Helsinki University Biomedical Dissertations No. 83

Molecular biology of progressive myoclonus epilepsy of Unverricht-Lundborg type (EPM1)

Kirsi Alakurtti

Folkhälsan Institute of Genetics

Neuroscience Center
and
Department of Medical Genetics
University of Helsinki

Academic Dissertation
To be publicly discussed with the permission of the Medical Faculty of the University of Helsinki, in Lecture Hall 2, Biomedicum Helsinki, on December 8th 2006, at 12 noon

Helsinki 2006

Supervised by

Professor Anna-Elina Lehesjoki
Folkhälsan Institute of Genetics
and
Neuroscience Center
University of Helsinki
Helsinki, Finland

Reviewed by

Docent Jaana Tyynelä-Vesterinen
Institute of Biomedicine/Biochemistry
and
Neuroscience Research Program
University of Helsinki
Helsinki, Finland

Docent Ismo Ulmanen
National Public Health Institute
Department of Molecular Medicine
Helsinki, Finland

Official opponent

Docent Kirsi Huoponen
Department of Medical Genetics
University of Turku
Turku, Finland

ISSN 1457-8433
ISBN 952-10-3506-4 (paperback)
ISBN 952-10-3507-2 (PDF)
http://ethesis.helsinki.fi
Helsinki University Printing House
Helsinki 2006

To my parents Annikki and Matti

CONTENTS

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following original articles, which are referred to in the text by their Roman numerals. In addition, unpublished data are presented.

I **Kirsi Alakurtti***, Kimmo Virtaneva*, Tarja Joensuu, Jorma J. Palvimo, Anna-Elina Lehesjoki
Characterization of the cystatin B gene promoter harboring the dodecamer repeat expanded in progressive myoclonus epilepsy, EPM1
Gene 2000; 242 (1-2), 65-73

II **Kirsi Alakurtti**, Nina Aula, Antti Aalto, Outi Kopra, Tarja Joensuu, Saara Tegelberg, Deying Guo, Kirsi Sainio, Anna-Elina Lehesjoki
Developmental expression of the cysteine protease inhibitor, *Cstb* mRNA, in mouse
Submitted

III **Kirsi Alakurtti**, Ekkehard Weber, Riitta Rinne, Gerit Theil, Gerrit-Jan de Haan, Dick Lindhout, Paula Salmikangas, Pekka Saukko, Ulla Lahtinen, Anna-Elina Lehesjoki
Loss of lysosomal association of cystatin B proteins representing progressive myoclonus epilepsy, EPM1, mutations
European Journal of Human Genetics 2005; 13 (2), 208-215
Corrigenda: European Journal of Human Genetics 2005; 13 (2), 264

IV Tarja Joensuu, Mervi Kuronen, **Kirsi Alakurtti**, Saara Tegelberg, Paula Hakala, Antti Aalto, Laura Huopaniemi, Nina Aula, Roberto Michellucci, Kai Eriksson, Anna-Elina Lehesjoki
Cystatin B: mutation detection, alternative splicing and expression in progressive myclonus epilepsy of Unverrict-Lundborg type (EPM1) patients
European Journal of Human Genetics, in press

* These authors contributed equally to the respective article.

ABBREVIATIONS

AGA	aspartylglucosaminidase
AP1	activating protein 1
ARE	androgen response element
BHK cells	baby hamster kidney cells
BLAST	basic local alignment search tool
bp	base pair
c.	coding DNA reference sequence
cDNA	complementary DNA
CEPH	Centre d'Etude du Polymorphisme Humain
CNS	central nervous system
COS-1 cells	African green monkey kidney cells
CSTA	cystatin A
CSTB / CSTB	human cystatin B gene / protein
Cstb / Cstb	mouse cystatin B gene / protein
CSTB2 / CSTB2	human cystatin B 2 splice variant transcript / protein
Cstb2 / Cstb2	mouse cystatin B 2 splice variant transcript / protein
CST3 / CSTC	cystatin C gene / protein
CTSB	cathepsin B
CTSD	cathepsin D
CTSK	cathepsin K
CTSL	cathepsin L
CTSS	cathepsin S
del	deletion
DNA	deoxyribonucleic acid
DRPLA	dentatorubral-pallidoluysian atrophy
E	embryonic day
ELISA	enzyme-linked immunosorbent assay
EPM1	progressive myoclonus epilepsy of Unverricht-Lundborg type
EPM1B	Unverricht-Lundborg disease locus B
EPM2	Lafora disease
EST	expressed sequence tag
fs	frameshift
HGP	human genome project
IVT	*in vitro* translation
kb	kilobase, 1 000 bp

kDa	kiloDalton
LAMP-1	lysosome-associated membrane protein 1
LD	Lafora disease
MHC	major histocompatibility complex
MPR	mannose-6-phosphate receptor
mRNA	messenger RNA
mSec13	mammalian secretory protein 13
mtDNA	mitochondrial DNA
NCL	neuronal ceroid lipofuscinosis
OMIM	Online Mendelian Inheritance in Man
p	short arm of a chromosome
P	postnatal day
p.	protein reference sequence
PCR	polymerase chain reaction
PDI	protein disulfide isomerase
PME	progressive myoclonus epilepsy
q	long arm of a chromosome
RNA	ribonucleic acid
RT-PCR	reverse transcriptase PCR
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
Sp1	specificity protein 1
TGN46	*trans*-Golgi network protein 46
ULD	Unverricht-Lundborg disease
UTR	untranslated region

In addition, standard one-letter abbreviations of nucleic acids and amino acids as well as three-letter abbreviations of amino acids are used.

ABSTRACT

Progressive myoclonus epilepsy of Unverricht-Lundborg type (EPM1) is an autosomal recessively inherited disorder characterized by stimulus-sensitive myoclonus, tonic-clonic epileptic seizures, age of onset at 6-15 years and a progressive course. Mutations in the cystatin B (*CSTB*) gene underlie EPM1. The most common underlying mutation is a dodecamer repeat expansion in the putative promoter region of the gene. In addition, prior to this work, seven other EPM1-associated *CSTB* mutations had been published. CSTB, a cysteine protease inhibitor, is a ubiquitously expressed tightly binding reversible inhibitor of cathepsins B, L, H and S, but its physiological function is unknown. In this study, we characterized the regulation and expression of the *CSTB* gene, determined the subcellular localization of the CSTB protein and elucidated the consequences of patient mutations on expression and subcellular localization of the CSTB gene and protein.

Two novel EPM1-associated *CSTB* mutations, c.149G>A and c.168+1_18del were identified.

Using promoter-luciferase reporter gene constructs, the basal *CSTB* promoter activity was located in the region from -670 to -1 bp from the translation initiation codon. The binding activity of transcription factors to five Sp1 and four AP1 sites as well as weak binding to an androgen response element (ARE) half site was demonstrated.

The mouse *Cstb* transcript was present in all embryonic stages and adult tissues examined, the expression being highest at embryonic day 7. In adult tissues highest expression was detected in thymus. In sectioned E17 embryos highest expression was seen in the thymus, intestine and neopallial cortex. In postnatal brain *Cstb* expression was cell-specific, being highest in the cortex, caudate putamen, thalamus and hippocampus, as well as in the Purkinje cell layer of the cerebellum. Cystatin B expression thus seems to be tightly temporally and spatially regulated.

Several alternatively spliced *CSTB* isoforms were identified, of these *CSTB2* was widely expressed, also in brain, but with very low levels compared to *CSTB* expression. The other splice variants showed more restricted tissue expression.

The endogenous CSTB protein localized to the nucleus, cytoplasm and lysosomes in COS-1 cells and human myoblasts. When myoblasts were differentiated into myotubes CSTB was excluded from the nucleus and lysosomes and merely

localized to the cytoplasm, suggesting that the subcellular distribution of CSTB is dependent on the differentation status of the cells.

Compatible with some 10-fold lowered *in vitro* transcriptional activity of the *CSTB* promoter containing a "premutation" length dodocamer repeat expansion, the *CSTB* expression was clearly decreased in individuals with "premutation" repeat expansions. The expression of *CSTB* mRNA and CSTB protein was markedly reduced in cells of EPM1 patients irrespective of the mutation type with the exception of one missense mutation. The proteins representing patient missense mutations were localized to cytoplasm and nucleus, but failed to associate with lysosomes, implying the importance of the lysosomal association for the proper physiological function of CSTB.

By quantitative real-time PCR the splice variant *CSTB2* was shown to be expressed at very low levels in all tissues. Whether CSTB2 protein is translated *in vivo* and its physiological relevance remain unknown. In patients the level of *CSTB2* expression was reduced similarly to that of *CSTB*. We were unable to detect endogenous CSTB2 protein in cell lines with the antibodies available. The overexpressed wild-type and p.Gly4Arg mutant CSTB2 proteins localized evenly to cytoplasm and nucleus, but not to lysosomes. Instead, the p.Gly50Glu CSTB2 protein aggregated in cytoplasm and nucleus.

The data produced in this study lay the basis for further understanding of the role of *CSTB* in health and disease.

