DYT6 dystonia: Review of the literature and creation of the UMD locus-specific database (LSDB) for mutations in the THAP1 gene

Arnaud Blanchard, Vuthy Ea, Agathe Roubertie, Mélanie Martin, Coline Coquart, Mireille Claustres, Christophe Béroud, Gwenaelle Collod-Beroud

DYT6 Dystonia: Review of the Literature and Creation of the UMD Locus-Specific Database (LSDB) for Mutations in the *THAP1* Gene

Arnaud Blanchard,[1,2] Vuthy Ea,[1,2] Agathe Roubertie,[1,2,3] Mélanie Martin,[1,2] Coline Coquart,[4] Mireille Claustres,[1,2,4] Christophe Béroud,[1,2,4] and Gwenaëlle Collod-Béroud[1,2]*

[1] INSERM U827, Montpellier, F-34000, France; [2] Université Montpellier1, UFR de Médecine, Montpellier, F-34000, France; [3] CHU Montpellier, Hôpital Gui de Chauliac, Service de Neuropédiatrie, Montpellier, F-34000, France; [4] CHU Montpellier, Hôpital Arnaud de Villeneuve, Laboratoire de Génétique Moléculaire, Montpellier, F-34000, France

ABSTRACT: **By family-based screening, first Fuchs and then many other authors showed that mutations in *THAP1* (THAP [thanatos-associated protein] domain-containing, apoptosis-associated protein 1) account for a substantial proportion of familial, early-onset, nonfocal, primary dystonia cases (DYT6 dystonia). THAP1 is the first transcriptional factor involved in primary dystonia and the hypothesis of a transcriptional deregulation, which was primarily proposed for the X-linked dystonia-parkinsonism (DYT3 dystonia), provided thus a new way to investigate the possible mechanism underlying the development of dystonic movements. Currently, 56 families present with a *THAP1* mutation; however, no genotype/phenotype relationship has been found. Therefore, we carried out a systematic review of the literature on the *THAP1* gene to colligate all reported patients with a specific *THAP1* mutation and the associated clinical signs in order to describe the broad phenotypic continuum of this disorder. To facilitate the comparison of the identified mutations, we created a Locus-Specific Database (UMD-*THAP1* LSDB) available at http://www.umd.be/THAP1/. Currently, the database lists 56 probands and 43 relatives with the associated clinical phenotype when available. The identification of a larger number of *THAP1* mutations and collection of high-quality clinical information for each described mutation through international collaborative effort will help investigating the structure–function and genotype–phenotype correlations in DYT6 dystonia.**

KEY WORDS: dystonia; DYT6; *THAP1*, database; UMD

Introduction

Dystonia defines a heterogeneous group of movement disorders due to central nervous system (CNS) dysfunction leading to sustained involuntary muscle contractions that cause abnormal postures and twisting movements. Since secondary dystonia is often linked to striatal injuries, it has been hypothesized that hereditary dystonia also could be due to a dysfunction of these structures and more broadly of the basal ganglia. Accordingly, many studies have reported electrophysiological, metabolic, and structural abnormalities in these brain areas [Vidailhet et al., 2009]. However, basal ganglia are not the only brain region that has been found to be altered in hereditary dystonia. Indeed, dysfunction of cerebral cortex, cerebellum, spinal cord, and other structures that are mainly involved in motor function has also been reported [Breakefield et al., 2008; Defazio et al., 2007]. Therefore, the concept that dystonia is a disorder linked to a specific CNS structure appears to be too restrictive and it should rather be considered as the result of a general disturbance of the motor circuit functions.

Hereditary dystonias are clinically and genetically heterogeneous. To date, 20 different genetically forms are known (DYT1 to DYT21; with DYT14 = DYT5) and mutations in nine genes have been involved [Breakefield et al., 2008; Norgren et al., 2011]. All the inheritance modes are found (autosomic recessive, autosomic dominant, and X-linked) but no epistatic relation between these genes was discovered except for DYT1 and DYT6 forms [Gavarini et al., 2010; Kaiser et al., 2010]. Since their products have different cell functions, it is very difficult up to now to propose a scheme of cell dysfunction common to all the dystonia forms and particularly to propose a unitary pathophysiological mechanism even if links with dopamine deregulation are often found [Blanchard et al., 2010; Park et al., 2005].

The last identified gene, *THAP1* for THAP [thanatos-associated protein] domain-containing, apoptosis-associated protein 1, was first found to be mutated in Amish–Mennonite families with primary dystonia and its implication was then extended to other

*Correspondence to: Gwenaëlle Collod-Béroud, INSERM U827, Institut Universitaire de Recherche Clinique, 641 av du doyen Gaston Giraud, 34093 Montpellier cedex 5, France. E-mail: gwenaelle.collod-beroud@inserm.fr

Contract grant sponsors: Université Montpellier1; INSERM; French Ministry of Health (National PHRC 2007-A00614-49); AMADYS-LFCD, Alliance France Dystonie; Lions Club La Grande Motte; the French Dystonia Network; the European Community's Seventh Framework Programme (FP7/2007–2013 under grant agreement no. 200754—the GEN2PHEN project).

populations [Almasy et al., 1997; Fuchs et al., 2009]. *THAP1* mutations cause DYT6 dystonia, an autosomal dominant primary form with about 60% penetrance. The main clinical characteristics of this form are early onset (mean = 17.8 years of age [calculated from 78 patients with available data]) and symptoms often localized to one upper limb at the beginning with tendency to extend to other body regions.

Fifty-three different mutations in the *THAP1* gene have been reported so far in 56 families (Table 1). These mutations are mainly private, essentially missense or out-of-frame deletions, usually nonrecurrent, and widely distributed throughout the gene. To date, no genotype/phenotype relationship has been observed. From this perspective, we have compiled a database that gather all the available information on this gene and on the patients with *THAP1* mutations and created a software package to facilitate the mutational analysis of *THAP1*. This database is available at http://www.umd.be/THAP1/.

THAP1 Expression

Two spliced mRNA variants that produce functional proteins have been reported (*THAP1a*: CCDS6136 and *THAP1b*: CCDS6137) [Girard et al., 2010]. The first 2.2-kb isoform contains three exons (Fig. 1), whereas the second corresponds to an alternatively spliced isoform that lacks exon 2 (2-kb mRNA). This second isoform encodes a truncated THAP1 protein without the C-terminus of the THAP domain. The two isoforms are expressed in many tissues, suggesting that THAP1 has a widespread (although not ubiquitous) distribution in humans.

In mouse brain tissue, immunoblot analysis revealed a highest concentration in embryonic whole brain tissue (at E16), which declines after birth in the different tested brain regions (at P60) [Gavarini et al., 2010].

THAP1a, which contains the THAP domain with DNA-binding properties, is the most studied.

THAP1 Is a Transcription Factor

THAP1 was first identified by two-hybrid screen of a high endothelial venule cell cDNA library using a cytokine bait [Roussigne et al., 2003]. THAP1 is a 213 amino-acid-long protein characterized by an N-terminal THAP domain (amino acids 1 to 81) [Bessiere et al., 2008] with DNA-binding properties, followed by a proline-rich region (amino acids 90 to 110), and a nuclear localization signal (amino acids 146 to 162) (Fig. 1A).

The protein is only found in the nucleus where it can display both a diffuse distribution and a discrete localization in nuclear dots. THAP1 colocalization with DAXX, a well-characterized protein that is expressed in promyelocytic leukemia (PML) nuclear bodies (NBs), indicates that these subnuclear regions are PML NBs. PML NBs are spheres of 0.1–1.0 μm, heterogeneous in composition, mobility, and function, that are present in most mammalian cell nuclei [Lallemand-Breitenbach and de The, 2010]. The PML protein is the key organizer that recruits an ever-growing number of proteins, residing constitutively or more often transiently. PML NBs facilitate partner protein posttranslational modification (acetylation, sumoylation, or phosphorylation) resulting in partner sequestration, activation, or degradation. They are proposed to fine-tune a wide variety of processes: induction of apoptosis and cellular senescence, inhibition of proliferation, maintenance of genomic stability, and antiviral responses [Bernardi and Pandolfi, 2007].

THAP1 interacts with the prostate apoptosis response-4 (PAR-4, HGNC-approved gene symbols Pawr), a pro-apoptotic protein that acts as a transcriptional regulator [Roussigne et al., 2003]. PAWR is present in low amounts in dendrites and synapses throughout the brain under normal conditions and its amount increases rapidly in neurons subjected to several triggers of apoptosis including trophic factor deprivation, oxidative stress, and excitotoxins [Duan et al., 1999a; Guo et al., 1998]. Pawr induction has been linked to neuronal death in various neurodegenerative diseases as Parkinson's [Duan et al., 1999b], Alzheimer [Guo et al., 1998], or amyotrophic lateral sclerosis [Pedersen et al., 2000]. More interestingly, Pawr has been shown to serve other important function in nervous system. It interacts with D2 dopamine receptor (D2DR) suggesting its potential role in modulating calcium-mediated dopaminergic signaling [Park et al., 2005].

THAP1 and PAWR exhibit pro-apoptotic activities (they increase the apoptosis sensitivity of mouse 3T3 fibroblasts to TNF-α and serum withdrawal when overexpressed) and they are both indirectly recruited to PML NBs by PML [Roussigne et al., 2003]. Roussigne and coworkers proposed that PML NBs regulate the functions of THAP1 and PAWR. They may be the site of assembly and/or post-translational modifications of the THAP1/PAWR complex or they may serve as a site of storage or degradation, thus regulating the free nucleoplasmic pool of PAWR and THAP1.

