ABSTRACT: **Background:** The THAP1 gene encodes a transcription factor, and pathogenic variants cause a form of autosomal dominant, isolated dystonia (DYT-THAP1) with reduced penetrance. Factors underlying both reduced penetrance and the disease mechanism of DYT-THAP1 are largely unknown.

**Methods:** We performed transcriptome analysis on 29 cortical neuronal precursors derived from human-induced pluripotent stem cell lines generated from manifesting and nonmanifesting THAP1 mutation carriers and control individuals.

**Results:** Whole transcriptome analysis showed a penetrance-linked signature with expressional changes more pronounced in the group of manifesting (MMCs) than in nonmanifesting mutation carriers (NMCs) when compared to controls. A direct comparison of the transcriptomes in MMCs versus NMCs showed significant upregulation of the DRD4 gene in MMCs. A gene set enrichment analysis demonstrated alterations in various neurotransmitter release cycle pathways, extracellular matrix organization, and deoxyribonucleic acid methylation between MMCs and NMCs. When specifically considering transcription factors, the expression of YY1 and SIX2 differed in MMCs versus NMCs. Further, THAP1 was upregulated in the group of MMCs.

**Conclusions:** To our knowledge, this is the first report systematically analyzing reduced penetrance in DYT-THAP1 in a human model using transcriptomes. Our findings indicate that transcriptional alterations during cortical development influence DYT-THAP1 pathogenesis and penetrance. We reinforce previously linked pathways including dopamine and eukaryotic translation initiation factor 2 alpha signaling in the pathogenesis of dystonia including DYT-THAP1 and suggest extracellular matrix organization and deoxyribonucleic acid methylation as mediators of disease protection.

© 2021 The Authors. Movement Disorders published by Wiley Periodicals LLC on behalf of International Parkinson and Movement Disorder Society

**Key Words:** reduced penetrance; dystonia; DYT-THAP1; induced pluripotent stem cell (iPSC); whole transcriptome analysis
Pathogenic variants in \textit{THAP1} have been identified as a cause of dystonia with remarkable intrafamilial clinical variability, ranging from severe generalized dystonia to the complete absence of signs and symptoms. In fact, around 50\% of all \textit{THAP1} mutation carriers do not manifest dystonia. The molecular basis of reduced penetrance is currently unsolved. Elucidating mechanisms underlying reduced penetrance has a high imperative as they may explain endogenous disease protection.

Dystonia is widely considered as a network disorder comprising defects in neural inhibitory circuits, sensorimotor integration, and maladaptive plasticity. On the molecular level, several neuronal subtypes have been implicated in DYT-THAP1 disease pathogenesis, including cerebellar Purkinje cells, GABAergic and dopaminergic neurons, and oligodendrocytes; however, the exact cause remains elusive. Neuropathological assessment of brains from DYT-THAP1 patients did not show any anatomically discrete abnormal regions, particularly no evidence of neurodegeneration, which is consistent with findings of patients with other forms of isolated dystonia. Despite the lack of gross neuropathological changes, neuroimaging of DYT-THAP1 patients showed that cortico-striato-pallido-thalamic-cortical and cerebello-thalamo-cortical brain circuits are specifically altered in manifesting mutation carriers (MMCs) compared to nonmanifesting mutation carriers (NMCs), and abnormalities in these circuits may also correlate with clinical severity and penetrance.

\textit{THAP1} has been demonstrated to autoregulate its own expression via binding to its own promoter, and some pathogenic variants affecting the THAP domain (eg, Arg13His) are associated with decreased DNA binding and elevated \textit{THAP1} expression. Animal models provide evidence that \textit{THAP1} expression is highest during prenatal brain development and declines with age in different brain regions, indicating an important role of \textit{THAP1} during development.

The lack of overt histopathological changes in the brains of dystonia patients creates a conceptual challenge with regard to which cell type is relevant for disease pathogenesis at which time point. We assumed that \textit{THAP1}, as a transcription factor, impacts on gene expression also in cortical neurons during early brain development. Under this assumption, an endogenous iPSC-derived neuron model using cortical neuronal precursor cells was utilized to study the expression consequences of pathogenic \textit{THAP1} mutations in MMCs and NMCs as well as in controls by global transcriptome analysis.

\section*{Patients and Methods}

\subsection*{Patients}

Written informed consent was obtained from all individuals, and the study was approved by the Ethics Committee at the University of Lübeck (approval number 16-039). All patients were diagnosed according to published clinical criteria for dystonia by movement disorder specialists (A.M., S.Z., C.K., and V.S.K.). Detailed phenotypic information on members of Family A and Family B was previously published. An overview of the clinical, genetic, and demographic data of \textit{THAP1} mutation carriers and controls is provided in Table 1, and pedigrees are shown in Figure 1A.

