4
165x199 mm (.17 x DPI)

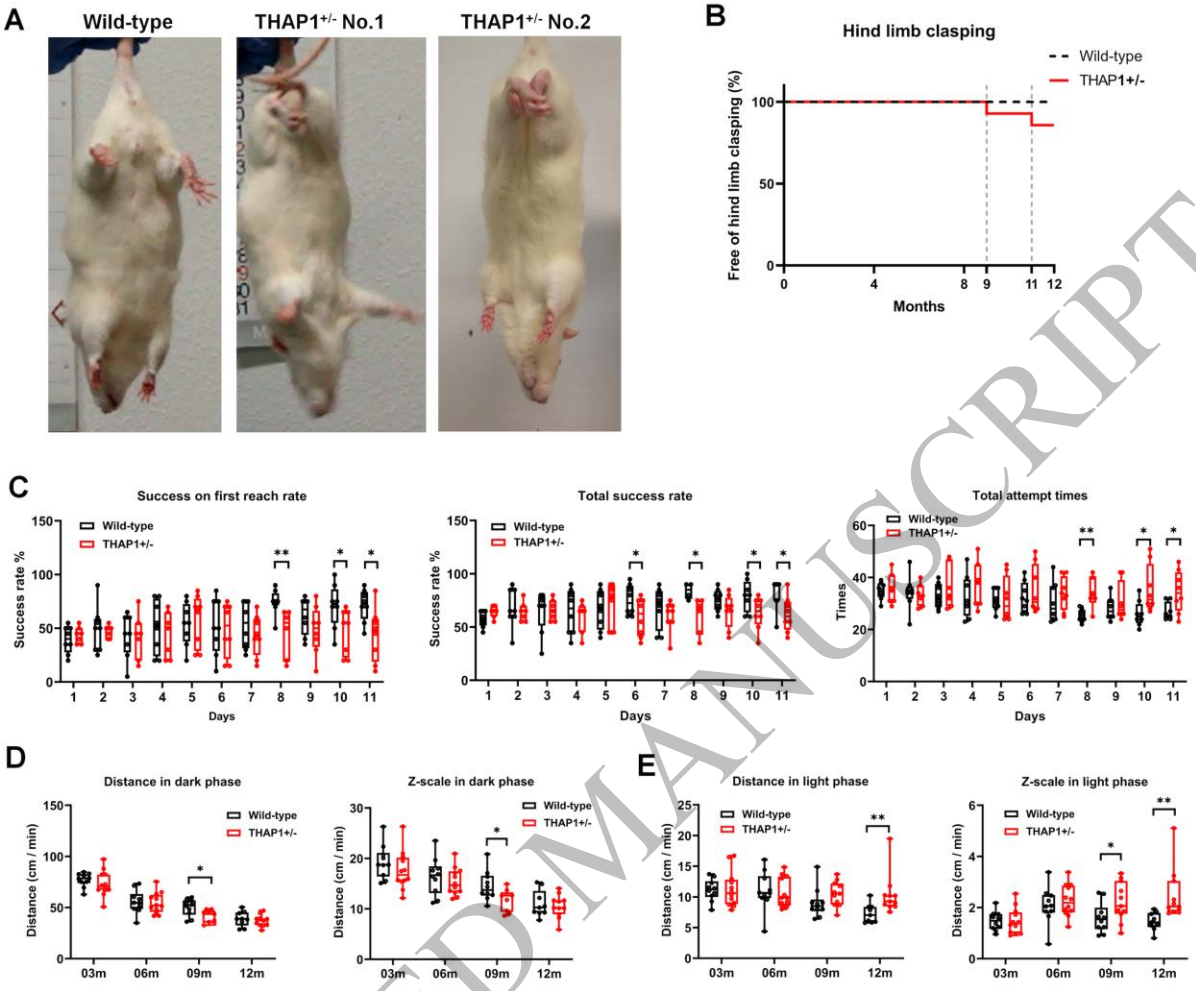

Figure 5
165x138 mm (.17 x DPI)