Extensive database search using the THAP domain identified about 100 THAP proteins among seven model animal organisms [Clouaire et al., 2005]. Interestingly, some of these proteins are tightly linked to the pRB (retinoblastoma protein)/E2F pathway, which regulates commitment of mammalian cells to DNA replication. The THAP domain is present in the zebrafish ortholog of E2F6 and in five *Caenorhabditis elegans* proteins (LIN-36, LIN-15B, LIN-15A, HIM-17, and GON-14), which are known to be genetically linked to LIN-35, the only Rb ortholog found in *C. elegans*. Based on these data, Cayrol et al. [2007] investigated whether THAP1 could be linked to the pRB/E2F pathway. They found that silencing as well as overexpression of *THAP1* led to cell cycle arrest at the G1/S transition. The similar effects following THAP1 silencing/overexpression suggest that, like many other proteins, THAP1 needs an optimal range of concentration to fulfill its physiological roles. DNA microarray analysis showed in both cases (upregulation or downregulation of *THAP1*) a reduction of the mRNA levels of cell cycle regulators and about 40 pRB/E2F-target genes [Cayrol et al., 2007], including *RRM1* (ribonucleotide reductase M1), *BIRC5* (survivin), *CDC2* (cell division cycle 2), *CCNB1* (cyclin B1) [Lacroix, 2007]. *RRM1* is activated at the G1/S transition and contains two predicted THAP1-binding sites (THABS) in its promoter. DNase I footprinting, electrophoretic mobility shift assay (EMSA) protein/DNA binding, and chromatin immunoprecipitation (ChIP) in proliferating endothelial cells demonstrated that endogenous THAP1 directly binds to the promoter of *RRM1* [Cayrol et al., 2007]. Moreover, Mazars et al. [2010] showed that THAP1 recruits HCF-1 (Host cell factor 1) to the *RRM1* promoter to allow its expression, thus adding further evidence for a role of THAP1 in cell cycle regulation. Although the list of cofactors that participate in THAP1 regulation of target genes is certainly not exhaustive, the available data clearly link THAP1 to apoptosis (through its interaction with PAWR) and to cell cycle (through HCF-1).

THAP1 Structure

The most extensive structural studies have focused on the THAP domain of THAP1 and implicitly serve as a model for the other hundred THAP proteins discovered in the animal kingdom. In many aspects, the THAP domain, which is about 80-amino-acid long,

Table 1. Mutations Reported in the *THAP1* Gene in Each Family/Proband

Mutation name	Nomenclature (protein)	Wild Type codon	Mutant codon	Mutational event	Amino acid position	Exon no.	Structure	HCD	Relatives in the database	Reference no.
Insdel mutation										
c.135_139delinsGGGTTTA,[a]	p.Phe45LeufsX28	CCC	Indels	Indels stop at 73	47	2	L3-THAP dom.		25	[Fuchs et al., 2009; Bressman et al., 2009]
Missense mutations										
c.1A>G	p.Met1?	ATG	GTG	A->G	1	1	THAP dom.		1	[De Carvalho Aguiar et al., 2010]
c.16T>C	p.Ser6Pro	TCC	CCC	T->C	6	1	L1-THAP dom.		1	[Clot et al., 2010]
c.17C>T	p.Ser6Phe	TCC	TTC	C->T	6	1	L1-THAP dom.		0	[Houlden et al., 2010]
c.23A>G	p.Tyr8Cys	ATG	TGC	A->G	8	1	L1-THAP dom.		0	[Houlden et al., 2010]
c.25G>T	p.Gly9Cys	GGC	TGC	G->T	9	1	L1-THAP dom.		1	[Xiao et al., 2010]
c.36C>A	p.Asn12Lys	AAC	AAA	C->A	12	1	L1-THAP dom.		0	[Bressman et al., 2009]
c.38G>A	p.Arg13His	CGC	CAC	G->A	13	1	L1-THAP dom.		1	[Zittel et al., 2010]
c.50A>G	p.Asp17Gly	GAC	GGC	A->G	17	1	L1-THAP dom.		0	[Xiao et al., 2010]
c.61T>A	p.Ser21Thr	TCT	ACT	T->A	21	1	L1-THAP dom.		0	[Bressman et al., 2009]
c.68A>C	p.His23Pro	CAC	CCC	A->C	23	1	BS1-THAP dom.		0	[Kaiser et al., 2010]
c.77C>G	p.Pro26Arg	CCT	CGT	C->G	26	2	L2-THAP dom.	DNA binding	0	[Houlden et al., 2010]
c.86G>A	p.Arg29Gln	CGA	CAA	G->A	29	2	L2-THAP dom.	DNA binding	0	[Paisan-Ruiz et al., 2009]
c.86G>A	p.Arg29Gln	CGA	CAA	G->A	29	2	L2-THAP dom.	DNA binding	0	[Paisan-Ruiz et al., 2009]
c.86G>C	p.Arg29Pro	CGA	CCA	G->C	29	2	L2-THAP dom.	DNA binding	0	[Bressman et al., 2009]
c.89C>G	p.Pro30Arg	CCC	CGC	C->G	30	2	L2-THAP dom.		1	[Jech et al., 2011]
c.95T>A	p.Leu32His	CTT	CAT	T->A	32	2	L2-THAP dom.	AA interactions	2	[Schneider et al., 2011]
c.115G>A	p.Ala39Thr	GCT	ACT	G->A	39	2	AH1-THAP dom.	AA interactions	0	[Bressman et al., 2009]
c.161G>A	p.Cys54Tyr	TGT	TAT	G->A	54	2	BS2-THAP dom.	C2CH motif	0	[Gavarini et al., 2010]
c.169C>A	p.His57Asn	CAC	AAC	C->A	57	2	AH2-THAP dom.	C2CH motif	0	[Söhn et al., 2010]
c.176C>T	p.Thr59Ile	ACT	ATT	C->T	59	2	THAP dom.		0	[Groen et al., 2010]
c.215T>G	p.Leu72Arg	CTG	CGG	T->G	72	2	L4-THAP dom.		0	[Clot et al., 2010]
c.224A>T	p.Asn75Ile	AAT	ATT	A->T	75	2	L4-THAP dom.		1	[Cheng et al., 2010]
c.241T>C	p.Phe81Leu	TTT	CTT	T->C	81	2	AH4-THAP dom.	AA interactions	3	[Fuchs et al., 2009]
c.247T>C	p.Cys83Arg	TGT	CGT	T->C	83	2			0	[Söhn et al., 2010]
c.266A>G	p.Lys89Arg	AAG	AGG	A->G	89	2			0	[Bressman et al., 2009]
c.395T>C	p.Phe132Ser	TTC	TCC	T->C	132	3			0	[Xiao et al., 2010]
c.407A>G	p.Asn136Ser	AAC	AGC	A->G	136	3		HCF-1 binding	0	[Houlden et al., 2010]
c.408C>G	p.Asn136Lys	AAC	AAG	C->G	136	3		HCF-1 binding	1	[Groen et al., 2010]
c.410A>G	p.Tyr137Cys	TAT	TGT	A->G	137	3		HCF-1 binding	0	[Söhn et al., 2010]
c.427A>G	p.Met143Val	ATG	GTG	A->G	143	3	Coiled-coil dom.		0	[Söhn et al., 2010]
c.446T>C[b]	p.Ile149Thr	ATT	ACT	T->C	149	3	NLS		1	[Xiao et al., 2010]
c.449A>C	p.His150Pro	CAT	CCT	A->C	150	3	NLS		0	[Cheng et al., 2010]
c.496G>A	p.Ala166Thr	GCA	ACA	G->A	166	3	Coiled-coil dom.		0	[Xiao et al., 2010]
c.506G>A	p.Arg169Gln	CGA	CAA	G->A	169	3	Coiled-coil dom.		0	[Houlden et al., 2010]
c.508T>C	p.Cys170Arg	ATG	CGC	T->C	170	3	Coiled-coil dom.		0	[Bonetti et al., 2009]
c.559C>A	p.Gln187Lys	CAG	AAG	C->A	187	3	Coiled-coil dom.		0	[Xiao et al., 2010]
c.574G>A	p.Asp192Asn	GAC	AAC	G->A	192	3			0	[Söhn et al., 2010]

(*Continued*)

Table 1. Continued

Mutation name	Nomenclature (protein)	Wild Type codon	Mutant codon	Mutational event	Amino acid position	Exon no.	Structure	HCD	Relatives in the database	Reference no.
Nonsense mutations										
c.7C>T	p.Gln3X	CAG	TAG	C->T	3	1	L1-THAP dom.		0	[Houlden et al., 2010]
c.85C>T	p.Arg29X	CGA	TGA	C->T	29	2	L2-THAP dom.	DNA binding	0	[Bressman et al., 2009]
c.150T>G	p.Tyr50X	TAT	TAG	T->G	50	2	L3-THAP dom.	DNA binding	0	[Houlden et al., 2010]
c.370C>T	p.Gln124X	CAG	TAG	C->T	124	3			0	[Söhn et al., 2010]
Small out-of-frame deletions										
c.2delT	p.Met1?	ATG	del1b	Stop at 72	1	1	THAP dom.		0	[Bressman et al., 2009]
c.20_33del	p.Ala7GlufsX23	GCC	del14b	Stop at 29	7	1	L1-THAP dom.		0	[Clot et al., 2010]
c.174delT	p.Phe58LeufsX15	TTT	del1c	Stop at 72	58	2	THAP dom.	DNA binding	0	[Houlden et al., 2010]
c.197_198delAG	p.Glu66ValfsX19	GAG	del2b	Stop at 84	66	2	L4-THAP dom.		0	[Söhn et al., 2010]
c.236delC	p.Thr79LysfsX41	ACA	del1b	Stop at 119	79	2	AH4-THAP dom.	AVPTIF motif	0	[Houlden et al., 2010]
c.377_378delCT	p.Pro126ArgfsX2	CCT	del2b	Stop at 127	126	3			1	[Blanchard et al., 2011]
c.388_389delTC	p.Val131PhefsX3	TCA	del2a	Stop at 133	130	3			0	[Söhn et al., 2010]
c.388_389delTC[b]	p.Val131PhefsX3	TCA	del2a	Stop at 133	130	3			0	[Djarmati et al., 2010]
c.436_443del	p.Arg146AspfsX9	CGG	del8a	Stop at 154	146	3	NLS		0	[Clot et al., 2010]
c.460delC	p.Gln154SerfsX27	CAG	del1a	Stop at 180	154	3	NLS		0	[Bressman et al., 2009]
c.474delA[b]	p.Lys158AsnfsX23	AAA	del1c	Stop at 180	158	3	NLS		2	[Djarmati et al., 2010]
Small in-frame deletion										
c.207_209delCAA	p.Asn69_Asn69del	AAC	del3c	In-frame del	69	2	L4-THAP dom.		0	[Groen et al., 2010]
c.207_209delCAA	p.Asn69_Asn69del	AAC	del3c	In-frame del	69	2	L4-THAP dom.		2	[Clot et al., 2010]
Small out-of-frame insertion										
c.514dup	p.Arg172LysfsX7	AGG	ins1b	Stop at 178	172	3	Coiled-coil dom.		0	[Blanchard et al., 2011]