INTRODUCTION

The human genome project (HGP), initiated formally in 1990, published the initial sequence of the 3 billion base pair human genome in 2001 by the publicly sponsored HGP (Lander et al., 2001b) and by the private company, Celera Genomics (Venter et al., 2001). The complete and high-quality sequence was reached in 2003, 50 years after the discovery of the structure of DNA (Watson and Crick, 1953). The most unexpected result of the HGP was that the estimated number of protein-coding genes in the human genome was reduced from an estimated 50000-100000 to only 20000-25000 genes (International Human Genome Sequencing Consortium, 2004), which lead to the realization that much of the diversity seen in higher vertebrates has to be generated by other mechanisms such as alternative splicing, post-translational modifications and multifunctional proteins (Levine and Tjian, 2003; Pennisi, 2005). In addition, the genomes of over one hundred eukaryotic model organisms have been sequenced (http://www.ncbi.nlm.nih.gov/), including the chimpanzee *Pan troglodytes* (Chimpanzee Sequencing and Analysis Consortium, 2005), the domestic dog *Canis familiaris* (Lindblad-Toh et al., 2005), the mouse *Mus musculus* (Waterston et al., 2002), the rat *Rattus norvegicus* (Gibbs et al., 2004), the fruitfly *Drosophila melanogaster* (Adams et al., 2000; Myers et al., 2000), the roundworm *Caenorhabditis elegans* (C. elegans Sequencing Consortium, 1998) and the baker's yeast *Saccharomyces cerevisiae* (Goffeau et al., 1996).

The HGP has had an enormous impact on the identification of genes that are associated with diseases. The Online Mendelian Inheritance in Man (OMIM) database lists more than 2000 of those currently known (http://www.ncbi.nlm.nih.gov/). In Finland, significant progress has been made especially in identifying the so-called Finnish disease heritage genes. This term, introduced by Drs. Nevanlinna, Norio and Perheentupa (Norio et al., 1973), refers to a collection of monogenic diseases that are more prevalent in Finland than in other populations. To date, the Finnish disease heritage is composed of 36 disorders (Norio, 2003), of which the disease gene has been identified in 33, and the disease gene locus is known for an additional one. Most of these diseases have an autosomal recessive mode of inheritance and one founder mutation typically accounts for 70-100% of these disease alleles (http://www.findis.org/).

The progressive myoclonus epilepsy of Unverricht-Lundborg type (EPM1) is one of the diseases belonging to the Finnish disease heritage. The study presented in this thesis is one part of a long-term research project aiming at understanding the molecular pathogenesis of EPM1. Before this study the research group had identified mutations in the underlying gene for EPM1 by positional cloning (Lehesjoki et al., 1991; Lehesjoki et al., 1992; Lehesjoki et al., 1993; Pennacchio et al., 1996; Virtaneva et al., 1996). The major mutation is an expansion of a dodecamer minisatellite repeat in the putative promoter region of the *CSTB* gene (Lalioti et al., 1997b). This mutation accounts for over 90% of the disease alleles (Virtaneva et al., 1997) and is unique in that it represented the first case of instability of a repeat unit other than trinucleotides associated with a human disease. The CSTB protein has been identified and much studied in Finland (Rinne, 1981; Järvinen and Rinne, 1982; Rinne et al., 2002; Laitala-Leinonen et al., 2006). CSTB, a cysteine protease inhibitor, is widely expressed (Pennacchio et al., 1996) and binds *in vitro* to cathepsins B, H, L and S (Green et al., 1984). This study was performed to investigate *CSTB* gene expression and CSTB protein functions in normal and pathological conditions.

REVIEW OF THE LITERATURE

1. Epilepsy and genetics of epilepsy

Human epilepsy is one of the most common neurological conditions with an active prevalence of 1% and a cumulative incidence by the age of 80 years of about 3% (Bate and Gardiner, 1999). Around 60 million people worldwide are affected and 45000 individuals in Finland are on anti-epileptic medication. The term "epilepsy" actually describes a very heterogeneous group of disorders characterized by predisposition to seizures that have different causes, ages of onset, clinical presentations and outcomes. The seizures themselves are the clinical outcome of cortical dysfunction due to a sudden, abnormal, excessive, and disorganized discharge of brain cells. Clinical manifestations include abnormal motor, sensory, cognitive, psychic and/or autonomic phenomena. A diagnosis of epilepsy requires at least two unprovoked seizures (Steinlein, 2004). The diagnostic scheme is divided into five parts, ictal phenomenology, seizure type, syndrome, etiology and impairment, to facilitate the clinical approach (Engel, 2001). Self-limited seizure types can be divided into generalized and focal seizures. Generalized seizures can be subdivided into e.g. tonic-clonic, clonic, myoclonic absence, tonic and myoclonic seizures. Focal seizures are divided into e.g. focal sensory and motor seizures (http://ilae-epilepsy.org/). The etiology of epilepsy can be symptomatic or idiopathic. Symptomatic epilepsies are caused by e.g. strokes, head injuries, developmental abnormalities, infections, presence of tumors or degenerative conditions. In idiopathic epilepsies no obvious reason or any other neurological abnormalities are found, and the origin is mainly genetic (Steinlein, 2004). All possible modes of inheritance are found in epilepsies, including autosomal, X-chromosomal, mitochondrial, and simple/Mendelian and complex inheritance. A genetic component has been estimated to be present in 40% of all epilepsies (Annegers et al., 1996). The majority of inherited epilepsies are believed to be multifactorial, with a contribution of several susceptibility genes and environmental factors (Mulley et al., 2005). Of all inherited epilepsies only a few have a known causative gene. These rare-monogenic epilepsies involve mutations in genes which lead, for example, to abnormal ion-channel conductance, deficits in energy or protein metabolism, or defects in cell migration.

2. Progressive myoclonus epilepsies

The progressive myoclonus epilepsies (PMEs) belong to a group of symptomatic epilepsies and are a heterogeneous group of inherited disorders characterized by myoclonic and tonic-clonic seizures and progressive neurological deterioration (Berkovic et al., 1986; Marseille Consensus Group, 1990).

PMEs are divided into five main diseases or disease groups: Unverricht-Lundborg disease (ULD/EPM1), Lafora disease (LD/EPM2), neuronal ceroid lipofuscinoses (NCLs), sialidoses, and myoclonic epilepsy with ragged-red fibers (MERRF) (Marseille Consensus Group, 1990). In all of these genetic mutations have been identified and account for the majority of PME cases. A summary of these disease genes is presented in Table 1. Four of the five disease entities have an autosomal recessive mode of inheritance while MERRF is maternally inherited in mitochondrial DNA. In addition, a number of rare disorders can cause the PME phenotype, for example the autosomal dominantly inherited dentatorubral-pallidoluysian atrophy (DRPLA), most prevalent in Japan, characterized by ataxia, myoclonus, choreoathetosis, epilepsy and dementia (Smith et al., 1958; Naito and Oyanagi, 1982), and is caused by an expanded trinucleotide CAG repeat in the *DRPLA* gene (Koide et al., 1994). Gaucher disease type III is characterized by ataxia, spastic paraplegia, grand mal and/or psychomotor seizures, supranuclear ophthalmoplegia and dementia, and is caused by mutations in the gene encoding acid beta-glucosidase (GBA) (Tsuji et al., 1987).

Lafora disease shows onset of grand mal seizures and/or myoclonus around the age of 15 years. Rapid and severe mental deterioration develops, often with psychotic symptoms, culminating in continuous seizures and death less than 10 years after onset. LD is histologically characterized with starch-like polyglucosan masses known as Lafora bodies (Lafora and Gluck, 1911; Chan et al., 2005). LD has the same kind of clinical manifestation as ULD, at least during the first years after onset, and before the demarcation of those two disorders on clinical grounds by Norio and Koskiniemi in 1979 there was confusion in their classification. The names of Unverricht, Lundborg and Lafora were coupled and separated in every possible way. Since in histological studies no Lafora bodies or Lafora-like deposits were found in Finnish patients or Swedish patients with Finnish ancestry (Haltia et al., 1969; Koskiniemi et al., 1974a), Norio and Koskiniemi proposed separate classifications for

Unverricht-Lundborg and Lafora diseases, as well as distinguished the clinical phenotypes from each other (Norio and Koskiniemi, 1979). Lafora disease can be caused by mutations in genes encoding either laforin (*EPM2A*) (Minassian et al., 1998; Serratosa et al., 1999), a protein composed of a carbohydrate binding domain and a dual-specific phosphatase domain (Ganesh et al., 2000; Minassian et al., 2000; Minassian et al., 2001; Chan et al., 2004a) or malin (*NHLRC1* [NHL repeat containing 1] or *EPM2B*) (Chan et al., 2003), a putative E3 ubiquitin ligase (Chan et al., 2003). To date, a total of 43 different *EPM2A* and 40 *EPM2B* associated mutations have been identified (http://projects.tcag.ca/lafora/; Ianzano et al., 2005). A third gene locus for LD is speculated to exist (Chan et al., 2004b). Two *Epm2a* mouse models have been created (Ganesh et al., 2002; Chan et al., 2004a). Interestingly, in the canine *Epm2b* gene an expanded dodecamer repeat mutation (Lohi et al., 2005) causes naturally occuring LD. The normal sequence of the 5' half of the *Epm2b* gene's single exon consists of two consecutive identical dodecamers and a third copy differing by a single nucleotide. The bi-allelic expansion mutation contains 19-26 copies of the dodecamer repeat and causes more than 900 times reduced mRNA levels of *Epm2b* in the skeletal muscle of affected domestic dogs (Lohi et al., 2005).