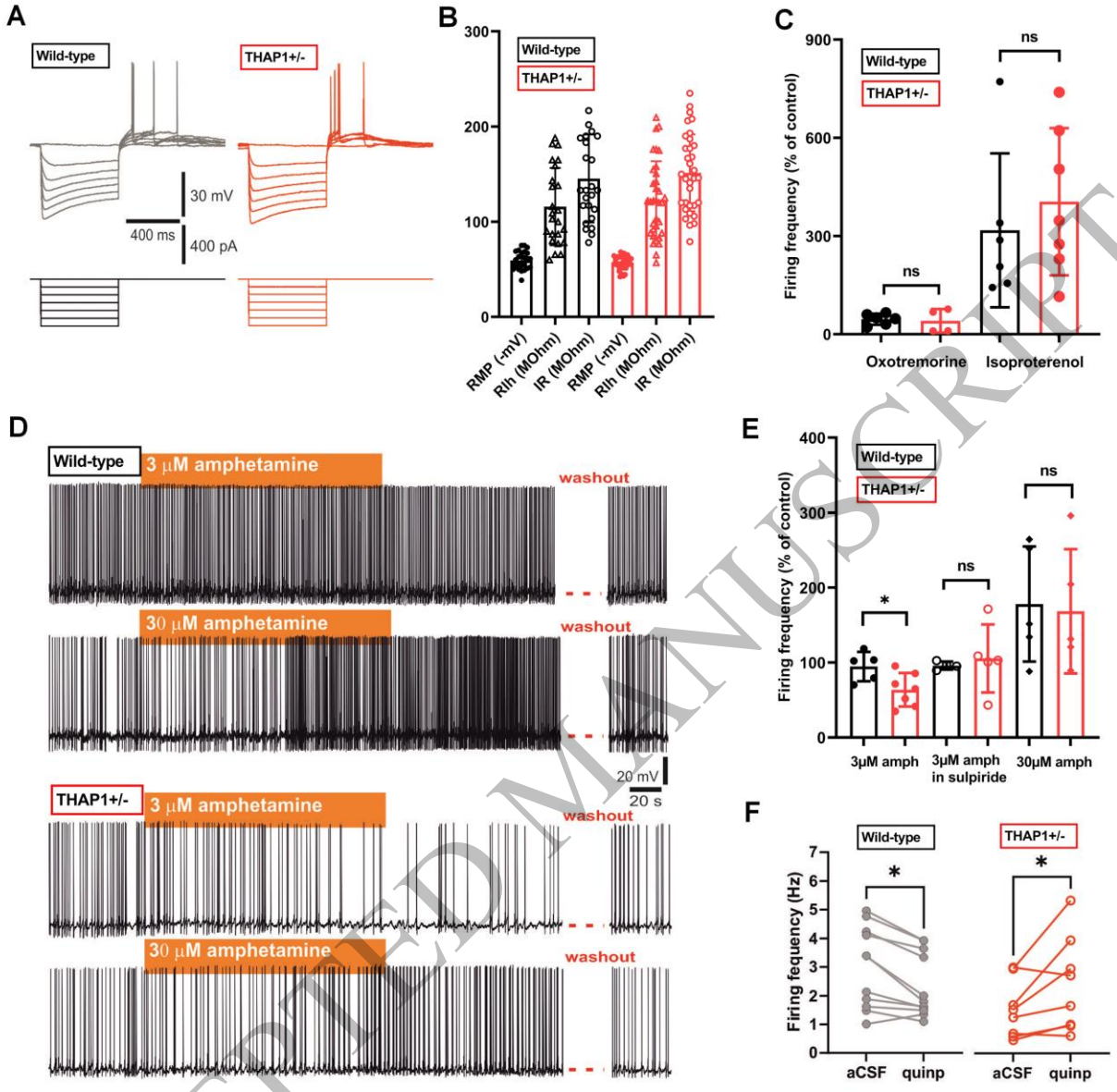

Figure 6
165x162 mm (.17 x DPI)