\section*{Reprogramming and Differentiation of iPSCs Into Cortical Neurons}

Reprogramming of patient fibroblasts into iPSCs was performed using nonintegrative Sendai virus (CytoTune-iPS 2.0 Reprogramming Kit, Thermo Fisher, Waltham, MA), and iPSC lines were comprehensively characterized as described. Control iPSC lines were derived from the StemBANCC consortium using the same method and distributed via EBiSC (https://cells.ebisc.org). Neural differentiation into cortical neurons was achieved by dual SMAD inhibition (selective inhibition of SMAD family members), replying of 2 rosette stages, and 1 dissociation step using an established protocol with minor modifications (Fig. 1B). The identifier of clones with information about differentiation batch and individual has been indicated in Supplementary Table S1.

\section*{Immunocytochemistry}

To confirm final differentiation into cortical neurons over 3 months, ~90-day-old cells were fixed for characterization of neuronal/cortical properties on glass coverslips with 4\% formaldehyde for 15 minutes at room temperature and rinsed with 1x PBS (phosphate-buffered saline). Coverslips were blocked with 10\% (v/v) normal goat serum and 0.1\% (v/v) Triton X-100 solved in 1x PBS for 45 minutes and then rinsed with 1x PBS. Primary antibodies BRN2 (1:500; Santa Cruz, CA), CTIP2 (1:500; Abcam, Cambridge, UK), MAP2 (1:200; Millipore, Danvers, MA), TUJ1 (1:1000; Covance, Princeton, NJ), and vGLUT1 (1:500; Abcam, Cambridge, MA) were applied overnight at 4°C. The cells were incubated with the respective species-specific fluorescent secondary antibodies (1:5000; Thermo Fisher), washed with 1x PBS, and mounted with a DAPI (4',6-diamidino-2-phenylindole)-containing medium. The cells were visualized using an LSM710 (Zeiss, Oberkochen, Germany) confocal laser-scanning microscope.

\section*{RNA Extraction}

RNA was extracted on day 44 of cortical differentiation for the transcriptome analysis. Cell pellets of 44-day-old iPSC-derived neurons were harvested using a spatula and rinsed with 1x PBS. After centrifugation at 800g for 5 minutes, the supernatant was aspirated,
and cell pellets were frozen at −80°C. Total RNA for bulk RNA sequencing was extracted using the RNeasy Mini Plus Kit (Qiagen, Venlo, The Netherlands). RNA quantity and quality were assessed by measuring the absorbance at 260/280 and 260/230 nm using the NanoDrop1000 spectrometer and RNA integrity number (RIN) using an Agilent Bioanalyzer. All samples exhibited RIN > 9.5 (Supplementary Table S1).

**Whole Transcriptome Analysis**

Library preparation and paired-end sequencing were performed at Novogene (Beijing, China) with 59–95 Mio reads per library. The company’s workflow included mRNA enrichment by oligo(dT) beads, rRNA removal using the Ribo-Zero kit, and random RNA fragmentation using fragmentation buffer. The double-stranded cDNA library was synthesized using random hexamer primer, and second strand synthesis was initiated using deoxynucleoside triphosphates, RNase H, and DNA polymerase I. On end repair, A-tailing, and adapter ligation, the double-stranded cDNA library was subjected to size selection (150 bp) and PCR (polymerase chain reaction) enrichment. Subsequently, the samples were sequenced by “sequencing by synthesis” on an Illumina machine.

Fastq reads were pseudo-aligned to the human transcriptome in Ensembl version 97 using Salmon\textsuperscript{20} (v0.14.1). Differential gene expression analysis was performed via the R library sleuth,\textsuperscript{21} for which read counts were aggregated to Ensembl Gene IDs. The significance and effect sizes of differential gene regulation were calculated via a Wald test using the family and/or sex as covariates, if applicable. Pathway enrichment analyses were performed based on the effect size of the differential gene expression analysis using the generally applicable gene set enrichment analysis (GSEA), GAGE,\textsuperscript{22} which determines whether a set of genes is systematically up- or downregulated as a whole. For gene set definitions, we used the Molecular Signatures Database (MSigDB)\textsuperscript{23} in v7.2. Gene sets with fewer than three or those with more than 500 members were discarded for statistical robustness and biological interpretation.

In addition, several R packages were used for visualization. Volcano plots were visualized using the R package EnhancedVolcano v1.2.0. Gene identifiers were retrieved using the R package BiomaRt v2.40.5. Principal component analysis (PCA) was performed in R using the prcomp on the scaled and median-centered log-transformed TPM (transcripts per million) values of the 1000 most variable genes according to the interquartile