The neuronal ceroid lipofuscinoses are clinically characterized by progressive mental and motor retardation, epileptic seizures, myoclonus, visual failure and premature death (Santavuori, 1988; Haltia, 2003). Histopathologically the diseases are characterized by the accumulation of autofluorescent storage material in several cell types, especially in neurons (Santavuori, 1988). The NCL disorders are divided depending on the age of onset, the rate of progression, the morphology of the storage material and the genetic basis into ten subgroups: infantile NCL (INCL/CLN1), classical late-infantile NCL (cLINCL/CLN2), juvenile NCL (JNCL/CLN3), adult NCL (CLN4), Finnish variant late-infantile NCL ($vLINCL_{Fin}$/CLN5), variant late-infantile NCL (vLINCL/CLN6), Turkish variant late-infantile NCL ($vLINCL_{Turk}$/CLN7), Northern epilepsy (EPMR/CLN8) (Haltia, 2003), variant juvenile NCL (CLN9) (Schulz et al., 2004) and congenital NCL (CLN10) (Siintola et al., 2006). Seven causative genes have been identified, four of them encode soluble lysosomal proteins and three transmembrane proteins. Genes encoding soluble lysosomal proteins are palmitoyl protein thioesterase 1 (*PPT1*/*CLN1*) (Järvelä et al., 1991; Vesa et al., 1995), tripeptidyl-peptidase I (*TPPI*/*CLN2*) (Sharp et al., 1997;

Sleat et al., 1997), *CLN5* (Klockars et al., 1996; Savukoski et al., 1998) and cathepsin D (*CTSD*/*CLN10*) (Siintola et al., 2006; Steinfeld et al., 2006). Genes encoding transmembrane proteins are *CLN3* (Gardiner et al., 1990; International Batten Disease Consortium, 1995), *CLN6* (Sharp et al., 1997; Gao et al., 2002; Wheeler et al., 2002) and *CLN8* (Haltia et al., 1969; Ranta et al., 1999). *CLN9* is a putative gene encoding a protein with proposed function as a regulator of dihydroceramide synthase (Schulz et al., 2004; Schulz et al., 2006). In Turkish vLINCL patients mutations in the *CLN8* (Ranta et al., 2004) and *CLN6* (Siintola et al., 2005) genes have been identified, though the "true" $vLINCL_{Turk}$ gene, *CLN7*, remains to be defined (Siintola et al., 2005). Recently mutations have been identified in the *CTSD* gene causing both congenital NCL (CLN10) (Siintola et al., 2006) and vLINCL (Steinfeld et al., 2006). The NCL Mutation Database summarises approximately 160 mutations causing NCL (http://www.ucl.ac.uk/ncl/). Targeted disruption of murine cathepsin F (Tang et al., 2006), and chloride channels ClC-3 (Stobrawa et al., 2001; Yoshikawa et al., 2002), ClC-6 (Poet et al., 2006) and ClC-7 (Kasper et al., 2005) have lead to NCL-like phenotypes. Several natural NCL animal models exist, for example the CLN2 dog (Awano et al., 2006a), CLN6 mouse (Gao et al., 2002; Wheeler et al., 2002), CLN8 mouse (Ranta et al., 1999) and dog (Katz et al., 2005), and CTSD sheep (Tyynelä et al., 2000) and dog (Awano et al., 2006b). In addition, several NCL mouse models have been created (Saftig et al., 1995; Bronson et al., 1998; Katz et al., 1999; Mitchison et al., 1999; Gupta et al., 2001; Cotman et al., 2002; Kopra et al., 2004; Sleat et al., 2004; Jalanko et al., 2005).

Sialidoses, characterized by the progressive lysosomal storage of sialidated glycopeptides and oligosaccharides, are caused by mutations in the gene encoding neuraminidase 1 (NEU1) (Bonten et al., 1996; Seyrantepe et al., 2003). Mutations in *NEU1* can cause both mild (type I) and severe (type II) disease phenotypes, depending on the residual activity of the mutated enzyme. Type I sialidosis occurs in the second decade of life and results in progressive loss of vision associated with nystagmus, ataxia, myoclonus and grand mal seizures. Type II sialidosis is characterized by the presence of abnormal somatic features and is divided into congenital, infantile and juvenile subtypes. The sialidosis mouse model has been generated by targeted disruption of the *Neu1* locus (de Geest et al., 2002).

Myoclonic epilepsy associated with ragged-red fibers is a maternally inherited disease characterized by myoclonic epilepsy and ragged-red fibres in the muscle. In addition, myoclonus, dementia, cardiomyopathy, pyramidal tract signs, neuropathy, optic nerve atrophy and neurosensory hearing loss are often present (Fukuhara et al., 1980; Rosing et al., 1985). The mutations in mitochondrial genes $tRNA^{Lys}$ (*TRNK*) (Shoffner et al., 1990) and $tRNA^{Phe}$ (*TRNF*) (Mancuso et al., 2004) affect the translation of all mtDNA-encoded genes and produce multiple deficiencies in the enzyme complexes of the respiratory chain, most prominently complex I (NADH-CoQ reductase) and complex IV (cytochrome-c oxidase) (Wallace et al., 1988; Bindoff et al., 1991).

In Lafora disease, NCLs and sialidoses the cellular mechanism of pathogenesis appear to be similar, with progressive formation of intracellular inclusion bodies (Harriman et al., 1955; Zeman and Albert, 1963; Carpenter et al., 1974; O'Brien, 1977). In MERRF abnormal mitochondrial proliferation, cytochrome-c oxidase (COX)-negative fibers and ragged-red fibers are detected (Nakamura et al., 1995; Mancuso et al., 2004; Melone et al., 2004). In contrast, EPM1 is not associated with inclusion bodies, storage material or mitochondrial defects (Lehesjoki and Koskiniemi, 1998; Pennacchio et al., 1998).

Table 1. The main human PME genes.

DISEASE	LOCUS	CHROMOSOMAL LOCALIZATION	GENE	REFERENCE
Unverricht-Lundborg disease	EPM1	21q22.3	*CSTB*	(Pennacchio et al., 1996)
	EPM1B	12p11-q13	?	(Berkovic et al., 2005)
Lafora disease	EPM2A	6q24	*EPM2A*	(Minassian et al., 1998; Serratosa et al., 1999)
	EPM2B	6p22.3	*EPM2B*	(Chan et al., 2003)
	EPM2C	?	?	
Neuronal ceroid lipofuscinoses	CLN1	1p32	*PPT1*	(Vesa et al., 1995)
	CLN2	11p15	*TPP1*	(Sleat et al., 1997)
	CLN3	16p12.1	*CLN3*	(IBDC*, 1995)
	CLN4 (putative)	?	?	
	CLN5	13q22.1-q32	*CLN5*	(Savukoski et al., 1998)
	CLN6	15q23	*CLN6*	(Gao et al., 2002)
	CLN7	?	?	
	CLN8	8p23	*CLN8*	(Ranta et al., 1999)
	CLN9	?	?	
	CLN10	11p15.5	*CTSD*	(Siintola et al., 2006; Steinfeld et al., 2006)
Sialidoses type I and II	NEU1	6p21.3	*NEU1*	(Bonten et al., 1996)
Myoclonic epilepsy with ragged-red fibers	TRNK	mtDNA	*TRNK*	(Shoffner et al., 1990)
	TRNF	mtDNA	*TRNF*	(Mancuso et al., 2004)

*International Batten Disease Consortium.

3. Progressive myoclonus epilepsy of Unverricht-Lundborg type (EPM1)

Unverricht-Lundborg disease was first described by Heinrich Unverricht in Estonia (Unverricht, 1891; Unverricht, 1895) and by Herman Lundborg in Sweden (Lundborg, 1903; Lundborg, 1912; Lundborg, 1913). Progressive myoclonus epilepsy of Unverricht-Lundborg type (EPM1; OMIM #254800; also known as Baltic or Mediterranean myoclonus epilepsy) is the single most common cause of progressive myoclonus epilepsy worldwide (Marseille Consensus Group, 1990). EPM1 is most prevalent in Finland and it belongs to the Finnish Disease Heritage (Norio, 2003). EPM1 is autosomal recessively inherited and the incidence in Finland is 1 in 20000 births (Norio and Koskiniemi, 1979; Eldridge et al., 1983; http://www.findis.org/), meaning around three to four new patient diagnoses per year. At present there are approximately 200 patients known in Finland (Lehesjoki and Koskiniemi, 1999; http://www.findis.org/). EPM1 is also prevalent in the Mediterranean region (Marseille Consensus Group, 1990; Malafosse et al., 1992; Lehesjoki et al., 1994; Parmeggiani et al., 1997) and is also found around the world (Eldridge et al., 1983; Marseille Consensus Group, 1990).

3. 1. Clinical symptoms

The onset of symptoms occurs between 6-15 years of age with myoclonus and/or generalized tonic-clonic epileptic seizures (Koskiniemi et al., 1974a). Stimulus-sensitive myoclonus is an essential feature for the diagnosis and the first symptom in approximately half of the patients (Norio and Koskiniemi, 1979; Koskiniemi, 1986; Koskiniemi, 1987). Myoclonus (brief contraction of a muscle or group of muscles) occurs spontaneously, or in association with action or external stimuli such as light, noise or touch (Berkovic et al., 1986; Marseille Consensus Group, 1990). Myoclonus may be focal or multifocal and may generalize to a shaking attack, to a tonic-clonic seizure attack and unconsciousness. The generalized tonic-clonic epileptic seizures are most frequent 3-7 years after the disease onset and may later, with anticonvulsive medication, cease entirely (Koskiniemi et al., 1974a; Norio and Koskiniemi, 1979; Koskiniemi, 1986). The clinical course of the disease progresses slowly, with ataxia, intentional tremor, incoordination and dysarthria manifesting later. No major intellectual decline is present, but a slow decrease in the IQ of about 10 points in a decade has been reported (Koskiniemi, 1974; Koskiniemi et al., 1974a; Koskiniemi,

1986). In addition, depression and emotional lability are often observed (Lehesjoki and Koskiniemi, 1999). Electroencephalogram (EEG) is abnormal even before the symptoms appear (Koskiniemi et al., 1974b; Lehesjoki and Koskiniemi, 1998) with typical spike-wave or polyspike-wave discharges (Koskiniemi et al., 1974b; Koskiniemi, 1986). Sodium valproate alone or combined with clonazepam or piracetam is the most effective therapy (Iivanainen and Himberg, 1982; Somerville and Olanow, 1982; Koskiniemi, 1986; Koskiniemi et al., 1998), whereas the previously used phenytoin is considered harmful due to toxic side effects (Eldridge et al., 1983). At present due to better medication and improvement in the whole therapy regimen the life span of EPM1 patients is probably normal (Lehesjoki, 2002).