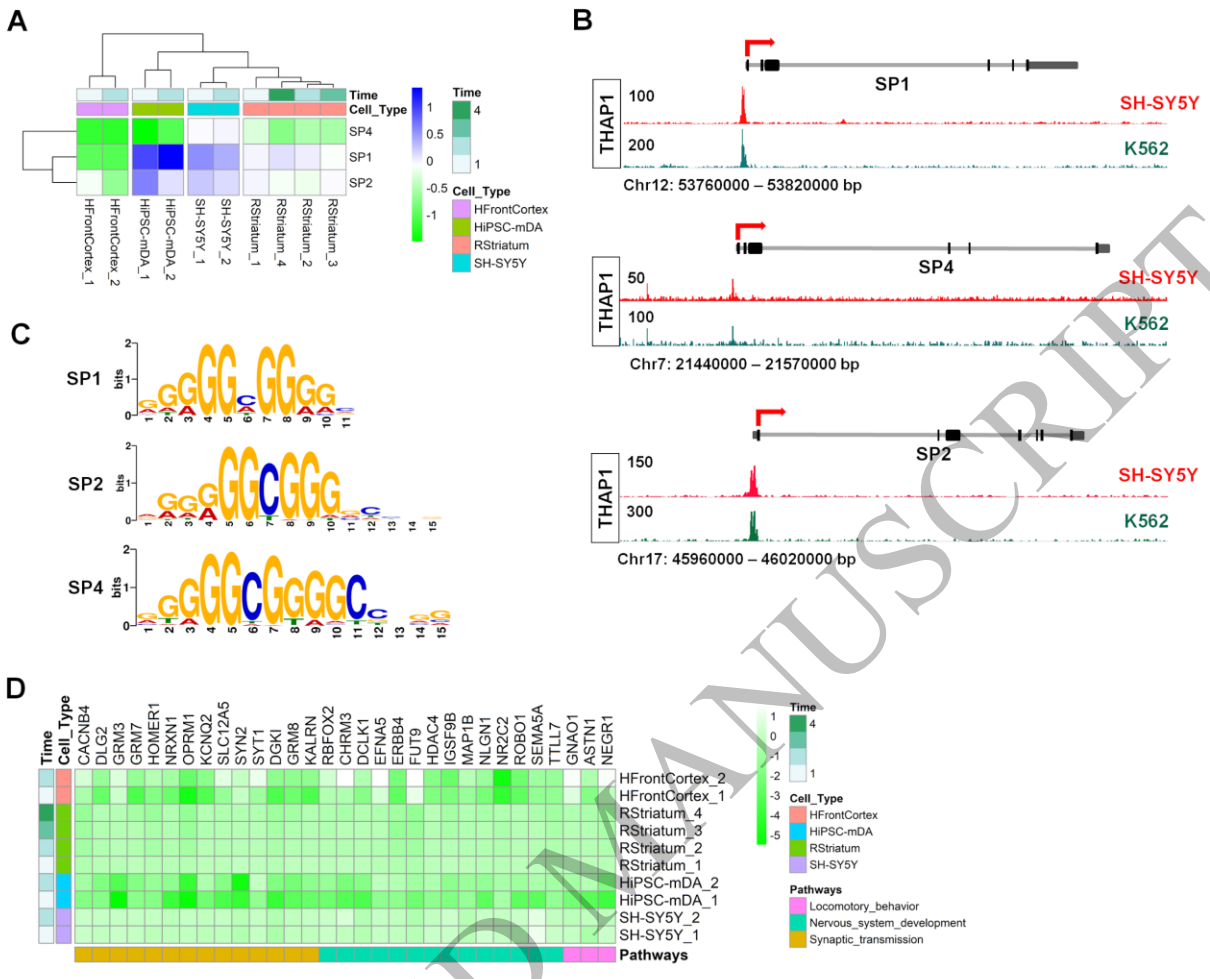

Figure 7

165x133 mm (.17 x DPI)