---

**TABLE 1.** Demographic and phenotypic information on DYT-THAP1 patients and controls

| Identifier (SFC-code) | Number of iPSC clones | Gender | Age at biopsy (yr) | Clinical status | Disease-causing THAP1 mutation (protein alteration) |
|-----------------------|-----------------------|--------|-------------------|----------------|---------------------------------------------------|
| **Family A**          |                       |        |                   |                |                                                   |
| L-3969                | 1                     | Male   | 30                | Manifesting (generalized dystonia) | c.38G > A (p.Arg13His) |
| L-4011                | 2                     | Female | 56                | Nonmanifesting | c.38G > A (p.Arg13His) |
| **Family B**          |                       |        |                   |                |                                                   |
| L-3736                | 1                     | Male   | 70                | Nonmanifesting | c.474delA (p.Lys158Asnfs*23)                      |
| L-3735\textsuperscript{a} | 2                   | Male   | 69                | Manifesting (Writer’s dystonia followed by laryngeal dystonia) | c.474delA (p.Lys158Asnfs*23) |
| L-3737                | 2                     | Female | 40                | Nonmanifesting | c.474delA (p.Lys158Asnfs*23)                      |
| **Family C**          |                       |        |                   |                |                                                   |
| L-10249               | 2                     | Female | 29                | Manifesting    | c.62C > G (p.Ser21Cys)                            |
| L-10250               | 2                     | Female | 52                | Nonmanifesting | c.62C > G (p.Ser21Cys)                            |
| L-10252               | 2                     | Female | 57                | Nonmanifesting | c.62C > G (p.Ser21Cys)                            |
| L-10253               | 2                     | Female | 32                | Manifesting    | c.62C > G (p.Ser21Cys)                            |
| L-10260               | 2                     | Female | 44                | Nonmanifesting | c.62C > G (p.Ser21Cys)                            |
| **Controls**          |                       |        |                   |                |                                                   |
| L-5991 (SFC062)       | 1                     | Female | 67                | Neurologically healthy control | No mutation |
| L-6534 (SFC065)       | 2                     | Male   | 65                | Neurologically healthy control | No mutation |
| L-5872 (SFC067)       | 1                     | Male   | 72                | Neurologically healthy control | No mutation |
| L-6353 (SFC068)       | 1                     | Female | 68                | Neurologically healthy control | No mutation |
| L-4993 (SFC084)       | 1                     | Female | 63                | Neurologically healthy control | No mutation |
| L-6396 (SFC086)       | 2                     | Female | 57                | Neurologically healthy control | No mutation |
| L-6069 (SFC156)       | 1                     | Male   | 75                | Neurologically healthy control | No mutation |
| L-4646 (SFC163)       | 1                     | Male   | 66                | Neurologically healthy control | No mutation |
| SFC841                | 1                     | Male   | 36                | Neurologically healthy control | No mutation |
| **                 | 29                    | 55% male| Mean: 63.2; range: 36–75 |                 |                                                   |

All individuals are of Caucasian ancestry.

\textsuperscript{a}Also known as L-2257 in Djarmati et al.\textsuperscript{15}
range (IQR) across all samples. The ovals denote the 0.68 normal probability of the respective groups.

Putative transcription factor activity from RNA-Seq data was assessed using the human gene set resource DoRothEA²⁴ v1, which provides a curated collection of transcription factors and target gene interactions (the regulon) from different sources. Only interactions with high, likely, and medium confidence (levels A, B, C) were considered. Regulons were statistically evaluated using the R package viper²⁵ (v1.22.0; row-wise t tests) based on the TPM values of the expressed genes (mean TPM > 0.5) in the MMC and NMC samples. Only regulons having at least 15 expressed gene targets were considered.

FIG. 1. Legend on next page.
Validation of RNA-Seq Data by Quantitative PCR

The expression levels of randomly selected genes from the RNA-seq were validated by quantitative real-time (RT)-PCR using material from the same RNA isolation that was also used for the transcriptome analysis. For qRT-PCR assays, first-strand cDNA synthesis was carried out using the Maxima First-Strand cDNA Synthesis Kit (Thermo Fisher). qRT-PCR was performed using SYBR-Green on an LC480 (Roche, Basel, Switzerland) device with gene-specific primers for GAPDH, ISG15, LAMP1, MX1, SETX, THAP1, USP18, and USP21. Primer sequences and PCR conditions are available on request. Target gene expression (concentration) was normalized to the reference gene GAPDH. Statistical analysis was performed by one-way analysis of variance followed by Bonferroni post hoc correction using GraphPad Prism v5.02.

Results

After 90 days, human iPSC lines were successfully differentiated into glutamatergic cortical neurons following the stereotypical temporal order of human corticogenesis in vivo as described in Grütz et al (2017). Neurons at this stage of differentiation showed expression of cortical markers (Supplementary Figure S1). Because we were interested in transcriptomic changes at early developmental stages, iPSC-derived maturing cortical precursor cells were harvested already on day 44. At this stage, neurons are in a proliferative state, gradually beginning to fire spontaneously. The identity of the cortical progenitor cells and cortical neurons was defined by the expression of a core set of markers. Notably, the expression of the neurodevelopmental marker gene panels was at comparable levels using the transcriptome data across the different families and controls indicating similar stages of maturation (Fig. 1D; Supplementary Figures S2–S6).