For each mutation, the nucleotide and protein names are given according to the nomenclature guidelines (http://www.hgvs.org/).
[a]Note that the c.134_135insGGGTT;137_139delAAC Amish mutation has been renamed in c.135_139delinsGGGTTTA to follow the guideline for "insdel" mutation (http://www.hgvs.org/).
[b]Same patient described in the two given references; "THAP dom.": THAP domain;
"L1-THAP dom.": THAP-domain loop 1;
"L2-THAP dom.": THAP-domain loop 2;
"L3-THAP dom.": THAP-domain loop 3;
"L4-THAP dom.": THAP-domain loop 4;
"AH1-THAP dom.": THAP-domain α-helix 1;
"H2-THAP dom.": THAP-domain 3_{10}helix 2;
"H3-THAP dom.": THAP-domain 3_{10}helix 3;
"H4-THAP dom.": THAP-domain 3_{10}helix 4;
"Coiled-coil dom.": coiled-coil domain;
"NLS": nuclear localization sequence;
"DNA binding -": amino acids proven to be involved in DNA-binding activity;
"C2CH motif": amino acids involved in the zinc-dependent DNA-binding activity of THAP1 through the C2CH motif;
"HCF-1 binding": amino acids involved in the interaction with HCF-1;
"AA interactions": interactions between THAP1 amino acids as described in Figure 1B. Example: THAP1 amino acid 39 is involved in interactions with THAP1 amino acids 40, 45, and 81.
For each proband, the number of relatives included in the database is given.

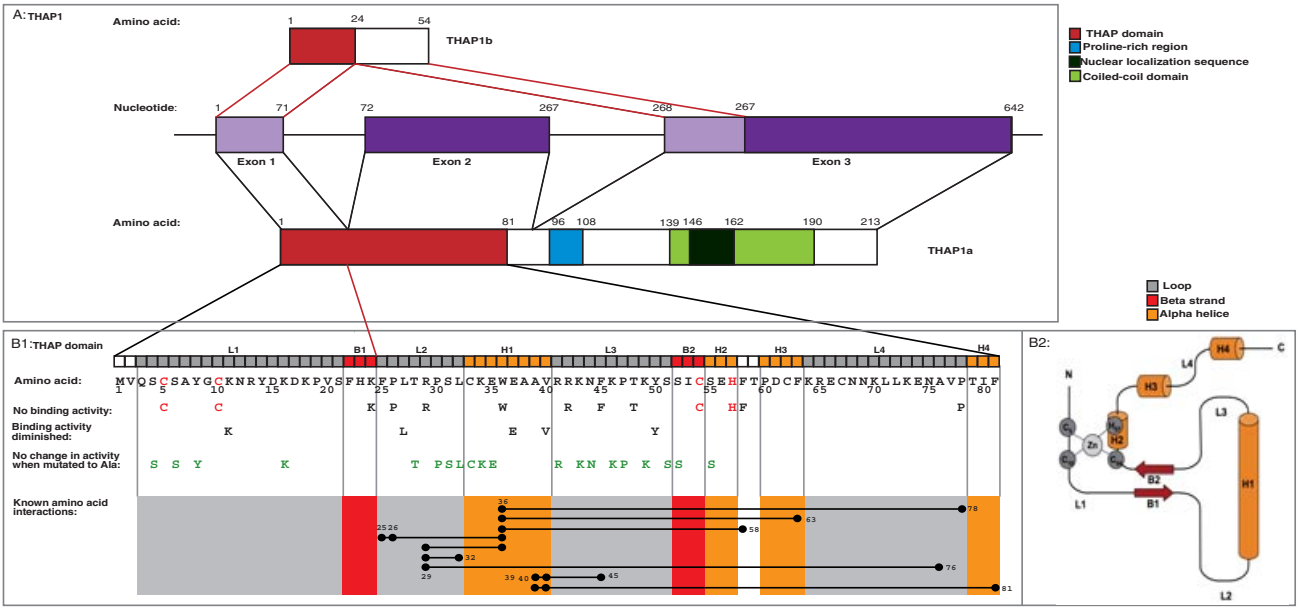

Figure 1. *THAP1* and its THAP domain. **A:** Schematic representation of the THAP1 gene and its two isoforms. Isoform 1 (THAP1a) and isoform 2 (THAP1b) result from alternative splicing of exon 2; **B1:** Schematic representation of the THAP domain with amino acid positions for each structure. The upper figure represents the α-helix H1, the three 3_{10}helices (H2, H3, and H4) (all in orange), the two β-sheets (B1 and B2) (in red), and the four loops (L1 to L4, in gray). "No binding activity": when these amino acids of the THAP domain are mutated to alanine, THAP1 shows loss of zinc-dependent DNA-binding activity; amino acids that are part of the C2CH motif are in red; "Binding activity diminished": when these amino acids are mutated to alanine, there is partial loss of DNA-binding activity; "No binding activity modification when mutated to Ala": mutation of these amino acids to alanine does not change the DNA-binding activity of THAP1 [Bessiere et al., 2008; Clouaire et al., 2005]; "Known amino acid interactions": structural representation of the known amino acid interactions [Bessiere et al., 2008]; **B2:** Topology diagram of the THAP domain of human THAP1 showing the secondary structure elements: the α-helix H1, the three 3_{10}helices (H2, H3, and H4) (all in orange), the two β-sheets (B1 and B2) (in red), and the location of the four loops (L1 to L4). The zinc finger and the four ligands are shown in gray. Modified from Bessiere et al. [2008].

exhibits some specific features (Fig. 1B1). It is characterized by a C2CH signature (Cys-X_{2-4}-Cys-X_{35-53}-Cys-X_2-His) associated with four invariant residues (Pro_{26}, Trp_{36}, Phe_{58}, Pro_{78}) in the *THAP1* sequence. Direct mutagenesis of each of these eight amino acids showed their critical identity for the zinc-dependent specific binding of the THAP domain to a precise DNA sequence (TXXXGGCA: THABS consensus sequence) [Clouaire et al., 2005] (Fig. 1B2). An AVPTIF box (Ala_{76}-Phe_{81}, in the THAP1 sequence), also essential for DNA binding, is located at the C-terminus of the THAP domain.

The NMR solution structure of the THAP zinc finger of THAP1 [Bessiere et al., 2008] indicates that this domain consists of a first loop (L1: Gln_3-Ser_{21}), followed by two short antiparallel β-sheets (B1: Phe_{22}-Lys_{24} and B2: Ser_{52}-Cys_{54}) separated by a loop (L2: Phe_{25}-Lys_{32})-helix (H1: Cys_{33}-Val_{40})-loop (L3: Arg_{41}-Ser_{51}) structure (Fig. 1B2). The four residues of the C2CH motif (Cys_5-Cys_{10}-Cys_{54}-H_{57}) are three-dimensionally brought together to form the zinc-dependent DNA-binding site. After the second β-strand, there are three 3_{10}helices: H2 (Ser_{55}-His_{57}), H3 (Arg_{60}-Phe_{63}), and H4 (Thr_{79}-Phe_{81}). H4 contains part of the AVPTIF motif and is preceded by a last mobile loop (L4: Lys_{64}-Pro_{78}).

The alpha-helix H1 can be considered as an important interaction hub and it is essential for the maintenance of the three-dimensional (3-D) shape of the THAP domain (Fig. 1B1). Indeed, it makes contacts with many residues distributed throughout the domain allowing a spatial connection between H1 (Trp_{36}, Ala_{39}, Val_{40}), L2 (Phe_{25}, Pro_{26}, Arg_{29}), L3 (Phe_{45}), ~H2 (Phe_{58}), H3 (Phe_{63}), L4 (Pro_{78}), and H4 (Phe_{81}). Residue Trp_{36}, which is located in the core of H1, seems to be a key amino acid for this function. It makes hydrophobic contacts with four amino acids (Phe_{25}, Arg_{29}, Phe_{58}, and Phe_{63}), and NMR data strongly suggest the presence of interaction also with the two highly conserved prolines (Pro_{26} and Pro_{78}). Four of these residues have been functionally tested (Pro_{26}, Arg_{29}, Phe_{58}, Pro_{78}) and their replacement by alanine strongly decreases the THAP domain ability to bind to DNA, suggesting that they have a critical role in the maintenance of THAP1 3-D shape [Clouaire et al., 2005].

The 3-D structure of the DNA–THAP domain complex has recently been solved [Campagne et al., 2010], showing that the THAP domain contacts DNA by insertion of its two antiparallel β-sheets into the major DNA groove. This ensures the connection of the second half of loop L3 (residues Lys_{46} to Ser_{51}) with the core of the THABS sequence. Upon binding to DNA, the folding of THAP1 is changed, enabling loop L4 to contact the DNA minor groove. Based on their nonconservation in the THAP family, specific base-contact with the core of the THABS motif and chemical behavior with non-specific DNA, some THAP domain residues have been proposed to play a crucial role in the specificity of DNA recognition: Tyr_{50} and Ser_{51} of loop L3, Ser_{52} of B2, Gln_3 and Ser_4 of the N-terminal tail and, finally Arg_{65} of L4. Other residues are involved in the nonspecific interaction with DNA, but their role remains absolutely necessary for the targeting of THAP1 to the THABS sequence. For example, Thr_{48}, which is located in the middle of the DNA-interaction area of loop L3, shows a chemical shift perturbation when the THAP domain is incubated with an unrelated DNA sequence, suggesting that Thr_{48} is involved in nonspecific interactions. However, mutation of this residue led to total loss of the DNA-binding activity of the THAP domain and it has been proposed that, with Lys_{46}, it may be important for positioning the protein onto DNA.