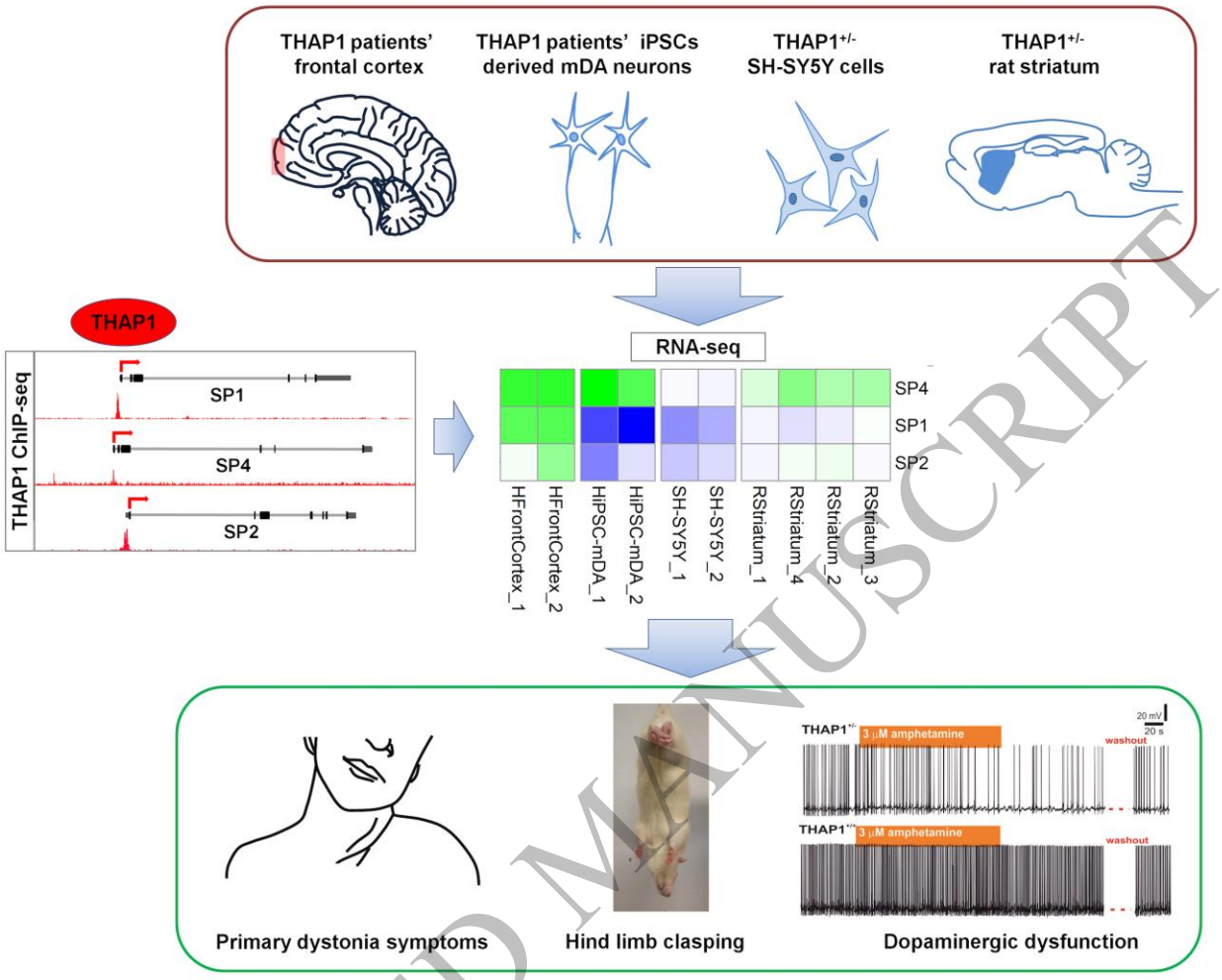

Figure 8
165x134 mm (.17 x DPI)

Using epigenetic and transcriptomic approaches combined with multiple model systems, Cheng *et al.* show that mutations in *THAP1* give rise to DYT6 dystonia via dysregulation of genes within the SP1 family. The latter could serve as therapeutic targets for a range of neurological diseases.