After assuring that the cells are in the expected and comparable differentiation stage, we first wanted to evaluate the impact of different THAP1 mutations on the transcription of THAP1 itself, as a previously reported target gene, using qRT-PCR. We observed a uniform pattern of transcriptional upregulation of THAP1 in MMCs compared to NMCs, suggesting a potential overcompensation via the autoregulation loop of THAP1 (Fig. 2A). Statistical significance was reached in 2 families (p.Lys158Asnfs*23 and p.Ser21Cys), whereas p.Arg13His mutation carriers showed a similar trend. Notably, the family harboring the p.Arg13His alteration was the smallest with the fewest samples (n = 3) that were studied.

Next, to evaluate the variability in transcriptomes of MMCs, NMCs, and controls, we performed a PCA of the transcriptome data on the 1000 most varying genes. A normal fit for the 3 groups showed overlapping clusters. However, a subtle distinction of the MMCs and NMCs was evident (Fig. 2B).

To assess the factors of penetrance in THAP1 mutation carriers, we used the transcriptome data to look for gene expression differences between the MMC and NMC groups using family and sex as covariates. The biggest covariate was actually the family. The intrafamilial differences in gene expression were large between MMCs and NMCs, in part probably because of the small numbers of individuals per family. However, differentially expressed genes (DEGs) differed largely depending on the family so that the shared differential gene expression across all three families remained small after adjusting for chance findings. A Wald test showed only DRD4 as significantly upregulated in MMCs after correction for multiple testing (using the Benjamini–Hochberg method). Based on the P-value <0.01 and effect size >1 or <−1, we found 65 DEGs (Fig. 2C; Supplementary Table S2).

FIG. 1. Pedigrees of THAP1 mutation carriers and maturation of iPSC-derived cortical neurons. (A) Pedigrees of Families A, B, and C. Filled symbols indicate manifesting THAP1 mutation carriers; a black dot in the middle of a symbol represents nonmanifesting THAP1 mutation carriers; the index patients are highlighted by an arrow pointing toward the symbol; the question mark indicates missing information; the L-identifier is located below a symbol; asterisks show individuals from whom fibroblasts were collected, reprogrammed, and differentiated into iPSC-derived cortical neurons. (B) Timeline of the cortical differentiation protocol adapted from Shi et al (2012) as described in Grütz et al (2017). BDNF, brain-derived neurotrophic factor; Dor, dorsomorphin; FGF2, human fibroblast growth factor 2; GDNF, glial cell-derived neurotrophic factor; SB, factor SB-431542; KSR, neuronal induction medium; mTeSR1, iPSC culture medium; NMM, neural maintenance medium; Y, factor Y-27632. (C) While iPSCs before neuronal induction exhibited a flat colony-forming character (day 0), a timely coordinated supplementation of factors necessary for human corticogenesis resulted in a morphologic change. Neutralization of human iPSCs was achieved by dual SMAD inhibition, followed by replating of neural rosettes (day 30). After single-cell replating on day 44, neural precursor cells were morphologically characterized by short cellular extensions. After several weeks of maturation, the cells were interconnected, with elongated cellular projections resembling axons and dendrites (day 90). Scale bar represents 300 μm. (D) Gene expression panels of neurodevelopmental markers shown per family and controls. The box-scatter plots display the gene expression in transcripts per million (TPM) of marker sets that represent different neurodevelopmental stages during the differentiation process. The presence of a small population of GABAergic neurons is monitored by the expression of CALB2, GAD2, PVALB, SLC6A1, and SS7. The expression of DLG4, HOMER1, and SLC17A7 indicates the identity of excitatory glutamatergic neurons. The general neuronal markers ENO2, MAP2, and TUBB3 are expressed at high levels, indicating a post-progenitor stage of the major cortical cell population. The identity of the primary progenitor cells is characterized by the expression primary progenitor marker genes, including Pax6 and Nestin (NES). Secondary progenitor marker genes for cortical origin become a major population around day 35 and express BCL11B and CUX1. The gene expression panels are at comparable levels across the different groups indicating similar stages of neurodevelopmental maturation. Individual expression levels of marker genes can be found in Supplementary Figures S2–S6.
Differential gene expression analysis was also performed pairwise between the groups of MMCs, NMCs, and all mutation carriers (aMCs), respectively, versus the group of controls. There were 113 DEGs comparing aMCs and controls (q-value <0.1; effect size >0.5/−0.5; Supplementary Table S2). When considering also the disease status, the analysis showed a distinct expression pattern (84 DEGs) in MMCs when compared to controls (Supplementary Figure S7A). Strikingly, the NMCs showed only minor expression differences when compared to controls (10 DEGs), with only 6 overlapping DEGs with MMCs versus controls (Supplementary Figure S7B,C). According to the assumption that all THAP1 mutations lead to a loss of function, the expression differences between missense and truncation variants were not further explored. A list of all DEGs from the four comparisons (aMC vs. controls, MMCs vs. NMCs, MMCs vs. controls, and NMCs vs. controls) is provided in Supplementary Table S2. The expression of a random subset of genes