Table 2. Clinical Characteristics of Carriers of *THAP1* Mutations

	Amish patients [a] $N = 25$	Non-Amish patients $N = 81$	All patients [b] $N = 106$
Sex:	F 15 (60%)	F 46[c] (57%)	F 61[c] (58%)
Age at onset (y):	Median 14.5 (5–38)	Median 13 (2–62)	Median 13 (2–62)
Age at last examination (y):	Median 40 (10–66)	Median 44 (7–86)	Median 44 (7–86)
Family history:	100%	58.5%	61.4%
Site at onset:	**25 patients**	**77 patients**[d]	**102 patients**[d]
Upper limb	11 (44%)[e]	36 (47%)[e]	48 (47%)
Lower limb	1 (4%)[e]	14 (18%)[e]	17 (17%)
Cervical	5 (20%)[e]	17 (22%)[e]	23 (23%)
Cranial	8 (32%)	17 (22%)	25 (25%)
Site at examination:	**25 patients**	**81 patients**	**106 patients**
Upper limb	22 (88%)	60 (74%)	82 (77%)
Lower limb	12 (48%)	39 (48%)	51 (48%)
Cervical	14 (56%)	59 (73%)	73 (69%)
Cranial	18 (72%)[e]	40 (49%)[e]	58 (55%)
Speech	17 (68%)[e]	48 (59%)[e,f]	65 (61%)[e,f]
Distribution:	**25 patients**	**81 patients**	**106 patients**
	FD: 3 (12%)	FD: 16 (20%)	FD: 19 (18%)
	SD: 10 (40%)	SD: 20 (25%)	SD: 30 (28%)
	MD: 4 (16%)	MD: 5 (6%)	MD: 9 (8%)
	GD: 8 (32%)	GD: 40 (49%)	GD: 48 (45%)

Gender: F, female; M, male. Dystonia distribution: "GD": generalized dystonia;
"FD": focal dystonia;
"MD": multifocal dystonia;
"SD": segmental dystonia.
[a]Four families with common ancestry, 25 patients. Data from Fuchs et al. [Söhn et al., 2010].
[b]Thirty-five families, 22 sporadic, total 106 patients.
[c]Sex is unknown for one patient from Söhn et al. [Söhn et al., 2010].
[d]For some patients, data were unavailable [Bressman et al., 2009; Gavarini et al., 2010; Kaiser et al., 2010].
[e]Data were estimated from Bressman et al. [2009] and Fuchs et al. [2009] when possible.
[f]Of note, two patients from Groen et al. [2010] have oromandibular dystonia without knowing if there was speech disturbance or not.

THAP1 Is Involved in DYT6 Dystonia

DYT6 dystonia has been first described in the genetically isolated Amish–Mennonite population [Almasy et al., 1997] as an autosomal dominant trait with incomplete penetrance. In two Amish–Mennonite families, the site of onset of the disease was localized with similar frequency to upper limb, cervical, or cranial muscles and symptoms began between 5 and 38 years of age (average 18.1) with a tendency to progressive involvement of other body regions. Hence, DYT6 dystonia was named "idiopathic torsion dystonia of 'mixed' type." These families were large enough to identify by linkage analysis the morbid locus, which maps in both families to chromosome 8p21-22.

The locus was further narrowed by the discovery of other affected Amish–Mennonite families [Saunders-Pullman et al., 2007] and, in 2009, Fuchs et al. [2009] demonstrated the involvement of a *THAP1* mutation in DYT6 dystonia: an insertion/deletion founder mutation found in four related Amish–Mennonite families. A point mutation was then identified in an unrelated non-Amish–Mennonite German family (family S, c.214T>C), indicating that DYT6 dystonia was present also in populations of different ancestry. Indeed, the complex insertion/deletion mutation was later identified also in a German family with no known Amish–Mennonites origin but sharing the same haplotype [Bressman et al., 2009].

Several teams have reported to date 53 different mutations in the coding sequence of *THAP1* in 56 families (56 probands and 43 relatives) of American [Bressman et al., 2009; Xiao et al., 2010], Chinese [Cheng et al., 2010; Xiao et al., 2010], Czech [Jech et al., 2011], Dutch [Groen et al., 2010], English [Houlden et al., 2010], French [Blanchard et al., 2011; Clot et al., 2010; Houlden et al., 2010], German [Bressman et al., 2009; Djarmati et al., 2009; Fuchs et al., 2009; Houlden et al., 2010; Söhn et al., 2010; Zittel et al., 2010], Greek [Bonetti et al., 2009; Houlden et al., 2010], Iranian [Schneider et al., 2011], Jewish [Houlden et al., 2010], Mauritian/Indian [Houlden et al., 2010], Spanish [Paisan-Ruiz et al., 2009], and Brazilian [De Carvalho Aguiar et al., 2010] ancestry. Unfortunately, clinical data are not available for 11 probands implemented in the database [Bressman et al., 2009; Gavarini et al., 2010; Kaiser et al., 2010] and insufficient clinical and molecular data for 10 relatives do not allow to implement them in the database [Bressman et al., 2009]. One of the mutation has been described in an asymptomatic individual control (c.197_198delAG) [Söhn et al., 2010].

The enlargement of the group of patients with *THAP1* mutations (108 up to now without the asymptomatic carrier) has provided new data for delineating the phenotype of DYT6 dystonia, which nevertheless has hardly changed since the original description (Table 2). DYT6 dystonia has been reported slightly more in females than males. The age distribution at onset is broad, ranging from 2 to 62 years (mean 17.8 years [calculated from 78 patients with available data]). Onset before 10 years of age occurs in 35.8% of the patients, and onset after 28 years of age in one-sixth (16%) of the patients. In almost half (47%) of the patients, the first symptoms of dystonia concern an upper limb, while cranial or cervical onset is almost equally observed in the other half of the patients. Onset in a lower limb is less common (17/102 patients reported in the literature). DYT6 dystonia tends to spread slowly to other body regions. Although the symptom distribution varies widely, the involvement of the cranial region is frequent, and disability mainly results from cranial dystonia with speech difficulties. At last follow-up, almost half of the reported patients with DYT6 dystonia suffered from generalized dystonia, but most of them remained ambulatory. To date, 9 patients underwent deep brain stimulation; one good response is reported in one patient [Jech et al., 2011], although deep brain stimulation (DBS) resulted in only mild or moderate improvement

in the other patients, with poor effect on speech or dysarthria [Clot et al., 2010; Djarmati et al., 2009; Groen et al., 2010; Jech et al., 2011; Zittel et al., 2010].

Early onset in a limb and progressive generalization in few years are shared features between DYT6 and DYT1 dystonias [Valente and Albanese, 2010]. Nevertheless, several differences emerge. In DYT6 dystonia, cervical or cranial muscles are involved most of the time, whereas lower limbs are frequently spared. The DYT6 clinical spectrum is more variable, and some adult-onset cases with stable focal dystonia are reported as well. Onset (mean age) is somewhat later and with a wider age range in DYT6 than in DYT1 patients. The penetrance in DYT6 is difficult to evaluate; it was performed only in the Amish–Mennonite population (around 60% [Saunders-Pullman et al., 2007]).

Recently, *TOR1A* has been demonstrated as a direct target of *THAP1* by EMSA, ChIP, and luciferase reporter gene assays [Gavarini et al., 2010; Kaiser et al., 2010]. Specific modulation of torsinA expression was not reported in nonneuronal cells after THAP1 knockdown or overexpression, nor in fibroblasts or lymphoblast cells from DYT6 patients [Gavarini et al., 2010; Kaiser et al., 2010]. The lack of effect could be explained by tissue and/or developmental stage specificity. However, THAP1 is predicted to potentially regulate numerous gene targets [data not published] and *TOR1A* deregulation in DYT6 patients should then be responsible of only one part of the phenotype. Nevertheless, human wild-type torsinA can lead to a cellular dysfunction of mice neurons when overexpressed at high levels (abundant inclusion-like structures, abnormalities of the NE ultrastructure, behavioral abnormalities, significant reduction of dopamine levels, serotonin, and homoranillic acid) as shown in transgenic mice line hWT24 [Grundmann et al., 2007].

The UMD-THAP1 Database

We have used the Universal Mutation Database (UMD) software [Beroud et al., 2000, 2005] to create a computerized locus-specific database that contains information about the published mutations of the *THAP1* gene. Codons were numbered (e.g., +1 = A of ATG) based on the cDNA sequence of the *THAP1* variant 1 (isoform with three exons) (accession number NM_018105.2, "Homo sapiens THAP domain containing, apoptosis associated protein 1 (*THAP1*), transcript variant 1, mRNA") obtained from the GenBank database. Intron–exon boundaries as well as intronic sequences were defined by matching the cDNA sequence to the corresponding genome sequence (NT_167187.1) and to the module organization from SwissProt (accession number Q9NVV9). The database follows the guidelines on mutation databases of the Hugo Mutation Database Initiative including the latest nomenclature (http://www.hgvs.org/).

The *THAP1* mutation database can list point mutations, large and small deletions, insertions, indels, and mutations affecting splicing (intronic mutations) in the *THAP1* gene. It cannot accommodate mutations from the UTR and promotor regions. In addition, two mutations that affect the same allele are entered as two different records linked by the same sample ID. It allows to link specific polymorphisms identified in the course of sequencing to one mutation for a specific patient. Separate entries are made for each probands and for each relative (with different sample ID) in order to describe specific phenotypes. If the same mutation in a single patient has been reported in different papers (example: the same patients 7021, L-2257, L-3736, and L-3637 were first reported in [Djarmati et al., 2009] and subsequently in [Zittel et al., 2010]), only one entry was made. If the same mutation has been reported in apparently unrelated patients (example c.207_209delCAA identified in French and Dutch patients [Clot et al., 2010; Groen et al., 2010]), separate entries were made for each patient as recurrent mutations, in the absence of haplotypes demonstrating common ancestor. For each mutation, different types of information are provided: genetic (exon and codon number, wild type and mutant codon, mutational event, mutation name), protein (wild type and mutant amino acids, affected domain, mutation name), clinical data (age of onset, site of onset, dystonia distribution, sites involved, genealogic tree when available), and experimental data.

We have annotated the *THAP1* sequence with indirect arguments in order to determine whether potential missense mutations are really causative using the UMD-Predictor® tool [Frederic et al., 2009]. The UMD-Predictor® tool is dedicated to the analysis of nucleotide substitutions in cDNA sequences. It provides a combinatorial approach that includes localization of the mutation in the protein, conservation, biochemical properties of the mutant and wild type residue, and potential impact of the variation on mRNA (http://www.umd.be). The indirect arguments, listed in the "Highly conserved domain" (HCD) of UMD-Predictor® are: (1) amino acids involved in the zinc-dependent DNA-binding activity of THAP1 through the C2CH motif [Clouaire et al., 2005]; (2) amino acids involved in DNA binding of the THAP domain of human THAP1 [Bessiere et al., 2008; Clouaire et al., 2005]; (3) amino acids involved in the interaction with HCF-1 (HBM consensus sequence) [Mazars et al., 2010]; and (4) THAP1 amino acids involved in interactions with other THAP1 amino acids [Bessiere et al., 2008] (Fig. 1B1).