3. 2. Histopathology

On histopathological examinations of the brain, widespread non-specific degenerative changes and loss of Purkinje cells in the cerebellum have been observed (Haltia et al., 1969; Koskiniemi et al., 1974a; Eldridge et al., 1983). Cerebellar atrophy and gliosis are present as well. Neuronal degenerative changes have been observed for example in cerebral cortex, striatum, medial and reticular thalamus, brainstem nuclei and spinal cord. In addition, proliferation of Bergmann glia has been detected. However, some of these changes detected in the patients may have resulted from treatment with phenytoin. Furthermore, no intracellular inclusion bodies or storage material have been observed (Haltia et al., 1969; Koskiniemi et al., 1974a; Eldridge et al., 1983).

3. 3. Molecular genetics

The mutated gene responsible for EPM1 was identified using a positional cloning approach. A genome-wide search for linkage localized the EPM1 gene to chromosome 21q22.3 (Lehesjoki et al., 1991). By linkage disequilibrium analysis and exploiting the historical recombinations in haplotype analyses the localization was narrowed to an approximately 175 kb region (Virtaneva et al., 1996). A systematic search for genes in this region resulted in identification of a gene encoding a previously described but unmapped protein, cystatin B (CSTB), a cysteine protease inhibitor (Pennacchio et al., 1996).

The *CSTB* gene is ubiquitously expressed with a transcript of approximately 0.8 kb in Northern blot analysis (Pennacchio et al., 1996). *CSTB* spans over a 2.5 kb genomic region, contains three small exons with coding sequences of 66 bp, 102 bp

and 129 bp (Fig. 1) (Pennacchio et al., 1996), and two transcription start sites -97 bp and -108 bp upstream from the translation initiation site (Lalioti et al., 1997b). The 3'untranslated region (UTR) is approximately 250 bp, and the introns 1422 bp and 326 bp in length (Pennacchio et al., 1996). Altogether ten EPM1-associated mutations (Fig. 1; Table 2, section 3.3.2) have been identified in the *CSTB* gene (Pennacchio et al., 1996; Lalioti et al., 1997a; Lalioti et al., 1997b; Kagitani-Shimono et al., 2002; de Haan et al., 2004; IV).

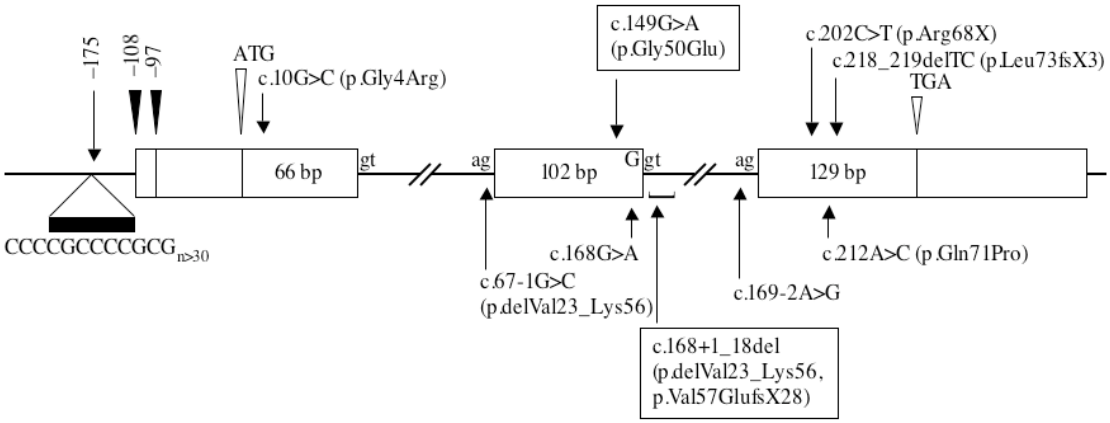


Figure 1. *CSTB* gene structure and EPM1-associated mutations. The *CSTB* gene contains three exons shown as white boxes with respective lengths of the coding sequences. The translation start (ATG) and stop (TGA) sites are indicated with white arrowheads. 5' and 3'UTRs are shown upstream off exon 1 and downstream from exon 3, respectively. The positions of the two transcription start sites, -108 bp and -97 bp from the translation initiation site, are indicated with black arrowheads. The positions of the EPM1-associated mutations are shown with arrows. The major mutation underlying EPM1, the dodecamer repeat expansion, is located in the putative promoter region -175 bp upstream off the translation initiation site. The novel mutations identified in study IV are framed. Picture modified from Lehesjoki (2003).

Molecular testing for EPM1 in the Helsinki University Central Hospital, Laboratory of Molecular Genetics, is available for the repeat expansion, and the c.10G>C, c.67-1G>C, c.169-2A>G, c.202C>T and c.218_219delTC mutations (http://www.hus.fi/). Gene testing for EPM1 is also performed in five other DNA diagnostic laboratories in Europe (http://www.findis.org/; http://www.eddnal.com/).

3.3.1. The dodecamer repeat expansion mutation

An unstable dodecamer (12 bp) 5'-CCCCGCCCCGCG-3' repeat expansion (Lalioti et al., 1997b; Virtaneva et al., 1997), the underlying mutation in the majority of EPM1 chromosomes, is detected worldwide in about 90% and in Finland in about 98% of disease alleles (Lafrenière et al., 1997; Lalioti et al., 1997b; Virtaneva et al., 1997). This normally polymorphic minisatellite repeat, located in the putative promoter of *CSTB* -175 bp upstream off the translation initiation codon (Fig. 1), occurs in 2 to 3

copies in normal chromosomes, whereas the repeat copy number in expanded alleles varies from 30 to 125, which translates to sizes from 360 to 1500 bp (Lalioti et al., 1997b; Virtaneva et al., 1997; Lalioti et al., 1998).

The diagnostic testing of the expansion mutation has been extremely difficult due to the minisatellite repeat consisting exclusively of G and C nucleotides. The detection has been largely based on Southern blot analysis of genomic DNAs digested with *Pst*I (2.6 kb fragment in normal chromosomes), *Eco*RI (9 kb) or *Hind*III (5.6 kb) restriction enzymes and hybridized with *CSTB* cDNA to detect enlarged fragments (Lalioti et al., 1997b; Virtaneva et al., 1997). In addition, several methods for the PCR amplification of the repeat expansion under special conditions have been generated: PCR amplification followed by hybridization with the oligonucleotide probe containing the repeat (Lalioti et al., 1998), fluorescent PCR genotyping using an automated DNA sequencer (Larson et al., 1999), and PCR with deaminated DNAs as templates (Weinhaeusel et al., 2003; Horiuchi et al., 2005).

The majority of EPM1 patients are homozygous for the repeat expansion (Lafrenière et al., 1997; Lalioti et al., 1997b; Virtaneva et al., 1997; Moulard et al., 2002). The origin of the expansion mutation has been investigated in at least three independent studies characterizing the EPM1 haplotype using microsatellite markers. The outcome of these studies has been controversial: suggesting only one, a small number or several origins for this mutation (Lafrenière et al., 1997; Virtaneva et al., 1997; Moulard et al., 2002). The definitive solution to this issue could be reached with haplotyping more stable SNP (single nucleotide polymorphism) markers in large and multinational patient material.

The so-called “premutation” alleles have been identified in the Venezuelan CEPH (Centre d’Etude du Polymorphisme Humain) families having Huntington’s disease, but in whom no signs or symptoms of EPM1 have been reported. In these families the repeat copy number ranges from normal to 17 (Lalioti et al., 1997b).

The normal range of repeats have been shown to form stable secondary structures that may facilitate strand slippage, thereby causing the expansion (Pataskar et al., 2001a; Pataskar et al., 2001b). The disease range repeats have been demonstrated to form tetraplex structures (Fig. 2), which might affect DNA replication and repair, and/or have an effect on local chromatin structure (Saha and Usdin, 2001).

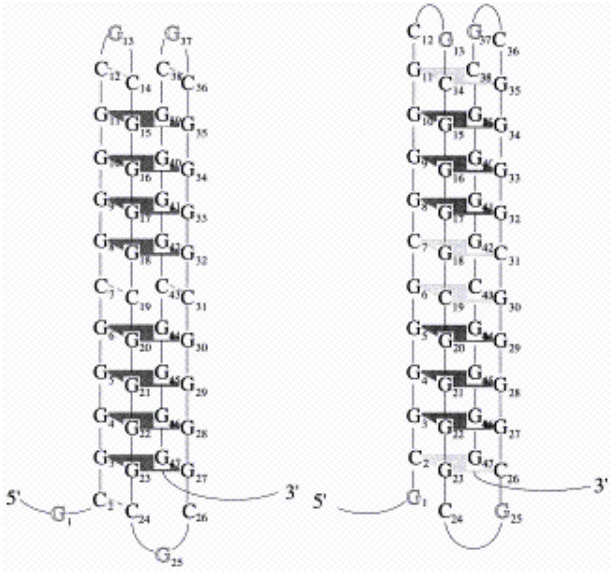


Figure 2. Two different models of the tetraplex structures formed by the bottom strand of the dodecamer repeat. Picture from Saha and Usdin (2001).

The meiotic instability of the repeat during transmission has been demonstrated in the EPM1 family pedigrees (Virtaneva et al., 1997; Lalioti et al., 1998; Larson et al., 1999). The mutation rate (47%) of the EPM1 expansion alleles is the highest yet observed for any human minisatellite (Larson et al., 1999). Both contraction and expansion of the repeats have been recognized, typically by a single repeat unit (Larson et al., 1999). The instability has been detected in both maternal and paternal transmissions, being greater in paternal ones (Larson et al., 1999), similarly to that observed with CAG repeats (Ikeuchi et al., 1996). This high mutation rate might be mediated by DNA replication errors due to the 100% GC content of considerable length (Larson et al., 1999).