FIG. 2. Penetrance-linked expression in cortical precursor neurons of THAP1 mutation carriers. (A) Results of qPCR analysis targeting THAP1. THAP1 expression is grouped by THAP1 mutations (based on the 3 families) and clinical status and normalized to the housekeeping gene GAPDH (9 lines of nonmanifesting carriers [NMCs]: gray bars, 9 lines of manifesting carriers [MMCs]: red bars, and 14 lines of control neurons: black bar). Experiments were carried out independently thrice with 2 replicates; values represent mean ± SEM. For statistical analysis, one-way analysis of variance followed by Bonferroni’s post hoc test was conducted. *P < 0.05; **P < 0.001; ns, non-significant. (B) Principal component analysis showing the second and third components of the 1000 most varying genes. The different groups are encircled (MMC = red, NMC = green, controls = blue). (C) Volcano plot of differentially expressed genes (DEGs) between MMC and NMC versus controls.
was validated by qRT-PCR, which was congruent with the RNA-Seq data (Supplementary Figure S8).

To evaluate the pathways involved in the transcriptional changes in MMCs and NMCs, we performed a GSEA using the Reactome pathways. This analysis showed 32 upregulated pathways in MMCs and 115 upregulated pathways in NMCs (Supplementary Table S3). Figure 3 shows the 20 most significant pathways for both MMC and NMC groups, including “extracellular matrix organization” and “DNA methylation” in NMCs, whereas the MMCs showed various neurotransmitter release cycle pathways, such as dopamine and eIF2α.

To obtain an insight into the transcription factor network driving the respective gene expression in either MMC or NMC sample, we tested the human regulon data from the DoroTheA library in comparison of the MMC to the NMC samples taking sex and families as covariates. In total, there were only 4 and 1 transcription factors, respectively, whose putative activity was significantly higher in MMCs or NMCs, including SIX2 as the most downregulated and YY1 as the most upregulated in MMCs (Fig. 4A). Of note, there were more diverse differences for the comparison between MMCs and controls with RUNX2 and SOX2 downregulated and YY1 and THAP1 upregulated in MMCs (Fig. 4B).

**Discussion**

To the authors’ knowledge, this is the first genome-wide transcriptome analysis to systematically investigate reduced penetrance in DYT-THAP1. In this study, differential gene expression analysis in cortical precursor neurons showed a transcriptional signature that distinguished MMCs from NMCs. DRD4 encoding the G-protein-coupled dopamine receptor D4 was found to be significantly upregulated in MMCs. DRD4 belongs to the dopamine D2-like receptor family that is characterized by its ability to inhibit adenylyl cyclase.26 The relationships between dystonia and dopamine are manifold,27 although isolated dystonia, especially DYT-THAP1, is usually not responsive to dopaminergic treatment.28 However, a form of combined
dystonia, dopa-responsive dystonia is, as the name implies, dopa-responsive, and dopamine receptor agonists may be an effective treatment in at least a subset of dystonia patients.28 Further, using a KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis in a conditional knockout mouse model, dopaminergic synaptic signaling was among the top hits, triggered mainly by DRD2 dysregulation.29 Dopamine receptor genes, including DRD1, DRD2, and DRD3, have been tested repeatedly for genetic associations with (idiopathic) dystonia, and at least one single-nucleotide variant in DRD1 might increase the risk to develop dystonia.30 Alteration of striatal dopaminergic neurotransmission was found in a knockout mouse model of myoclonus dystonia, which is another form of combined dystonia.31 Our data now suggest that DRD4 expression levels might play a role in mediating reduced penetrance in THAP1 mutation carriers. The GSEA analysis further underlined dopamine signaling as a pathway involved in disease expression as one of the top hits among several neurotransmitter-linked pathways upregulated in MMCs (Fig. 3). Other pathways upregulated in MMCs included the response of EIK2AK1 and EIK2AK4. Interestingly, Thap1 mutant mice showed evidence of abnormal eIF2α signaling; the eIF2α signaling pathway was one of the top differentially regulated signaling pathways in the striatum and cerebellum of Thap1<sup>+/-</sup> mice.32 The eIF2α pathway is a key effector of the cellular response to several stressors, including the accumulation of misfolded proteins in the endoplasmic reticulum (ER).32 This pathway has also been linked to other genetic forms of dystonia, including DYT-TOR1A32 and DYT-PRKRA.33 Mutations in PRKRA cause a rare form of dystonia, with sometimes additional parkinsonian features.33 PRKRA encodes a stress-activated modulator of the eIF2α kinase PKR, with evidence of abnormal phosphorylation of PKR and eIF2α in patient fibroblasts under ER stress.34 Thus, it is plausible that eIF2α signaling is involved in the expressivity of THAP1 mutations.