Mutations that potentially affect splicing can be highlighted with dedicated algorithms that have been recently implemented in the Human Splicing Finder (HSF) tool [Desmet et al., 2009]. HSF is also available for any gene as a web resource (http://www.umd.be/HSF/) and is dedicated to the prediction of the impact of a mutation on splicing signals and/or to the identification of these motifs in any human sequence. It contains all available matrices for auxiliary sequences prediction, as well as new position weight matrices to evaluate the strength of 5′ and 3′ splice sites and of branch point sequences.

The UMD-THAP1 Database can be used with different routines described at http://www.umd.be/THAP1/. It is possible to analyze the mutation distribution by exon or by mutation type and the mechanism for deletion/insertion. With the "HCD" annotation, it is possible to know whether an amino acid has a known reported function in the protein. The UMD software has already been used successfully since 1994 for the analysis of multiple genes implicated in genetic diseases, such as *LDLR* [Villeger et al., 2002], *TGFBR2* [Frederic et al., 2008], *DYS* [Krahn et al., 2009], and in cancer, such as *TP53* [Cariello et al., 1994a, b; Hamroun et al., 2006] or *MEN1* [Wautot et al., 2002]. Twenty-nine UMDs are available through the web. More information concerning the UMD software is available at http://www.umd.be.

The current database and the ensuing updated versions are available online at http://www.umd.be. Notifications of omissions and errors in the current version as well as specific phenotypic data will be gratefully received by the corresponding author. The software will be expanded as the database grows and according to the users' requirements, and new functions may be implemented.

Mutation Analysis

To date, a total of 53 different *THAP1* mutations are known (Table 1) since two of the three patients and their specific mutation reported by Zittel et al. [2010] were already described by Djarmati

et al. [2009], and the patient reported by Van Gerpen et al. [2010] is part of the cohort described by Xiao et al. [2010]. These mutations are mainly private, missense, and widely distributed throughout the gene. Only three apparently recurrent mutations have been described. For the c.86G>A, described in two patients by Paisan-Ruiz et al. [2009], no other polymorphism has been described ruling out common ancestor [Coro Paisan-Ruiz, personal communication]. The c.388_389delTC mutation [Djarmati et al., 2009; Söhn et al., 2010] and the c.207_209delCAA mutation [Clot et al., 2010; Groen et al., 2010] have been characterized by different teams, but until haplotype analyses become available, it is unclear whether these are truly recurrent mutations or whether they result from a founder effect. These mutations have then been considered as recurrent. It is noteworthy that two mutations, c.407A>G [Houlden et al., 2010] and c.95T>A [Schneider et al., 2011] have been found to be homozygous.

Insdel Mutations

The heterozygous Amish mutation, described as a 5-base (GGGTT) insertion followed by a 3-base deletion (AAC) (c.134_135insGGGTT;137_139delAAC) [Fuchs et al., 2009], has been detected in four related Amish–Mennonite families (families M, C, R, and W). This mutation was later identified also in a German family without known Amish–Mennonite ancestry (family 2) but sharing the Amish haplotype [Bressman et al., 2009]. This patient has then been implemented as a relative. To follow the international guideline for "insdel" mutation nomenclature (http://www.hgvs.org/), the name of the mutation has been corrected to c.135_139delinsGGGTTTA, that is deletion of five bases (TAAAC) and insertion of seven bases (GGGTTTA), which results in the same predicted mutant protein.

Nonsense and Missense Mutation

Four reported mutations are nonsense mutation: c.7C>T [Houlden et al., 2010], c.85C>T [Bressman et al., 2009], c.150T>G [Houlden et al., 2010], and c.370C>T [Söhn et al., 2010]. Thirty-seven are missense mutations. Five mutations are predicted to impair THAP1 DNA-binding activity: c.77C>G [Houlden et al., 2010]; the two recurrent mutations c.86G>A [Paisan-Ruiz et al., 2009]; c.86G>C [Bressman et al., 2009]; and c.169C>A in the C2CH motif [Söhn et al., 2010]. Three mutations are located in the sequence involved in THAP1 binding to HCF-1: c.407A>G [Houlden et al., 2010], c.408C>G [Groen et al., 2010], and c.410A>G [Söhn et al., 2010]. One mutation (c.241T>C) is located in the AVPTIF motif [Fuchs et al., 2009]. The homozygous mutation c.407A>G [Houlden et al., 2010] creates a potential acceptor splice site in exon 3, but its strength and the presence of many upstream good candidate acceptor splice sites make it unlikely that it will be used.

The c.266A>G mutation [Bressman et al., 2009] is located in the penultimate nucleotide of exon 2. This mutation does not impair the donor acceptor splice site, neither auxiliary splicing sequences (exonic sequence enhancer [ESE] or silencer [ESS]).

For the 24 remaining mutations, the mutated codon does not seem to be clearly involved in a particular function. Among all the missense mutations, only nine were not predicted by the UMD-Predictor® tool to be pathogenic based on a combinatorial approach [Frederic et al., 2009] (Table 3): c.38G>A [Zittel et al., 2010], c.61T>A [Bressman et al., 2009], c.169C>A [Söhn et al., 2010], c.241T>C [Fuchs et al., 2009], c.266A>G [Bressman et al., 2009], c.427A>G [Söhn et al., 2010], c.506G>A [Houlden et al., 2010], c.559C>A [Xiao et al., 2010], and c.574G>A [Söhn et al., 2010]. Nevertheless, mutation c.241T>C, reported in dbSNP without frequency data, shows a decreased DNA-binding activity in vitro [Fuchs et al., 2009]. EMSA experiments for c.61T>A show that the mutation abolishes binding of THAP1 to *TOR1A* [Bressman et al., 2009; Gavarini et al., 2010], while mutation c.38G>A is shown to significantly reduce THAP1-mediated repression of *TOR1A* (57% of residual activity)[Kaiser et al., 2010; Zittel et al., 2010].

UMD-Predictor® tool has a very high positive predictive value and a high negative predictive value. Mutations experimentally demonstrated by gel-shift assays (Table 1) [Bessiere et al., 2008; Clouaire et al., 2005] to have (1) no binding activity (p.Cys5Ala, p.Cys10Ala, p.Lys24Ala, p.Pro26Ala, p.Arg29Ala, p.Trp36Ala, p.Arg42Ala, p.Phe45Ala, p.Thr48Ala, p.Cys54Ala, p.His57Ala, p.Phe58Ala, and p.Pro78Ala) are all predicted as pathogenic; (2) a diminished binding activity were also all predicted as pathogenic (p.Glu37Ala, p.Lys11Ala, p.Leu27Ala, and p.Tyr50Ala) or probably pathogenic (p.Val40Ala); (3) no change in activity were either predicted: as pathogenic (p.Ser4Ala, p.Tyr8Ala, p.Lys16Ala, p.Leu32Ala, p.Arg41Ala, p.Asn44Ala, p.Lys46Ala, p.Lys49Ala); as probably pathogenic (p.Thr28Ala, p.Pro30Ala, p.Glu35Ala, p.Lys34Ala, p.Lys43Ala); as polymorphism (p.Ser55Ala); or as probable polymorphism (p.Ser6Ala, p.Pro47Ala, p.Ser31Ala, p.Cys33Ala, p.Ser51Ala, p.Ser52Ala). Nevertheless, these gel-shift assays demonstrate no modification of the DNA-binding activity "within the limits of detection of the present assay" and do not exclude "a role in the binding affinity and selectivity" [Bessiere et al., 2008]. Amino acids Ser_4, Ser_{51}, and Ser_{52} are described to be involved in the specificity of DNA recognition and Lys_{46} proposed, with Trp_{48}, to be important for positioning the protein onto DNA [Campagne et al., 2010]. Therefore, the predictions are in full agreement with in vitro studies for pathogenic mutations (18/18). For the alanine substitutions classified as "non modifying binding activities," some discrepancies are observed but their interpretation is difficult as they could result from wrong predictions by UMD-predictor or wrong interpretation of experimental data.

These experimental results show that in spite of the high efficiency of the UMD-predictor tool [Frederic et al., 2009] and other prediction tools [SIFT, PolyPhen, data not shown] for several genes (*FBN1*, *FBN2, TGFBR1*, *TGFBR2* [Frederic et al., 2009], *DYS* [Krahn et al., 2009], *CDKN2A* [Kannengiesser et al. 2009], and *ALAS2* [Ducamp et al. 2011]), more functional data will have to be collected to improve the prediction of missense variations in *THAP1* gene.

Insertion/Duplication

Only one reported mutation is an insertion that corresponds to the duplication of a single base (A_{514}). This insertion is predicted to create a premature termination codon (PTC), codon seven amino acids downstream of the mutation (position 178) [Blanchard et al., 2011].