In Northern analyses of mRNA extracted from lymphoblastoid cells, dramatically reduced mRNA levels were observed in EPM1 patients with homozygous repeat expansions. This suggests reduced *CSTB* gene transcription due to the expansion in the promoter region (Pennacchio et al., 1996; Bespalova et al., 1997a; Lafrenière et al., 1997; Lalioti et al., 1997b). Decreased cathepsin inhibitory activity of the CSTB protein in these cells has been demonstrated as well (Rinne et al., 2002). On the contrary, Lalioti et al. (1997a) detected normal *CSTB* expression levels with RNase protection assays of patient lymphoblastoid and fibroblast cell lines, while in blood leukocytes *CSTB* levels were markedly reduced, suggesting a cell-specific regulation of *CSTB* gene expression.

The hypermethylation status of the promoter and surrounding CpG-elements has been investigated by Southern blot analysis with methylation sensitive *Hpa*II and *Msp*I restriction enzymes (Lalioti et al., 1997b). In blood leukocytes and lymphoblastoid cell lines of controls, carriers and patients no obvious differences in DNA methylation was revealed, thus reduction of *CSTB* expression is presumably not associated with methylation differences (Lalioti et al., 1997b). However, the methylation status of the repeat itself has not yet been determined.

No correlation between the expansion size and the age of onset or clinical severity has been observed (Lafrenière et al., 1997; Lalioti et al., 1997b; Virtaneva et al., 1997; Lalioti et al., 1998), suggesting that once the size expands beyond a critical threshold, sufficiently reducing *Cstb* gene expression, the pathological consequences are fully developed. The differences in the disease manifestation even among siblings is likely to be due to the influence of modifier genes and/or environmental factors (Lalioti et al., 1998).

3.3.2. Other EPM1-associated mutations

The so-called minor EPM1 mutations together account for a small proportion, less than 10%, of disease alleles worldwide. In compound heterozygous form with the repeat expansion occur six published mutations (Table 2): the nonsense c.202C>T (p.Arg68X) (Pennacchio et al., 1996), missense c.212A>C (p.Gln71Pro) (de Haan et al., 2004), frameshift c.218_219delTC (p.Leu73fsX3) (Lalioti et al., 1997a), and three splice site mutations, c.67-1G>C (p.delVal23_Lys56) (Pennacchio et al., 1996), c.168G>A (Kagitani-Shimono et al., 2002) and c.169-2A>G (Lalioti et al., 1997a). In addition, one EPM1 patient has the missense c.10G>C (p.Gly4Arg) mutation, the only minor mutation identified in homozygous form (Lalioti et al., 1997a).

The splice site c.67-1G>C mutation, detected in several mostly caucasian EPM1 patients (Pennacchio et al., 1996; Bespalova et al., 1997b; Lafrenière et al., 1997; Lalioti et al., 1997a), results in skipping of exon 2 with consequent in-frame deletion of 34 amino acids (p.delVal23_Lys56) (Bespalova et al., 1997b). This mutation is the most common of the minor mutations as it has been detected in several disease chromosomes, whereas the other minor mutations have each been described in only one or a few chromosomes (Lehesjoki and Koskiniemi, 1999). The exon-intron boundary c.168G>A mutation, identified in a Japanese patient, causes no amino acid change, but affects a conserved splice site sequence predicting severe splicing defects

that are, however, unconfirmed (Kagitani-Shimono et al., 2002). The splice site c.169-2A>G mutation, found in a French patient, probably results in abnormal splicing, but this has not been verified (Lalioti et al., 1997a).

Table 2. EPM1-associated mutations in the *CSTB* gene.

MUTATION	CONSEQUENCES	POPULATION	REFERENCE
Dodecamer repeat expansion	Reduced mRNA transcription	Several populations	(Lafrenière et al., 1997; Lalioti et al., 1997b; Virtaneva et al., 1997)
c.10G>C	p.Gly4Arg	Moroccan	(Lalioti et al., 1997a)
c.67-1G>C	p.delVal23_Lys56	Several populations	(Pennacchio et al., 1996; Bespalova et al., 1997b; Lafrenière et al., 1997; Lalioti et al., 1997a)
c.149G>A	p.Gly50Glu	Finnish	IV
c.168G>A	Aberrant splicing?	Japanese	(Kagitani-Shimono et al., 2002)
c.168+1_18del	p.delVal23_Lys56 p.Val57GlufsX28	Italian	IV
c.169-2A>G	Aberrant splicing?	French	(Lalioti et al., 1997a)
c.202C>T	p.Arg68X	Finnish Swedish Dutch	(Pennacchio et al., 1996; Lafrenière et al., 1997; de Haan et al., 2004)
c.212A>C	p.Gln71Pro	Dutch	(de Haan et al., 2004)
c.218_219delTC	p.Leu73fsX3	French American	(Bespalova et al., 1997a; Lafrenière et al., 1997; Lalioti et al., 1997a)

The missense c.10G>C mutation identified in a Moroccan patient (Lalioti et al., 1997a) results in the substitution of a highly conserved glycine at amino acid position four (p.Gly4Arg). On the basis of the known crystal structure of CSTB with papain (Fig. 3, section 4.1), the glycine residue is part of the papain-binding pocket (Stubbs et al., 1990). The substitution of this glycine by an arginine, with a long and charged side chain, causes major steric hindrance that is likely to strongly affect the papain binding of CSTB (Lalioti et al., 1997a). The other published missense mutation c.212A>C, found in a Dutch patient, results in substitution of a glutamine at position 71 by a proline (p.Gln71Pro) (de Haan et al., 2004). This mutation site is located proximal to the second hairpin loop that contributes indirectly to protease binding (Stubbs et al., 1990; Pol and Björk, 1999), and the mutation is likely to affect the structure of the fourth beta strand within the second hairpin loop by two residues, most likely resulting in altered binding affinities.

The nonsense c.202C>T mutation causing an early termination codon and predicting a truncated p.Arg68X CSTB protein has been identified in Finnish, Swedish and Dutch patients (Pennacchio et al., 1996; Lafrenière et al., 1997; de Haan

et al., 2004). The c.218_219delTC mutation causes the deletion of TC in a short dinucleotide repeat AATCTCTCCC, deleting amino acid Leu73 and creating a frameshift with two out-of-frame amino acids (proline-serine) before an early stop codon (p.Leu73fsX3). A truncated protein only 74 amino acids in length, 24 amino acid shorter than the normal 98 amino acid CSTB protein, is predicted to be produced (Lalioti et al., 1997a). This mutation has been detected in French and American patients (Bespalova et al., 1997a; Lafrenière et al., 1997; Lalioti et al., 1997a).

3.3.3. EPM1B, a new locus for Unverricht-Lundborg disease

A systematic study of inherited epilepsies performed in Israel identified mutations in the *CSTB* gene in an inbred Arab family (Mazarib et al., 2001), but in another Arab family living in a nearby town and suffering from similar symptoms mutations were, unexpectedly, not detected (Berkovic et al., 2005). This family contained eight affected individuals with progressive myoclonus and ataxia, seven of whom also had tonic-clonic epileptic seizures. Magnetic resonance imaging (MRI), muscle and skin biopsies were unremarkable. The disease phenotype was typical of Unverricht-Lundborg disease, but the mean age of onset was earlier than in EPM1 (7.3 years versus 11 years for EPM1) (Berkovic et al., 2005). With a genome-wide screen a new locus for EPM1, named EPM1B, was mapped to chromosome 12 (Berkovic et al., 2005). This 15-megabase region on 12p11-q13 with a maximum LOD (logarithm of odds) score of 6.32 contains 47 known genes and numerous predicted genes, and does not contain any known PME genes or genes known to be related to *CSTB* (Berkovic et al., 2005). Identification of this new *EPM1B* gene may contribute to understanding the molecular pathogenesis of ULD.

4. Cystatins

Cystatins are inhibitors of papain family cysteine proteases, cathepsins, which they bind tightly and reversibly (Abrahamson et al., 1986; Brömme et al., 1991; Machleidt et al., 1991; Lenarcic et al., 1996). Cystatins are thought to participate in the protection against unwanted proteolysis, through controlling the activities of endogenous lysosomal proteases by rapidly trapping these enzymes in tight complexes (Barrett et al., 1986b). They also protect against proteases released from metastasizing cancer cells (Sloane, 1990) and harmful effects of free cysteine proteases released from invading microorganisms (North et al., 1990; Stoka et al.,

1995). Chicken egg white cystatin is the prototype of the group (Fossum and Whitaker, 1968). All cystatins contain a conserved N-terminal Gly amino acid residue and a conserved polypeptide sequence of Gln-Val-Val-Ala-Gly (QVVAG) (Barrett et al., 1986b; Barrett, 1987). Cystatins are divided into three groups on the basis of functional and structural similarities (Barrett et al., 1986a; Rawlings and Barrett, 1990; http://merops.sanger.ac.uk/). Type 1 cystatins consist of intracellular cystatins A and B (CSTA and CSTB) that are small molecular weight, 98 amino acid, 12 kDa proteins with no disulfide bonds or carbohydrate side chains. Cystatins C (CSTC), D, E/M, F, S, SA and SN, the type 2 cystatins, are larger in size (approximately 120 residues, 13 kDa), and contain two disulfide bridges and a signal sequence for extracellular targeting, but still no carbohydrate side chains. The plasma L- and H-kininogens, type 3 cystatins, are large (50-114 kDa) multifunctional glycoproteins. They consist of three divergent copies of type 2 like cystatin sequences, and contain several disulfide bonds.