On the contrary, extracellular matrix organization and DNA methylation were among the upregulated pathways based on the GSEA analysis in NMCs (Fig. 3). Collagens are extracellular matrix proteins with diverse functions, including structural adhesion, neural circuit formation, and maintenance.35 Type IV collagens are involved in the coordination of synaptogenesis and the stability of the synaptic networks.36 Mutations in a collagen VI gene, COL6A3, have been reported as a cause of dystonia.37 Using a zebrafish model, defective axonal targeting has been demonstrated in knockdowns of COL6A3 orthologs, suggesting that COL6A3 mutations disrupt the establishment of normal neural circuits during development.37 This suggests that the opposite scenario, upregulation of collagens and other extracellular matrix proteins, may be involved in reducing the penetrance of THAP1 mutations. Further, DNA methylation is an important epigenetic DNA modification and relevant for the regulation of gene expression. Time- and tissue-specific methylation regulates the development and function of different cells. Variable degrees of promoter methylation may result in subtle changes of gene expression and thus could influence the clinical expressivity in THAP1 mutation carriers.
Because THAP1 encodes a transcription factor and disruption of the transcriptional network has been implicated in several forms of dystonia,\(^9\) we also specifically looked for differences in transcription factor activity using human regulon data from the DoroThEA library and found YY1 to be upregulated in MMCs when compared to NMCs or controls. YY1 encodes a transcription factor with an established role in myelination, and YY1 mutations cause a neurodevelopmental disorder.\(^38\) Recently, it was shown that THAP1 regulates the DNA occupancy of the oligodendrocyte maturation factor YY1 and that the loss of THAP1 impaired myelination in the central nervous system through a cell-autonomous role in the oligodendrocyte lineage and decreased DNA occupancy.\(^39\) In contrast, SIX2 was upregulated in NMCs. SIX2 belongs to a gene family that plays an important role in brain development.\(^40,41\) Interestingly, a SIX2-mediated protective effect of glial cell line–derived neurotrophic factor on damaged dopaminergic neurons was recently demonstrated using a rat model in vivo and in vitro.\(^42\) Knockdown of SIX2 resulted in decreased cell viability and increased apoptosis of damaged dopaminergic neurons, whereas the overexpression of SIX2 increased cell viability and decreased cell apoptosis.\(^42\) With regard to reduced penetrance, the presumably increased cell viability in NMCs might be beneficial in THAP1 mutation carriers. Although our data provide new insights into the disease mechanism and mediation of reduced penetrance of THAP1 mutations, the cause of the dysregulated pathways and genetic network in maturing cortical neurons in MMCs needs to be elucidated further. It also seems to be linked to the altered transcription factor activity of THAP1 itself. This view is supported by the dysregulated THAP1 mRNA expression in MMCs observed in this study. The possible explanations for the altered THAP1 activity include sequence changes in regulatory DNA elements, epigenetic regulation, environmental factors, or merely stochastic events. The observed expression changes in our study are rather small but are in line with the hypothesis that mild expression changes during early brain development might be linked to a dysregulated cortical network in DYT-THAP1 patients, as suggested by repeatedly described cortico-striato-pallido-thalamic and cerebello-thalamo-cortical projection alterations in dystonia patients.\(^9,11,43\)

The previous pathway analysis of transcriptome data in animal or overexpressing cell models of THAP1 mutations or knockouts used Gene Ontology (GO) terms or KEGG pathways and implied cytoskeleton, dopamine signaling, eIF2α signaling, mitochondrial dysfunction, neuron projection development, axonal guidance, synaptic long-term depression, cell adhesion, and inflammatory response in the pathogenesis of DYT-THAP1.\(^29,32,35,39,44,45\) We also found most of these pathways among the top hits in the GSEA, although pathways are differently named (see Supplementary Table S3). The lack of a complete overlap of transcriptional changes among the different studies is most likely based on the different models used. Our human iPSC-derived model is based on the endogenous genetic background of the mutation carriers, which might contribute to the phenotypic expression. Further, we focused on cortical neurons in an early developmental stage in contrast to the previously used animal model.

Despite the many strengths of our study (unique model, quite large sample size, and inclusion of affected and unaffected mutation carriers), a limitation of the present study is the variability of the iPSC-derived cells that we used as a model. Even among the lines derived from the same individual, expression changes were detectable. For this reason, we included a relatively high number of samples (n = 29) with biological and technical replicates, matched for the ancestral background and gender, and reprogrammed all fibroblasts into iPSC lines using the same method. Human iPSC-derived neurons carry the same repertoire of genetic variants as the donor and therefore represent the best model to study the pathophysiologic mechanisms of reduced penetrance, especially during human brain development. RNA-Seq is a highly sensitive and versatile method to detect small expression differences, including the comparison of regions in postmortem brains,\(^46\) indicating the suitability of this approach to detect genetic modifiers in neurologic diseases.