Deletions

Among the 13 small deletions, 11 create a PTC and two are in-frame deletions resulting in loss of the ASN_{69} residues (c.207_209delCAA). Two single-base deletions, c.174delT [Houlden et al., 2010] and c.474delA [Djarmati et al., 2009], might be the result of slipped mispairing. Five mutations are deletions of a repeated sequence: c.197_198delAG [Söhn et al., 2010], the recurrent mutations c.388_389delTC [Djarmati et al., 2009; Söhn et al., 2010] and c.207_209delCAA [Clot et al., 2010; Groen et al., 2010]. One mutation, c.436_443del, is flanked by direct repeats [Clot et al., 2010]. For the five other mutations, the mechanisms are unknown

Table 3. Analysis of Missense Mutations with the UMD-Predictor® Tool

Nomenclature c.	Nomenclature p.	Structure	HCD	Conservation	SIFT probability	BLOSUM 62	Biochemical values	ESE modification	Splice site	Pathogenicity	Conclusion
c.17C>T	p.Ser6Phe	L1-THAP dom.		0.71	0.06	**-2.00**	**0.33**		No impact	82	Pathogenic
c.23A>G	p.Tyr8Cys	L1-THAP dom.		0.71	0.08	**-2.00**	**0.38**		PASSC [73.14]	76	Pathogenic
c.25G>T	p.Gly9Cys	L1-THAP dom.		0.71	**0.02**	**-3.00**	**0.46**		No impact	93	Pathogenic
c.50A>G	p.Asp17Gly	L1-THAP dom.		0.71	**0.05**	**-1.00**	**0.29**		No impact	79	Pathogenic
c.68A>C	p.His23Pro	BS1-THAP dom.		0.5	**0.00**	**-2.00**	**0.25**		No impact	88	Pathogenic
c.77C>G	p.Pro26Arg	L2-THAP dom.	DNA binding -	0.79	**0.02**	**-2.00**	**0.17**	-SC35 [2.73]	No impact	100	Pathogenic
c.86G>C	p.Arg29Pro	L2-THAP dom.	DNA binding -	0.79	**0.04**	**-2.00**	**0.17**		PASSC [75.5]	100	Pathogenic
c.89C>G	p.Pro30Arg	L2-THAP dom		0.79	**0.02**	**-2.00**	**0.17**		No impact	87	Pathogenic
c.95T>A	p.Leu32His	L2-THAP dom	L2-THAP dom	0.79	**0.01**	**-3.00**	**0.33**		No impact	100	Pathogenic
c.115G>A	p.Ala39Thr	AH1-THAP dom.	L2-THAP dom	0.79	**0.19**	**0.00**	**0.29**	-SF2/ASF [2.48]	No impact	88	Pathogenic
c.161G>A[a]	p.Cys54Tyr	BS2-THAP dom	C2CH motif.	0.79	**0.01**	**-2.00**	**0.38**		No impact	99	Pathogenic
c.215T>G	p.Leu72Arg	L4-THAP dom.		0.79	**0.13**	**-2.00**	**0.25**		No impact	76	Pathogenic
c.224A>T	p.Asn75Ile	L4-THAP dom.		0.71	**0.02**	**-3.00**	**0.33**		No impact	93	Pathogenic
c.247T>C	p.Cys83Arg			0.86	0.35	**-3.00**	**0.25**		No impact	82	Pathogenic
c.395T>C	p.Phe132Ser			1	0.77	**-2.00**	**0.33**		No impact	76	Pathogenic
c.408C>G	p.Asn136Lys		HCF-1 binding	1	0.06	0.00	**0.46**		No impact	76	Pathogenic
c.410A>G	p.Tyr137Cys		HCF-1 binding	1	0.10	**-2.00**	**0.38**		No impact	88	Pathogenic
c.446T>C	p.Ile149Thr	NLS		0.93	**0.02**	**-3.00**	**0.42**		No impact	87	Pathogenic
c.449A>C	p.His150Pro	NLS		0.93	0.24	**-2.00**	**0.25**		No impact	76	Pathogenic
c.508T>C	p.Cys170Arg	Coiled-coil dom.		0.86	**0.00**	**-3.00**	**0.25**		No impact	94	Pathogenic
c.1A>G	p.Met1?	THAP dom.		0.71	**0.00**	1.00	0.71		No impact	65	Probably pathogenic
c.16T>C	p.Ser6Pro	L1-THAP dom.		0.71	0.20	**-1.00**	**0.38**		No impact	71	Probably pathogenic
c.36C>A	p.Asn12Lys	L1-THAP dom.		0.71	0.22	0.00	**0.46**		No impact	65	Probably pathogenic
c.86G>A	p.Arg29Gln	L2-THAP dom.	DNA binding -	0.79	0.06	1.00	0.50		PASSC [71.63]	65	Probably pathogenic
c.176C>T	p.Thr59Ile	THAP dom.		0.79	0.07	**-1.00**	**0.42**		No impact	65	Probably pathogenic
c.407A>G	p.Asn136Ser		HCF-1 binding	1	0.13	1.00	0.63		**PASSC [85.88]**	71	Probably pathogenic
c.496G>A	p.Ala166Thr	Coiled-coil dom.		0.93	0.42	0.00	**0.29**		No impact	65	Probably pathogenic
c.38G>A[a]	p.Arg13His	L1-THAP dom.		0.71	**0.01**	0.00	0.58		No impact	64	Probable polymorphism
c.169C>A	p.His57Asn	AH2-THAP dom.	C2CH motif	0.79	0.05	1.00	0.58		No impact	59	Probable polymorphism
c.241T>C[a]	p.Phe81Leu	AH4-THAP dom.	AA interactions	0.79	0.45	0.00	0.71		No impact	59	Probable polymorphism
c.61T>A[a]	p.Ser21Thr	L1-THAP dom.		0.71	**0.00**	1.00	0.79		No impact	47	Polymorphism
c.266A>G	p.Lys89Arg			0.79	0.34	2.00	0.96		**PASSC [72.98]**	35	Polymorphism
c.427A>G	p.Met143Val	Coiled-coil dom.		0.93	1.00	1.00	0.71		No impact	35	Polymorphism
c.506G>A	p.Arg169Gln	Coiled-coil dom.		0.86	0.18	1.00	0.50		No impact	47	Polymorphism
c.559C>A	p.Gln187Lys	Coiled-coil dom.		0.86	1.00	1.00	0.54		PASSC [78.26]	41	Polymorphism
c.574G>A	p.Asp192Asn			0.93	0.43	1.00	0.75		No impact	35	Polymorphism

The UMD-predictor® tool computes all elements (structure, biochemistry, splicing, and conservation) and provides for each of them a specific strength based on its relative impact.
"Conservation": The score ranges from 0 for nonconserved to 1 for fully conserved.
"SIFT": (sorting intolerant from tolerant): uses sequence homology to predict whether an amino acid substitution will affect protein function and, hence, potentially confer a phenotype [Ng and Henikoff, 2003]. Scores that are ≤0.05 are considered to be deleterious (in bold).
"BLOSUM62" is an amino acid substitution matrix based on local multiple alignments of an unselected protein set of related sequences [Henikoff and Henikoff, 1992]. Positive scores are associated with conservative changes and negative scores (in bold) with less conservative changes.
"Biochemical value" is an amino acid substitution matrix depending on 48 qualitative physicochemical properties describing side chain structure and functional groups, optical properties, hydrophobicity/charge/acid–base properties, and size (volume and side chain length) [Yu, 2001]. A value below 0.05 defines an invalid substitution concerning the physico-chemical properties (in bold).
"ESE" (exonic splicing enhancer) and "Splice site" modifications are evaluated by the UMD tool.
"PASSC": potential acceptor splice site created; potential deleterious modifications are shown in bold.
"Pathogenicity": the UMD-predictor® tool computes all the elements (structure, biochemistry, splicing, and conservation) providing for each a specific strength based on their relative impact. Predictions are normalized on a scale from 0 to 100. In "Conclusion": a value of less than 50 is associated with the prediction of a nonpathogenic mutation annotated as "polymorphism"; a value of 50–64 is associated with the prediction of a "probable polymorphism"; a value of 65–74 is associated with the prediction of a "probably pathogenic" mutation; while a value above 74 is associated with the prediction of a "pathogenic" mutation. For "Structure" and "HCD" descriptions see the legends to Table 1. For more details see Frederic et al. [2009].
[a]These mutations have been studied in vitro [Fuchs et al., 2009; Gavarini et al., 2010; Kaiser et al., 2010].

and have to be determined by searching, among others, for the presence of quasi-palindromic sequences, inverted repeats, or symmetric elements that facilitate the formation of secondary-structure intermediates [Krawczak and Cooper, 1991].

Splice-Site Mutations

Only one splice site variation is described in the literature: c.72-4T>C (or c.IVS1-4T>C) in association with the missense mutation c.77C>G [Houlden et al., 2010]. The splice site variation has been analyzed with the UMD algorithm implemented in HSF tool [Desmet et al., 2009]. The software computed the variation of the consensus value (CV) of the mutant "splice site" (CV = 87.25) relative to the wild-type site (CV = 86.59). As the variation between these values was not higher than 10%, it is highly probable that such CV variation does not disrupt the wild-type splice site and thus this variation corresponds to a polymorphism as suggested also by the authors [Houlden et al., 2010]. The pathogenic mutation is thus likely to be c.77C>G and consequently the c.72-4T>C (or c.IVS1-4T>C) polymorphism was associated to the c.77C>G mutation in the database.

Conclusion

The UMD-*THAP1* database contains 53 different mutations (Table 1) in 56 probands and 43 relatives. These mutations are mainly private, essentially missense, usually nonrecurrent, and widely distributed in the THAP domain and the rest of the *THAP1* gene. The currently reported *THAP1* mutations are mainly missense mutations (64.9%, 37/57) and small out-of-frame deletions (19.3%, 11/57), with a minority of other mutation types (7% of nonsense [4/57], 3.5% of small in-frame deletions [2/57], 1.8% of small out-of-frame insertions [1/57], and 3.5% of complex mutations). To date, no clear genotype/phenotype relationship has been identified. This is not surprising when considering past experience with other disease genes, such as the *Fibrillin-1* gene (*FBN1*), for which many mutations were accumulated before the genotype/phenotype relationships emerged. Indeed, the UMD-*FBN1* database was created in 1995 [Collod et al., 1996], and the first report about the correlations between genotype and phenotype was published in 2007 [Faivre et al., 2007] and more followed [Detaint et al., 2010; Faivre et al., 2007, 2008, 2009a, b, c; Stheneur et al., 2009]. From this perspective, the UMD-*THAP1* locus-specific database will facilitate the mutational analysis of *THAP1* gene as previously demonstrated for other genes.

Acknowledgments

A.B. and V.E. are supported by PhD studentships from MESR (Ministère de l'Enseignement Supérieur et de la Recherche). The research leading to these results has also received funding from the Université Montpellier1, INSERM, the French Ministry of Health (National PHRC 2007-A00614-49), AMADYS-LFCD, Alliance France Dystonie, Lions Club La Grande Motte, the French Dystonia Network, and from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 200754—the GEN2PHEN project.

References

Almasy L, Bressman S, Raymond D, Kramer P, Greene P, Heiman G, Ford B, Yount J, de Leon D, Chouinard S, Saunders-Pullman R, Brin M, Kapoor R, Jones A, Shen H, Fahn S, Risch N, Nygaard T. 1997. Idiopathic torsion dystonia linked to chromosome 8 in two Mennonite families. Ann Neurol 42:670–673.

Bernardi R, Pandolfi PP. 2007. Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. Nat Rev Mol Cell Biol 8:1006–1016.

Beroud C, Collod-Beroud G, Boileau C, Soussi T, Junien C. 2000. UMD (Universal mutation database): a generic software to build and analyze locus-specific databases. Hum Mutat 15:86–94.