CSTA inhibits lysosomal cathepsins B, C, H, L and S, as well as cysteine proteases from parasites e.g. cruzipain, and some plant proteases e.g. papain and actinidin (Green et al., 1984; Abrahamson et al., 1986; Pol et al., 1995; Stoka et al., 1995; Leonardi et al., 1996). The three-dimensional structure of CSTA has been determined and similarity between free CSTA and CSTB in complex with papain is considerable, but there is also conformational differencies in their protease binding regions (Martin et al., 1995). The expression pattern of CSTA is quite restricted, with high levels in skin, where it takes part in the epidermal barrier formation (Nemes and Steinert, 1999), and in some blood cells. CSTA is not associated with any disease phenotype and is little studied.

The mature CSTC protein is composed of 120 amino acid residues, but is synthesized as a preprotein with a 26 amino acid signal peptide (Barrett et al., 1984; Abrahamson et al., 1987; Merz et al., 1997). CSTC is a strong inhibitor of cathepsins B, H, L and S, and is distributed in most human tissues and fluids (Löfberg and Grubb, 1979; Grubb et al., 1983; Abrahamson et al., 1986; Barrett, 1986) with particular abundance in the cerebrospinal fluid (Abrahamson et al., 1986). The CSTC protein has amyloidogenic properties and colocalizes with beta-amyloid in degenerated neurons in Alzheimer's disease, suggesting a role in neuronal degeneration (Deng et al., 2001; Levy et al., 2001). Defects in the CSTC protein cause

hereditary cerebral hemorrhage with amyloidosis (HCHWA), which is characterized by amyloid fibril deposits in the cerebral arteries leading to massive brain hemorrhage and death in young adults (Ghiso et al., 1986; Palsdottir et al., 1988). In *Cstc*-deficient mice no gross pathological abnormalities have been detected (Huh et al., 1999; Olsson et al., 2004).

4. 1. The cystatin B (CSTB) protein

CSTB, also known as Neutral Cysteine Protease Inhibitor (NCPI) and stefin B, was first found and characterized in lymphatic tissue (Rinne, 1981; Järvinen and Rinne, 1982) and liver (Ritonja et al., 1985). CSTB has isoelectric variants (Rinne, 1981; Järvinen and Rinne, 1982) and often forms dimers, which are stabilized by noncovalent forces. It is found mainly intracellularly and is widely distributed among different cell types and tissues. CSTB binds *in vitro* tightly and reversibly to cathepsins H, L and S, and less to cathepsin B (Green et al., 1984; Abrahamson et al., 1986; Barrett, 1986; Brömme et al., 1991; Machleidt et al., 1991; Lenarcic et al., 1996). In addition, CSTB has been shown to inhibit cathepsin K activity in osteoclasts (Laitala-Leinonen et al., 2005).

In EPM1 patient lymphoblastoid cell lines biochemical studies have revealed that the CSTB inhibitory activity is significantly reduced (or absent) and correlates with increased cathepsins B, L and S activity (Rinne et al., 2002), while cathepsin H activity is not significantly altered (Rinne et al., 2002). These findings demonstrate that the activity of these cathepsins is regulated *in vivo* by CSTB. In the rat kindling model of epilepsy, markedly increased Cstb mRNA and protein expression was detected 6 and 24 hours after kindling-induced epileptic seizures in forebrain neurons. This implies that upregulation of *Cstb* following seizures may counteract apoptosis by binding cathepsins, supporting the hypothesis of a neuroprotective role for CSTB (D'Amato et al., 2000).

Cstb expression has been investigated in two large-scale expression analyses of mouse orthologues of the entire transcriptome of human chromosome 21 with emphasis on identifying candidate genes for the pathogenesis of Down syndrome (Gitton et al., 2002; Reymond et al., 2002). Reymond et al. (2002) presented an expression atlas of 161 genes using RNA *in situ* hybridization on whole mount embryos at embryonic day 9.5 (E9.5) and E10.5 and in tissue sections at E14.5, and by RT-PCR in adult tissues. Gitton et al. (2002) studied expression of 158 genes using

RNA *in situ* hybridization at E9.5 embryos and in postnatal day 2 (P2) brain and *in silico* computer mining of expressed sequence tags. *Cstb* was expressed in several embryonal tissues (Reymond et al., 2002) or alternatively in the entire embryos (Gitton et al., 2002). In adult mouse, *Cstb* was expressed in almost all tissues examined (Reymond et al., 2002). In adult rat brain *Cstb* was detected to be differentially expressed in several brain areas with RNA *in situ* hybridization and immunohistochemistry. *Cstb* expression was highest in the hippocampal formation and reticular thalamic nucleus (D'Amato et al., 2000).

On the cellular level CSTB protein has been thought to localize to several intracellular compartments. At first CSTB was considered to be a soluble cytoplasmic protein (Stubbs et al., 1990). Later, using immunofluorescence analysis, CSTB was shown to be concentrated at the membranes of many vesicles in hepatoma cells (Calkins et al., 1998) and also to localize in the nucleus (D'Amato et al., 2000). It was also detected in cytoplasmic granular structures, representing either the Golgi complex or lysosomes, in a variety of cell lines (Riccio et al., 2001).

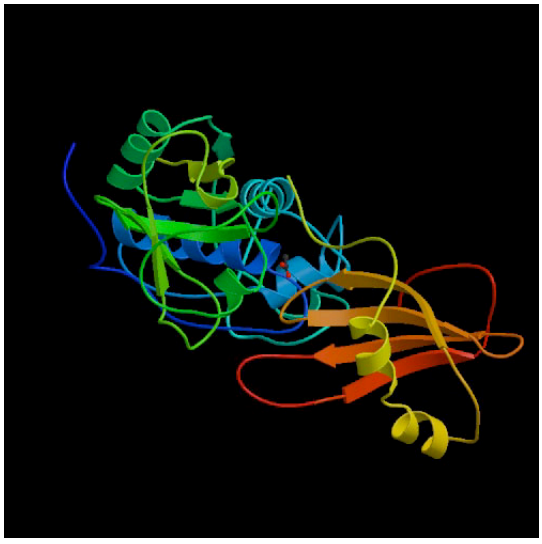


Figure 3. Ribbon model of a complex of CSTB (N-terminal region and the α-helix, shown with yellow color, and the β-sheets with red and orange colors) with papain (shown with green and blue colors). Modified from Stubbs et al. (1990).

The 3D-structure of the recombinant human CSTB-papain complex has been determined (Fig. 3) (Stubbs et al., 1990). CSTB consist of a five-stranded β-sheet wrapped around a five-turn α-helix and a C-terminal strand running along the convex side of the sheet. CSTB interacts with papain through a tripartite wedge formed by conserved residues at the N-terminal end, the first hairpin loop containing the highly

conserved QVVAG sequence and the second hairpin loop. The C-terminus provides an additional interaction region. The interactions are mostly hydrophobic contacts (Stubbs et al., 1990).

CSTB may counteract inappropriate proteolysis by inhibiting cathepsins that leak out into the cytoplasm from the lysosomes, but may also be involved in other cellular prosesses either through cathepsin interactions or other mechanisms. For example, growing evidence for nuclear localization has led to the idea that CSTB might be involved in prosesses other than proteolysis. Therefore, the full physiological function of CSTB is still unknown.

4.1.1. Mammalian model for EPM1: *Cstb*-deficient mouse

A mammalian model for the human EPM1 disease has been produced by targeted disruption of the mouse *Cstb* gene (Pennacchio et al., 1998). The mouse *Cstb* gene encodes a 98 amino acid polypeptide identical in length to the human and rat proteins, with 79% and 86% amino acid similarity, respectively (Pennacchio and Myers, 1996). *Cstb*-deficient mice develop a phenotype that resembles the human disease, with progressive ataxia and myoclonic seizures. However, no tonic-clonic seizures or photosensitivity, typical in human EPM1 patients, have been observed (Pennacchio et al., 1998). Myoclonic seizures occured only during sleep and developed by 1 month of age. While signs of ataxia were detected at 6 months of age (Pennacchio et al., 1998). Analysis of mice revealed apoptotic death of cerebellar granule cells (Pennacchio et al., 1998). Less marked neuronal apoptosis was detected in the hippocampal formation and entorhinal cortex in 3- to 4-month-old mice (Shannon et al., 2002). In older mice, gliosis was present in the hippocampal formation, entorhinal cortex, neocortex and striatum. Widespread gliosis was also present in the white matter, which may be a secondary phenomenon. The superficial neurons of the prosubiculum in the cerebral cortex displayed prominent cellular atrophy (Shannon et al., 2002). Myoclonic seizures were only visible in an isogenic 129Sv background, while ataxia and abnormal brain histology were present in both mixed and isogenic backgrounds (Pennacchio et al., 1998; Shannon et al., 2002).

Transcriptional profiling of *Cstb*-deficient mice brains have shown increased expression of seven genes: cathepsin S (*Ctss*), C1q B-chain of complement (*C1qB*), β2-microglobulin, glial fibrillary acidic protein (*Gfap*), apolipoprotein D, fibronectin 1 and metallothionein II, which are involved in proteolysis, apoptosis and glial

activation (Lieuallen et al., 2001). Ctss was the known target of Cstb, while the other transcripts found encode proteins that are known to be involved in responding to neuronal damage. The increased levels of *Ctss* were a consistent alteration, while cathepsins B, H and L levels were found to vary inconsistently (Lieuallen et al., 2001) Therefore Ctss was suggested to be involved in the apoptosis seen in EPM1. However, no improvement of EPM1 phenotype was seen in a mouse model doubly deficient for *Cstb* and *Ctss* (see below, section 5.2) (Houseweart et al., 2003a).