Taken together, we identified DEGs and pathways exclusively upregulated in MMCs compared to NMCs and controls (but not in NMCs vs. controls), suggesting that these transcriptional alterations may be related to disease penetrance. Some of these DEGs are implicated in neuronal signaling, including dopamine transmitter and eIF2α signaling, as well as cortical network formation. This altered expression pattern might influence neural connectivity during early corticogenesis in mutation carriers. Although neurite extension during early development might induce structural changes in the brain, it is plausible that the disruption of these connections impairs motor learning and motor plasticity later in life, as shown for a DYT-TOR1A mouse model.\(^47\) Our results provide first insights into a possible disease mechanism in cortical neurons underlying DYT-THAP1 dystonia. Further studies, including systems biology approaches, are warranted to shed further light on the molecular basis of the differential gene expression in the THAP1 pathway. Advances in the understanding of DYT-THAP1 pathogenesis and its translation will help to develop novel therapeutic strategies in the future.
Acknowledgments: We thank the patients and their families for kindly donating biospecimen samples and for participating in this study.

References

1. Gajos A, Golańska E, Sieruta M, Szybka M, Liberski PP, Bogacki A. High variability of clinical symptoms in a Polish family with a novel THAP1 mutation. Int J Neurosci 2015;125:755–759.
2. 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:2098–2105.
3. Schirinzì T, Sciamanna G, Mercuri NB, Pisani A. Dystonia as a network disorder: a concept in evolution. Curr Opin Neurol 2018;31:498–503.
4. Gilbertson T, Humphries M, Steele JD. Maladaptive striatal plasticity and abnormal reward-learning in cervical dystonia. Eur J Neurosci 2019;50:3191–3204.
5. Jinnah HA, Neychev V, Hess EJ. The anatomical basis for dystonia: the motor network model. Tremor Other Hyperkinet Mov (N Y) 2017;7:506.
6. Balint B, Mencacci NE, Valente EM, et al. Dystonia. Nat Rev Dis Primers 2018;4:25.
7. Paudel R, Li A, Hardy J, Bhatia KP, Houlden H, Holton J. DYT6 Dystonia: a Neuropathological study. Neurodegener Dis 2016;16:273–278.
8. Sharma N. Neuropathology of Dystonia. Tremor Other Hyperkinet Mov (N Y) 2019;9:569.
9. Niethammer M, Carbon M, Argylean M, Eidelberg D. Hereditary dystonia as a neurodevelopmental circuit disorder: evidence from neuroimaging. Neurobiol Dis 2011;42:202–209.
10. Vo A, Sako W, Niethammer M, et al. Thalamocortical connectivity correlates with phenotypic variability in Dystonia. Cereb Cortex 2015;25:3086–3094.
11. Argylean M, Carbon M, Niethammer M, et al. Cerebellotalthalamocortical connectivity regulates penetrance in Dystonia. J Neurosci 2009;29:9740–9747.
12. Erogulları A, Hollstein R, Seibler P, et al. THAP1, the gene mutated in DYT6 dystonia, autoregulates its own expression. Biochim Biophys Acta 2014;1839:1196–1204.
13. Gavarini S, Cayrol C, Fuchs T, et al. Direct interaction between causative genes of DYT1 and DYT6 primary Dystonia. Ann Neurol 2010;68:349–353.
14. Zhao Y, Xiao J, Gong S, Clara JA, Ledoux MS. Neural expression of the transcription factor THAP1 during development in rat. Neuroscience 2013;231:282–295.
15. Djarrami A, Schneider SA, Lohmann K, et al. Mutations in THAP1 (DYT6) and generalised dystonia with prominent spasmodic dysphonia: a genetic screening study. Lancet Neurol 2009;8:447–452.
16. Zittel S, Moll CK, Brüggemann N, et al. Clinical neuroimaging and electrophysiological assessment of three DYT6 dystonia families. Mov Disord 2010;25:2405–2412.
17. Shi Y, Kirwan P, Livesey FJ. Directed differentiation of human pluripotent stem cells to cerebral cortex neurons and neural networks. Nat Protoc 2012;7:1836–1846.
18. Grutz K, Seibler P, Weissbach A, et al. Faithful SGCE imprinting in iPSC-derived cortical neurons: an endogenous cellular model of myoclonus-dystonia. Sci Rep 2017;7:41156.
19. Baumann H, Jahn M, Muenchau A, Trlacka-Winkler M, Lohmann K, Seibler P. Generation and characterization of eight human-derived iPSC lines from affected and unaffected THAP1 mutation carriers. Stem Cell Res 2018;33:60–64.
20. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods 2017;14:417–419.
21. Pimentel H, Bray NL, Puente S, Melsted P, Pachter L. Differential analysis of RNA-seq incorporating quantification uncertainty. Nat Methods 2017;14:687–690.