Beroud C, Hamroun D, Collod-Beroud G, Boileau C, Soussi T, Claustres M. 2005. UMD (Universal mutation database): 2005 update. Hum Mutat 26:184–191.

Bessiere D, Lacroix C, Campagne S, Ecochard V, Guillet V, Mourey L, Lopez F, Czaplicki J, Demange P, Milon A, Girard JP, Gervais V. 2008. Structure-function analysis of the THAP zinc finger of THAP1, a large C2CH DNA-binding module linked to Rb/E2F pathways. J Biol Chem 283:4352–4363.

Blanchard A, Roubertie A, Frederic MY, Claustres M, Collod-Beroud G. 2010. Monogenetic dystonia: revisiting the dopaminergic hypothesis. Rev Neurol (Paris) 166:389–399.

Blanchard A, Roubertie A, Simonetta-Moreau M, Ea V, Coquart C, Frederic M, Gallouedec G, Adenis J, Benatru I, Borg M, Burbaud P, Calvas P, Cif L, Damier P, Destee A, Faivre L, Guyant-Marechal L, Janik P, Janoura S, Kreisler A, Lusakowska A, Odent S, Potulska-Chromik A, Rudzińska M, Thobois S, Vuillaume I, Tranchant C, Tuffery-Giraud S, Coubes P, Sablonnière B, Claustres M, Collod-Béroud G. 2011. Singular DYT6 phenotypes in association with new THAP1 frameshift mutations. Mov Disord, DOI: 10.1002/mds.23641 in press.

Bonetti M, Barzaghi C, Brancati F, Ferraris A, Bellacchio E, Giovanetti A, Ialongo T, Zorzi G, Piano C, Petracca M, Albanese A, Nardocci N, Dallapiccola B, Bentivoglio AR, Garavaglia B, Valente EM. 2009. Mutation screening of the DYT6/THAP1 gene in Italy. Mov Disord 24:2424–2427.

Breakefield XO, Blood AJ, Li Y, Hallett M, Hanson PI, Standaert DG. 2008. The pathophysiological basis of dystonias. Nat Rev Neurosci 9:222–234.

Bressman SB, Raymond D, Fuchs T, Heiman GA, Ozelius LJ, Saunders-Pullman R. 2009. Mutations in THAP1 (DYT6) in early-onset dystonia: a genetic screening study. Lancet Neurol 8:441–446.

Campagne S, Saurel O, Gervais V, Milon A. 2010. Structural determinants of specific DNA-recognition by the THAP zinc finger. Nucleic Acids Res 38:3466–3476.

Cariello N, Cui L, Béroud C, Soussi T. 1994a. Database and software for the analysis of mutations in the human p53 gene. Cancer Res 54:4454–4460.

Cariello NF, Beroud C, Soussi T. 1994b. Database and software for the analysis of mutations at the human p53 gene. Nucleic Acids Res 22:3549–3550.

Cayrol C, Lacroix C, Mathe C, Ecochard V, Ceribelli M, Loreau E, Lazar V, Dessen P, Mantovani R, Aguilar L, Girard JP. 2007. The THAP-zinc finger protein THAP1 regulates endothelial cell proliferation through modulation of pRB/E2F cell-cycle target genes. Blood 109:584–594.

Cheng FB, Wan XH, Feng JC, Wang L, Yang YM, Cui LY. 2010. Clinical and genetic evaluation of DYT1 and DYT6 primary dystonia in China. Eur J Neurol 18:497–503.

Clot F, Grabli D, Burbaud P, Aya M, Derkinderen P, Defebvre L, Damier P, Krystkowiak P, Pollak P, Leguern E, San C, Camuzat A, Roze E, Vidailhet M, Durr A, Brice A. 2010. Screening of the THAP1 gene in patients with early-onset dystonia: myoclonic jerks are part of the dystonia 6 phenotype. Neurogenetics 12: 87–89.

Clouaire T, Roussigne M, Ecochard V, Mathe C, Amalric F, Girard JP. 2005. 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 102:6907–6912.

Collod G, Beroud C, Soussi T, Junien C, Boileau C. 1996. Software and database for the analysis of mutations in the human FBN1 gene. Nucleic Acids Res 24:137–140.

De Carvalho Aguiar P, Fuchs T, Borges V, Lamar KM, Silva SM, Ferraz HB, Ozelius L. 2010. Screening of Brazilian families with primary dystonia reveals a novel THAP1 mutation and a de novo TOR1A GAG deletion. Mov Disord 25:2854–2857.

Defazio G, Berardelli A, Hallett M. 2007. Do primary adult-onset focal dystonias share aetiological factors? Brain 130:1183–1193.

Desmet FO, Hamroun D, Lalande M, Collod-Beroud G, Claustres M, Beroud C. 2009. Human splicing finder: an online bioinformatics tool to predict splicing signals. Nucleic Acids Res 37:e67.

Detaint D, Faivre L, Collod-Beroud G, Child AH, Loeys BL, Binquet C, Gautier E, Arbustini E, Mayer K, Arslan-Kirchner M, Stheneur C, Halliday D, Beroud C, Bonithon-Kopp C, Claustres M, Plauchu H, Robinson PN, Kiotsekoglou A, De Backer J, Ades L, Francke U, De Paepe A, Boileau C, Jondeau G. 2010. Cardiovascular manifestations in men and women carrying a FBN1 mutation. Eur Heart J 31:2223–2229.

Djarmati A, Schneider SA, Lohmann K, Winkler S, Pawlack H, Hagenah J, Bruggemann N, Zittel S, Fuchs T, Rakovic A, Schmidt A, Jabusch HC, Wilcox R, Kostic VS, Siebner H, Altenmuller E, Munchau A, Ozelius LJ, Klein C. 2009. Mutations in THAP1 (DYT6) and generalised dystonia with prominent spasmodic dysphonia: a genetic screening study. Lancet Neurol 8:447–452.

Duan W, Rangnekar VM, Mattson MP. 1999a. Prostate apoptosis response-4 production in synaptic compartments following apoptotic and excitotoxic insults: evidence for a pivotal role in mitochondrial dysfunction and neuronal degeneration. J Neurochem 72:2312–2322.

Duan W, Zhang Z, Gash DM, Mattson MP. 1999b. Participation of prostate apoptosis response-4 in degeneration of dopaminergic neurons in models of Parkinson's disease. Ann Neurol 46:587–597.

Ducamp S, Kannengiesser C, Touati M, Garcon L, Guerci-Bresler A, Guichard JF, Vermylen C, Dochir J, Poirel HA, Fouyssac F, Mansuy L, Leroux G, Tertian G, Girot R, Heimpel H, Matthes T, Talbi N, Deybach JC, Beaumont C, Puy H, Grandchamp B. 2011. Sideroblastic anemia: molecular analysis of the ALAS2 gene in a series of 29 probands and functional studies of 10 missense mutations. Hum Mutat 32:590–597.

Faivre L, Collod-Beroud G, Callewaert B, Child A, Binquet C, Gautier E, Loeys BL, Arbustini E, Mayer K, Arslan-Kirchner M, Stheneur C, Kiotsekoglou A, Comeglio P, Marziliano N, Wolf JE, Bouchot O, Khau-Van-Kien P, Beroud C, Claustres M, Bonithon-Kopp C, Robinson PN, Ades L, De Backer J, Coucke P, Francke U, De Paepe A, Jondeau G, Boileau C. 2009a. Clinical and mutation-type analysis from an international series of 198 probands with a pathogenic FBN1 exons 24-32 mutation. Eur J Hum Genet 17:491–501.

Faivre L, Collod-Beroud G, Callewaert B, Child A, Loeys BL, Binquet C, Gautier E, Arbustini E, Mayer K, Arslan-Kirchner M, Kiotsekoglou A, Comeglio P, Grasso M, Beroud C, Bonithon-Kopp C, Claustres M, Stheneur C, Bouchot O, Wolf JE, Robinson PN, Ades L, De Backer J, Coucke P, Francke U, De Paepe A, Boileau C, Jondeau G. 2009b. Pathogenic FBN1 mutations in 146 adults not meeting clinical diagnostic criteria for Marfan syndrome: further delineation of type 1 fibrillinopathies and focus on patients with an isolated major criterion. Am J Med Genet A 149A:854–860.

Faivre L, Collod-Beroud G, Child A, Callewaert B, Loeys BL, Binquet C, Gautier E, Arbustini E, Mayer K, Arslan-Kirchner M, Stheneur C, Kiotsekoglou A, Comeglio P, Marziliano N, Halliday D, Beroud C, Bonithon-Kopp C, Claustres M, Plauchu H, Robinson PN, Ades L, De Backer J, Coucke P, Francke U, De Paepe A, Boileau C, Jondeau G. 2008. Contribution of molecular analyses in diagnosing Marfan syndrome and type I fibrillinopathies: an international study of 1009 probands. J Med Genet 45:384–390.

Faivre L, Collod-Beroud G, Loeys BL, Child A, Binquet C, Gautier E, Callewaert B, Arbustini E, Mayer K, Arslan-Kirchner M, Kiotsekoglou A, Comeglio P, Marziliano N, Dietz HC, Halliday D, Beroud C, Bonithon-Kopp C, Claustres M, Muti C, Plauchu H, Robinson PN, Ades LC, Biggin A, Benetts B, Brett M, Holman KJ, De Backer J, Coucke P, Francke U, De Paepe A, Jondeau G, Boileau C. 2007. Effect of mutation type and location on clinical outcome in 1,013 probands with Marfan syndrome or related phenotypes and FBN1 mutations: an international study. Am J Hum Genet 81:454–466.

Faivre L, Masurel-Paulet A, Collod-Beroud G, Callewaert BL, Child AH, Stheneur C, Binquet C, Gautier E, Chevallier B, Huet F, Loeys BL, Arbustini E, Mayer K, Arslan-Kirchner M, Kiotsekoglou A, Comeglio P, Grasso M, Halliday DJ, Beroud C, Bonithon-Kopp C, Claustres M, Robinson PN, Ades L, De Backer J, Coucke P, Francke U, De Paepe A, Boileau C, Jondeau G. 2009c. Clinical and molecular study of 320 children with Marfan syndrome and related type I fibrillinopathies in a series of 1009 probands with pathogenic FBN1 mutations. Pediatrics 123:391–398.