A hypothesis of whether cathepsins in the absence of the proapoptotic Bcl-2 family member Bid promote apoptosis *in vivo* has been tested in mice doubly deficient for *Cstb* and *Bid* (Houseweart et al., 2003b). *In vitro* studies had suggested that cathepsins can induce apoptosis by cleaving Bid (Stoka et al., 2001). In these mice the phenotypes of apoptosis, ataxia and seizure were not decreased, and cathepsins still promote apoptosis, indicating that apoptosis is mediated via a different pathway or some other molecule can substitute the function of Bid (Houseweart et al., 2003b).

5. Cathepsins

Lysosomal proteases, cathepsins, are involved in intracellular and extracellular protein degradation as well as in a range of cellular processes such as major histocompatibility complex antigen presentation and bone remodelling (Riese and Chapman, 2000). There are two aspartic (cathepsins D and E) and 13 known cysteine (cathepsins B, C, F, H, K, L, N, O, S, T, V, X and W) lysosomal proteases. Cathepsins B, D and L are the most abundant. From now onwards only cathepsins belonging to cysteine proteases are discussed. The cathepsins have acidic pH-optima, with the exception of CTSS which retains its proteolytic activity also at physiological pH (Kirschke et al., 1989; Kirschke and Wiederanders, 1994). Cathepsins B, C, F, H, L, O and X are ubiquitously expressed in mammalian tissues, while CTSK is mainly found in osteoclasts, CTSS mainly in spleen, lymphocytes, monocytes and peripheral antigen-presenting cells, CTST in kidney, spleen and liver, CTSV in thymus and testis, and CTSW in T lymphocytes (Brömme et al., 1996; Linnevers et al., 1997; Santamaria et al., 1998; Turk et al., 2000).

Defects in *CTSK* cause pycnodysostosis, an osteochondrodysplasia characterized by osteopetrosis, short stature due to premature closure of long bone

growth and other severe bone abnormalities (Andren et al., 1962; Maroteaux and Lamy, 1962), suggesting CTSK to be a major protease in bone resorption (Gelb et al., 1996). Interestingly, CSTB has been recently considered to have regulatory functions in osteoclasts, where it could inhibit bone resorption by down-regulating intracellular CTSK activity (Laitala-Leinonen et al., 2005).

5. 1. Cathepsin-deficient mice

The molecular roles of cathepsins have been studied through generating gene targeted mice, e.g. *Ctsb*-, *Ctsk*-, *Ctsl*- and *Ctss*-deficient mice. The *Ctsb*-deficient mice have no spontaneous pathological phenotype, suggesting redundancy of *Ctsb* function (Deussing et al., 1998; Halangk et al., 2000). The *Ctsk*-deficient mice displayed an osteopetrotic phenotype with excessive trabeculation of the bone-marrow space, further confirming the importance of *Ctsk* in bone resorption (Saftig et al., 1998; Saftig et al., 2000; Kiviranta et al., 2005). The *Ctsl*-deficient mice developed periodic hair loss, and epidermal hyperplasia, acanthosis and hyperkeratosis (Roth et al., 2000), and significant ventricular and atrial enlargements that were associated with a comparatively small increase in relative heart weight (Stypmann et al., 2002). In addition, both heterozygote and homozygote *Ctsl*-deficient mice had significantly decreased trabecular, but not cortical, bone volume (Potts et al., 2004). These phenotypes indicate the importance of *Ctsl* for instance in epidermal homeostasis, regulation of hair follicle morphogenesis and cycling, cardiac morphology and function, and controlling bone turnover. Further, a missense mutation in *Ctsl* was identified as the molecular basis (Roth et al., 2000) of the previously-known *furless* mouse (Green, 1954). The *Ctss*-deficient mice displayed profound inhibition of invariant chain (Ii) degradation resulting in an impaired MHC class II peptide loading (Nakagawa et al., 1999; Shi et al., 1999) and defective microvessel development during wound repair (Shi et al., 2003). These findings suggest that *Ctss* is essential for effective Ii proteolysis and has a role in extracellular matrix degradation during microvessel formation.

The functional redundancy and function of cathepsins was investigated by combining *Ctsb*- and *Ctsl*-deficient mouse strains. A double deficiency of *Ctsb* and *Ctsl* was lethal during the second to fourth week of life, and associated with a degree of brain atrophy not previously seen in mice (Felbor et al., 2002). A massive amount of apoptosis of certain neurons was detected in the cerebral cortex, and in the Purkinje

and granular cell layers of the cerebellum. In addition to neurodegeneration, pronounced reactive astrocytosis, accumulation of ultrastructurally and biochemically unique lysosomal bodies in large cortical neurons, and axonal enlargements were detected. These findings indicate the essential role of *Ctsb* and *Ctsl* in the maintenance of the central nervous system (Felbor et al., 2002).

5. 2. Mice double deficient for *Cstb* and cathepsins B, L or S

To find out the contribution of cathepsins to the EPM1 phenotype mice doubly deficient for *Cstb* and either *Ctsb*, *Ctsl* or *Ctss* were created (Houseweart et al., 2003a). The removal of *Ctsl* and *Ctss* from *Cstb*-deficient mice did not improve any aspects of the EPM1 phenotype, whereas the removal of *Ctsb* resulted in 36-89% reduction in the amount of cerebellar granule cell apoptosis depending on the mouse age, as well as reduced the incidence of an incompletely penetrant eye phenotype. The ataxia and seizure phenotypes were not diminished. These results indicated that in the absence of an endogenous inhibitor, Ctsb can initiate or propagate apoptosis. Since the ataxia and seizure phenotypes were not dimished, another molecule, in addition to Ctsb, is probably involved in the pathogenesis of EPM1 or can partially compensate for its function. In addition, as the removal of any of these cathepsins did not totally improve the EPM1 phenotype, Cstb may have other functions beside protease inhibition (Houseweart et al., 2003a).

6. Features of eukaryotic promoters

Most biological processes are regulated at the level of gene transcription. Transcription in eukaryotic cells is conducted by three different RNA polymerases. RNA polymerase II (RNAP II) transcribes promoters of mRNA encoding genes. These promoters contain three types of DNA elements. Basal promoter elements are located near the transcription start site and bind the general transcription machinery proteins. Sequence-specific transcription factors bind to proximal promoter elements that lie upstream of the gene and to distal enhancer elements that can be located several kilobases away from the start site.

General transcription factors (GTFs) include TFIIA, -B, -D, -E, -F and -H that are needed for RNAP II mediated basal transcription. TFIID is composed of TATA-binding protein (TBP) and TBP-associated factors (TAFs) (Orphanides et al., 1996;

Woychik and Hampsey, 2002). TBP binds to a TATA box, an element located -31 to -26 bp upstream of the transcription start site. The TFIIB recognition element (BRE) is located -37 to -32 bp, initiator (Inr) -2 to +4 bp, and downstream promoter element (DPE) +28 to +32 bp from the start site. Different TAFs interact with DPE and Inr. Most core promoters contain a TATA element, and the minority of them that do not are called TATA-less promoters, and usually contain a compensatory DPE motif. The TATA box can function in the absence of BRE, Inr and DPE motifs, but DPE requires the presence of Inr (Butler and Kadonaga, 2001; Smale and Kadonaga, 2003).

The general transcription machinery containing RNAP II and GTFs is sufficient to promote basal gene transcription *in vitro*, whereas additional factors are required for the response to activators. The mediator complex is unable to bind specific DNA sequences, but transmits regulatory information from gene specific transcription factors to the core transcription machinery and is required for regulated transcription initiation (Flanagan et al., 1991). Insulator DNAs are needed to prevent an enhancer associated with one gene from inappropriately regulating neighboring genes (Burgess-Beusse et al., 2002). Sequence-specific factors are required to achieve efficient transcription and specific regulation of gene expression within a specific tissue or cell type (Orphanides et al., 1996; Woychik and Hampsey, 2002). There may be as many as 3000 transcription factors in humans (Lander et al., 2001a), highlighting the complexity of transcriptional regulation. In the following paragraphs some of the essential sequence-specific transcription factor families are introduced as examples.

AP1 (activating protein 1) consists of dimeric transcription factors composed of Jun, Fos or ATF (activating transcription factor) subunits that bind to common DNA sites, the AP1 sites, also known as the TPA (12-O-tetradecanoylphorbol 13-acetate)-responsive elements (TRE). The common feature among AP1 family members is the evolutionarily conserved bZIP (basic region leucine zipper) domain. The consensus sequence of AP1 sites is 5'-TGAG/CTCA-3' (Vogt and Bos, 1990; Angel and Karin, 1991; Karin et al., 1997). These proteins are differentially expressed and regulated, which results in different cell types having a complex mixture of AP1 dimers with different functions (Wagner, 2001).

The GC-boxes or CpG islands containing the sequence 5'-GGGGCGGGG-3' are important DNA regulatory elements widely distributed in promoters and

enhancers of housekeeping and tissue-specific genes. The human genome contains approximately 29000 of these CpG islands, which can also be modified by methylation. Sp1 (specificity protein 1) was the first identified transcription factor that binds to the GC-box regulatory elements (Dynan and Tjian, 1983a; Dynan and Tjian, 1983b). However, nowadays there are at least eight known transcription factors which share common structural features and DNA binding specificities with Sp1 (Zhao and Meng, 2005). All of these members contain a conserved DNA-binding domain consisting of three zinc fingers (Philipsen and Suske, 1999; Suske, 1999). Sp1 and Sp3 are both ubiquitously expressed, and while Sp3 also has the capacity to repress transcription, they compete for the same binding sites and it has been assumed that their relative abundance allows regulation of gene activities of different cell types (Suske, 1999). Sp4, a tissue restricted activator, is predominantly expressed in the CNS, but is also detectable in epithelial tissues, testis and developing teeth (Hagen et al., 1992). Proteins Sp5 through Sp9 are shorter in their amino acid lengths and are expressed in tissue-restricted manners. Sp1-like transcription factor family members have diverse functions during embryonic development, though the specific function of each member of the family is difficult to clarify due to overlapping expression and similar DNA binding properties (Zhao and Meng, 2005).