22. Luo W, Friedman MS, Sheldon K, Hankenson KD, Woolf PJ. GAGE: generally applicable gene set enrichment for pathway analysis, BMC Bioinformatics 2009;10:161.
23. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. Cell Syst 2015;1:417–425.
24. Garcia-Alonso L, Holland CH, Ibrahim MM, Turei D, Saez-Rodriguez J. Benchmark and integration of resources for the estimation of human transcription factor activities. Genome Res 2019;29:1363–1375.
25. Alvarez MJ, Shen Y, Giorgi FM, et al. Functional characterization of somatic mutations in cancer using network-based inference of protein activity. Nat Genet 2016;48:838–847.
26. Oldenhof J, Vickerzy R, Anaﬁ M, et al. SH3 binding domains in the dopamine D4 receptor. Biochemistry 1998;37:15726–15736.
27. Ribot B, Auppy J, Vidalhét M, et al. Dystonia and dopamine: from phenomenology to pathophysiology. Prog Neurobiol 2019;182:101678.
28. Fan X, Donsante Y, Jinnah HA, Hess EJ. Dopamine receptor agonist treatment of idiopathic dystonia: a reappraisal in humans and mice. J Pharmacol Exp Ther 2018;365:20–26.
29. Frederick NM, Shah PV, Didonna A, Langley MR, Kanashian YM, Seibler P. Loss of the dystonia gene THAP1 leads to transcriptional deﬁcits that converge on common pathogenic pathways in dystonic syndromes. Hum Mol Genet 2019;28:1343–1356.
30. Ohle÷ O, Dobricic V, Lohmann K, Klein C, Lill CM, Bertram L. Field synopsis and systematic meta-analyses of genetic association studies in isolated dystonia. Parkinsonism Relat Disord 2018;57:50–57.
31. Zhang L, Yokoi F, Parsons DS, Standaert DG, Li Y. Alteration of striatal dopaminergic neurotransmission in a mouse model of DYT11 myoclonus-dystonia. PLoS One 2012;7:e33669.
32. 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:e1007169.
33. Camargos S, Scholz S, Simon-Sanchez J, et al. DYT16, a novel young-onset dystonia-parkinsonism disorder: identiﬁcation of a segregating mutation in the stress-response protein PRKRA. Lancet Neurol 2008;7:207–215.
34. Bragg DC, Armana IA, Nery FC, Takeﬁeld XO, Sharma N. Molecu-lar pathways in dystonia. Neurobiol Dis 2011;42:136–147.
35. Domingo A, Erro R, Lohmann K. Dystonia novel genes: clues on disease mechanisms and the complexities of high-throughput sequencing. Mov Disord 2016;31:471–477.
36. Fox MA, Sanes JR, Borza DB, et al. Distinct target-derived signals organize formation, maturation, and maintenance of motor nerve terminals. Cell 2007;129:179–193.
37. Zech M, Lam DD, Francescatto L, et al. Recessive mutations in the alpha3 (VI) collagen gene COL6A3 cause early-onset isolated dystonia. Am J Hum Genet 2015;96:883–893.
38. Gabriele M, Vitulo-van Silfhout AT, Germain PL, et al. YY1 homeodomain of the human SIX3 gene cause holoprosencephaly. Brain 2007;130:147–156.
39. Yellajoshyula D, Liang CC, Pappas SS, et al. The DYT6 dystonia protein THAP1 regulates myelination within the Oligodendrocyte lineage. Dev Cell 2017;42:52–67 e54.
40. Lachabwan F, Solomon BD, Roessler E, et al. Clinical spectrum of SIX3-associated mutations in holoprosencephaly: correlation between genotype, phenotype and function. J Med Genet 2009;46:389–398.
41. Wallis DE, Roessler E, Hehr U, et al. Mutations in the homeodomain of the human SIX3 gene cause holoprosencephaly. Nat Genet 1999;22:196–198.
42. Gao J, Kang XY, Sun S, et al. Transcription factor Six2 mediates the protection of GDNF on 6-OHDA lesioned dopaminergic neurons by regulating Smurf1 expression. Cell Death Dis 2016;7:e2217.
43. Ramdhani RA, Simonyan K. Primary dystonia: conceptualizing the disorder through a structural brain imaging lens. Tremor Other Hyperkinet Mov (N Y) 2013;3:1–11.
44. Aguilo F, Zakirova Z, Nolan K, et al. THAP1: role in mouse embryonic stem cell survival and differentiation. Stem Cell Reports 2017; 9:92–107.

45. Cheng F, Walter M, Wassouf Z, et al. Unraveling molecular mechanisms of THAP1 missense mutations in DYT6 dystonia. J Mol Neurosci 2020;70:999–1008.

46. Monzon-Sandoval J, Poggiolini I, Ilmer T, Wade-Martins R, Webber C, Parkkinen L. Human-specific Transcriptome of ventral and dorsal midbrain dopamine neurons. Ann Neurol 2020;87:853–868.

47. Maltese M, Stanic J, Tassone A, et al. Early structural and functional plasticity alterations in a susceptibility period of DYT1 dystonia mouse striatum. Elife 2018;7:e33331.

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site.