Frederic MY, Hamroun D, Faivre L, Boileau C, Jondeau G, Claustres M, Beroud C, Collod-Beroud G. 2008. A new locus-specific database (LSDB) for mutations in the TGFBR2 gene: UMD-TGFBR2. Hum Mutat 29:33–38.

Frederic MY, Lalande M, Boileau C, Hamroun D, Claustres M, Beroud C, Collod-Beroud G. 2009. UMD-predictor, a new prediction tool for nucleotide substitution pathogenicity—application to four genes: FBN1, FBN2, TGFBR1, and TGFBR2. Hum Mutat 30:952–959.

Fuchs T, Gavarini S, Saunders-Pullman R, Raymond D, Ehrlich ME, Bressman SB, Ozelius LJ. 2009. Mutations in the THAP1 gene are responsible for DYT6 primary torsion dystonia. Nat Genet 41:286–288.

Gavarini S, Cayrol C, Fuchs T, Lyons N, Ehrlich ME, Girard JP, Ozelius LJ. 2010. Direct interaction between causative genes of DYT1 and DYT6 primary dystonia. Ann Neurol 68:549–553.

Girard JP, Amalric F, Roussigne M, Clouaire T. 2010. Novel death associated proteins, and THAP1 and PAR4 pathway in apoptosis control. United States, Patent version number 20100021482.

Groen JL, Ritz K, Contarino MF, van de Warrenburg BP, Aramideh M, Foncke EM, van Hilten JJ, Schuurman PR, Speelman JD, Koelman JH, de Bie RM, Baas F, Tijssen MA. 2010. DYT6 dystonia: mutation screening, phenotype, and response to deep brain stimulation. Mov Disord 25:2420–2427.

Grundmann K, Reischmann B, Vanhoutte G, Hubener J, Teismann P, Hauser TK, Bonin M, Wilbertz J, Horn S, Nguyen HP, Kuhn M, Chanarat S, Wolburg H, Van der Linden A, Riess O. 2007. Overexpression of human wildtype torsinA and human DeltaGAG torsinA in a transgenic mouse model causes phenotypic abnormalities. Neurobiol Dis 27:190–206.

Guo Q, Fu W, Xie J, Luo H, Sells SF, Geddes JW, Bondada V, Rangnekar VM, Mattson MP. 1998. Par-4 is a mediator of neuronal degeneration associated with the pathogenesis of Alzheimer disease. Nat Med 4:957–962.

Hamroun D, Kato S, Ishioka C, Claustres M, Beroud C, Soussi T. 2006. The UMD TP53 database and website: update and revisions. Hum Mutat 27:14–20.

Henikoff S, Henikoff JG. 1992. Amino acid substitution matrices from protein blocks. Proc Natl Acad Sci USA 89:10915–10919.

Houlden H, Schneider SA, Paudel R, Melchers A, Schwingenschuh P, Edwards M, Hardy J, Bhatia KP. 2010. THAP1 mutations (DYT6) are an additional cause of early-onset dystonia. Neurology 74:846–850.

Jech R, Bares M, Krepelova A, Urgosik D, Havrankova P, Ruzicka E. 2011. DYT 6—a novel THAP1 mutation with excellent effect on pallidal DBS. Mov Disord 26:924–925.

Kaiser FJ, Osmanoric A, Rakovic A, Erogullari A, Uflacker N, Braunholz D, Lohnau T, Orolicki S, Albrecht M, Gillessen-Kaesbach G, Klein C, Lohmann K. 2010. The dystonia gene DYT1 is repressed by the transcription factor THAP1 (DYT6). Ann Neurol 68:554–559.

Kannengiesser C, Brookes S, del Arroyo AG, Pham D, Bombled J, Barrois M, Mauffret O, Avril MF, Chompret A, Lenoir GM, Sarasin A, Peters G, Bressac-de Paillerets B. 2009. Functional, structural, and genetic evaluation of 20 CDKN2A germ line mutations identified in melanoma-prone families or patients. Hum Mutat 30:564–574.

Krahn M, Beroud C, Labelle V, Nguyen K, Bernard R, Bassez G, Figarella-Branger D, Fernandez C, Bouvenot J, Richard I, Ollagnon-Roman E, Bevilacqua JA, Salvo E, Attarian S, Chapon F, Pellissier JF, Pouget J, Hammouda el H, Laforet P, Urtizberea JA, Eymard B, Leturcq F, Levy N. 2009. Analysis of the DYSF mutational spectrum in a large cohort of patients. Hum Mutat 30:E345–E375.

Krawczak M, Cooper D. 1991. Gene deletions causing human genetic disease: mechanisms of mutagenesis and the role of the local DNA sequence environment. Hum Genet 86:425–441.

Lacroix C. 2007. THAP1, un régulateur clé de la prolifération des cellules endothéliales: relations structure/fonction et gènes cibles. Toulouse: Université Toulouse III-Paul Sabatier . 154 p.

Lallemand-Breitenbach V, de The H. 2010. PML nuclear bodies. Cold Spring Harb Perspect Biol 2:a000661.

Mazars R, Gonzalez-de-Peredo A, Cayrol C, Lavigne AC, Vogel JL, Ortega N, Lacroix C, Gautier V, Huet G, Ray A, Monsarrat B, Kristie TM, Girard JP. 2010. The THAP-zinc finger protein THAP1 associates with coactivator HCF-1 and O-GlcNAc transferase: a link between DYT6 and DYT3 dystonias. J Biol Chem 285:13364–13371.

Ng PC, Henikoff S. 2003. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res 31:3812–3814.

Norgren N, Mattson E, Forsgren L, Holmberg M. 2011. A high-penetrance form of late-onset torsion dystonia maps to a novel locus (DYT21) on chromosome 2q14.3-q21.3. Neurogenetics 12:137–143.

Paisan-Ruiz C, Ruiz-Martinez J, Ruibal M, Mok KY, Indakoetxea B, Gorostidi A, Masso JF. 2009. Identification of a novel THAP1 mutation at R29 amino-acid residue in sporadic patients with early-onset dystonia. Mov Disord 24:2428–2429.

Park SK, Nguyen MD, Fischer A, Luke MP, Affar el B, Dieffenbach PB, Tseng HC, Shi Y, Tsai LH. 2005. Par-4 links dopamine signaling and depression. Cell 122:275–287.

Pedersen WA, Luo H, Kruman I, Kasarskis E, Mattson MP. 2000. The prostate apoptosis response-4 protein participates in motor neuron degeneration in amyotrophic lateral sclerosis. FASEB J 14:913–924.

Roussigne M, Cayrol C, Clouaire T, Amalric F, Girard JP. 2003. THAP1 is a nuclear proapoptotic factor that links prostate-apoptosis-response-4 (Par-4) to PML nuclear bodies. Oncogene 22:2432–2442.

Saunders-Pullman R, Raymond D, Senthil G, Kramer P, Ohmann E, Deligtisch A, Shanker V, Greene P, Tabamo R, Huang N, Tagliati M, Kavanagh P, Soto-Valencia J, Aguiar Pde C, Risch N, Ozelius L, Bressman S. 2007. Narrowing the DYT6 dystonia region and evidence for locus heterogeneity in the Amish-Mennonites. Am J Med Genet A 143A:2098–2105.

Schneider SA, Ramirez A, Shafiee K, Kaiser FJ, Erogullari A, Bruggemann N, Winkler S, Bahman I, Osmanovic A, Shafa MA, Bhatia KP, Najmabadi H, Klein C, Lohmann K. 2011. Homozygous THAP1 mutations as cause of early-onset generalized dystonia. Mov Disord 26:858–861.

Söhn AS, Glockle N, Doetzer AD, Deuschl G, Felbor U, Topka HR, Schols L, Riess O, Bauer P, Muller U, Grundmann K. 2010. Prevalence of THAP1 sequence variants in German patients with primary dystonia. Mov Disord 25:1982–1986.

Stheneur C, Collod-Beroud G, Faivre L, Buyck JF, Gouya L, Le Parc JM, Moura B, Muti C, Grandchamp B, Sultan G, Claustres M, Aegerter P, Chevallier B, Jondeau G, Boileau C. 2009. Identification of the minimal combination of clinical features in probands for efficient mutation detection in the FBN1 gene. Eur J Hum Genet 17:1121–1128.

Valente EM, Albanese A. 2010. Advances in the genetics of primary torsion dystonia. F1000 Biol Rep 2:41.

Van Gerpen JA, Ledoux MS, Wszolek ZK. 2010. Adult-onset leg dystonia due to a missense mutation in THAP1. Mov Disord 25:1306–1307.

Vidailhet M, Grabli D, Roze E. 2009. Pathophysiology of dystonia. Curr Opin Neurol 22:406–413.

Villeger L, Abifadel M, Allard D, Rabes JP, Thiart R, Kotze MJ, Beroud C, Junien C, Boileau C, Varret M. 2002. The UMD-LDLR database: additions to the software and 490 new entries to the database. Hum Mutat 20:81–87.

Wautot V, Vercherat C, Lespinasse J, Chambe B, Lenoir GM, Zhang CX, Porchet N, Cordier M, Beroud C, Calender A. 2002. Germline mutation profile of MEN1 in multiple endocrine neoplasia type 1: search for correlation between phenotype and the functional domains of the MEN1 protein. Hum Mutat 20:35–47.

Xiao J, Zhao Y, Bastian RW, Perlmutter JS, Racette BA, Tabbal SD, Karimi M, Paniello RC, Wszolek ZK, Uitti RJ, Van Gerpen JA, Simon DK, Tarsy D, Hedera P, Truong DD, Frei KP, Dev Batish S, Blitzer A, Pfeiffer RF, Gong S, LeDoux MS. 2010. Novel THAP1 sequence variants in primary dystonia. Neurology 74:229–238.

Yu K. 2001. Theoretical determination of amino acid substitution groups based on qualitative physicochemical properties. http://cmgm.stanford.edu/biochem218/Projects%202001/Yu.pdf.

Zittel S, Moll CK, Bruggemann N, Tadic V, Hamel W, Kasten M, Lohmann K, Lohnau T, Winkler S, Gerloff C, Schonweiler R, Hagenah J, Klein C, Munchau A, Schneider SA. 2010. Clinical neuroimaging and electrophysiological assessment of three DYT6 dystonia families. Mov Disord 25:2405–2412.