The basic-helix-loop-helix (bHLH) transcriptional regulatory proteins are key players in numerous developmental processes. The family consists of seven classes of bHLH proteins that bind to a consensus sequence 5'-CANNTG-3', called the E-box. The basic region acts as sequence-specific DNA-binding domain, which recognizes the E-box binding site and has additional preference for internal and flanking sequences (Blackwell and Weintraub, 1990). bHLH proteins form dimers through the interaction of their HLH domains and each half of the dimer recognizes one half of the E-box (Murre et al., 1989). Specific subfamilies of bHLH proteins act in the differentiation of various cell types. Myogenic regulatory factors (MRFs), bHLH class II proteins, are critical for the determination and terminal differentiation of skeletal muscle cells. MRFs consist of MyoD, Myf5, myogenin and Mrf4 (Lassar et al., 1986; Molkentin and Olson, 1996). The NeuroD subfamily, also bHLH class II proteins, is largely restricted to neural and neuroendocrine cells, and regulates neuronal specification and differentiation (Bertrand et al., 2002). This subfamily contains NeuroD, Mash1, Neurogenin1-3 (Ngn1-3) and Math1/3 (Anderson et al., 1997; Lee,

1997; Schuurmans and Guillemot, 2002). The E-protein subfamily, bHLH class I proteins, which include E12, E47, HEB and E2-2, are expressed in many tissues and are capable of forming either homo- or heterodimers. With few exceptions, class II proteins are incapable of forming homodimers and preferentially heterodimerize with an E-protein (Murre et al., 1989).

The nuclear steroid receptors form a large group of transcription factors with high homology in the DNA-binding domain and lower homology in the carboxyterminal ligand-binding domain (Evans, 1988; Lubahn et al., 1988). The androgen receptor (AR), glucocorticoid receptor (GR), progesterone receptor (PR) and mineralocorticoid receptor (MR) form a subgroup within the nuclear receptor family (Glass, 1994). These receptors bind to the glucocorticoid response element (GRE), also called progesterone response element (PRE) or androgen response element (ARE), whose consensus sequence is 5'-AGAACAnnnTGTTCT-3' (Nordeen et al., 1990; Roche et al., 1992; Lieberman et al., 1993; Lombes et al., 1993).

The nuclear factor (NF1) family, also known as CAAT box transcription factors (CTFs), are composed of four members NFI-A, -B, -C and -X. NFI proteins can both activate and repress expression on consensus binding sites 5'-TTGGC(N5) GCCAA-3', which have been characterized from genes expressed in almost every organ system and tissue (Gronostajski, 2000).

7. Disease associated repeat expansion mutations

Repeat expansion is a unique form of mutation that is linked to more than 40 neurological, neurodegenerative and neuromuscular disorders. The first identified pathogenic repeat mutations were in fragile X syndrome (FRAXA) (Kremer et al., 1991; Verkerk et al., 1991), and in spinal and bulbar muscular atrophy (SBMA) (La Spada et al., 1991). Repeat expansion disorders not only include trinucleotide repeats, but also tetranucleotide (myotonic dystrophy type 2 [DM2]), pentanucleotide (spinocerebellar ataxia type 10 [SCA10]), minisatellite (EPM1, Creutzfeldt-Jakob disease [CJD] and diabetes) and megasatellite repeats (facioscapulohumeral muscular dystrophy type 1A [FSHMD1A]) (Wijmenga et al., 1992; Owerbach and Gabbay, 1993; Lalioti et al., 1997b; Matsuura et al., 2000; Liquori et al., 2001).

Beyond expansion of a repeat, the genetics of these diseases seem to have little in common. Genes carrying expanded repeats can be situated on the X chromosome

or autosomes, and the pattern of inheritance can be dominant, recessive or multifactorial (Table 3). Target genes have no evident functional similarities and repeats are located in various parts of genes resulting in either loss or gain of function. Repeat expansions in non-coding regions cause disorders resulting either from a loss-of-function through to inhibiting gene expression (Pandolfo, 1998; Jin and Warren, 2000) or from an RNA-mediated gain-of-function mechanism, in which the expanded RNA transcripts form imperfect double-stranded structures that accumulate and disrupt RNA-processing functions (Ranum and Cooper, 2006). Repeat expansions in coding regions can lead to change of function of the mutant protein or to a protein gain-of-function through poly-amino acid stretches, which may lead to self-aggregation and aggregation with other proteins (Orr and Zoghbi, 2001). In the following paragraphs specific features of some of these diseases associated with repeat expansion mutations in non-coding, coding and other chromosomal locations are highlighted (Table 3).

Table 3. Diseases associated with repeat expansion mutations.

DISEASE	INHERITANCE	GENE	LOCALIZATION	REPEAT SEQUENCE	NORMAL	MUTATION
NON-CODING						
DM1[1]	AD	*DMPK*	3'UTR	CTG	5-38	50-2000
DM2[2]	AD	*ZNF9*	intron	CCTG	10-26	75-11000
EPM1[3]	AR	*CSTB*	promoter	CCCCGCCCCGCG	2-3	30-125
FRDA[4]	AR	*FXN*	intron	GAA	6-36	90-1700
FRAXA FXTAS/POF[5]	XD	*FMR1*	5'UTR	CGG	6-60	200-2000 55-200
FRAXE[6]	XR	*FMR2*	5'UTR	CCG	4-39	200-900
SCA10[7]	AD	*ATXN10*	intron	ATTCT	10-22	500-4500
SCA12[8]	AD	*PPP2R2B*	promoter and 5'UTR	CAG	7-45	55-78
CODING						
CJD[9]	AD	*PRNP*	exon	24 bp	5	6-13
DRPLA[10]	AD	*ATN1*	exon	CAG	7-34	49-88
HD[11]	AD	*HD*	exon	CAG	11-34	40-121
SBMA[12]	XR	*AR*	exon	CAG	9-36	38-62
BOTH CODING AND NON-CODING						
SCA8[13]	AD	*ATXN8OS* *ATXN8*	3'UTR exon	CTG CAG	16-34	74-1300
HDL2[14]	AD	*JPH3*	exon and 3'UTR	CTG	6-28	43-78
OTHER CHROMOSOMAL LOCATIONS						
					Class I	Class III
DM[15]	MF	*INS*	VNTR	14-15 bp	26-63	141-209

The pattern of inheritance can be X-linked (X), or autosomal (A) and dominant (D) or recessive (R) as well as multifactorial (MF). [1](Brook et al., 1992); [2](Liquori et al., 2001); [3](Lalioti et al., 1997b); [4](Campuzano et al., 1996); [5](Kremer et al., 1991; Verkerk et al., 1991; Oostra and Willemsen, 2003); [6](Knight et al., 1993; Gecz et al., 1996; Gu et al., 1996); [7](Matsuura et al., 2000); [8](Holmes et al., 1999); [9](Owen et al., 1989; Owen et al., 1990); [10](Koide et al., 1994); [11](Huntington's Disease Collaborative Research Group, 1993); [12](La Spada et al., 1991); [13](Koob et al., 1999; Moseley et al., 2006); [14](Holmes et al., 2001b); [15](Owerbach and Gabbay, 1993).

Myotonic dystrophy type 1 (DM1) is caused by a CTG expansion, located in the 3'UTR of the dystrophia myotonica-protein kinase (*DMPK*) gene (Brook et al., 1992). DMPK is a serine/threonine protein kinase detected only in muscles (Brook et al., 1992). Myotonic dystrophy type 2 (DM2) is caused by a CCTG expansion in intron 1 of the zinc finger 9 (*ZNF9*) gene (Liquori et al., 2001). ZNF9 is a nucleic acid binding protein that is highly expressed in skeletal muscle and has been suggested to modulate the activity of the beta-myosin heavy chain (Flink and Morkin, 1995). Both diseases are mediated by an RNA-mediated gain-of-function mechanism (Liquori et al., 2001; Ranum and Day, 2004a).

Most Friedreich ataxia (FRDA) patients have homozygous GAA repeat expansions within the first intron of the frataxin (*FXN*) gene, which results in loss-of-function (Campuzano et al., 1996; Campuzano et al., 1997). FXN is involved in mitochondrial iron metabolism (Rotig et al., 1997). The loss of FXN leads to accumulation of mitochondrial iron, increased susceptibility to oxidative stress and reduction in oxidative phosphorylation (Wilson, 2003).

Fragile X mental retardation (FRAXA) is caused by a CGG expansion of greater than 200 repeats in the 5'UTR of the fragile X mental retardation 1 (*FMR1*) gene through a loss-of-function mechanism (Kremer et al., 1991), whereas smaller "premutation" expansions (55-200 repeats) can cause fragile X tremor ataxia syndrome (FXTAS) and possibly premature ovarian failure (POF) through an RNA gain-of-function mechanism (Oostra and Willemsen, 2003; Hagerman and Hagerman, 2004). The FMR1 protein is proposed to act as a regulator of mRNA transport and/or translation (Corbin et al., 1997). FRAXE patients have expansion of a CCG repeat in the 5'UTR of the fragile X mental retardation 2 (*FMR2*) gene, which leads to hypermethylation of the CpG island and transcriptional silencing (Knight et al., 1993; Gecz et al., 1996; Gu et al., 1996). FRM2 is a potential transcription factor (Gecz et al., 1997).

Spinocerebellar ataxia type 8 (SCA8), a dominantly inherited disease of reduced penetrance, is caused by a CTG expansion located in the 3'UTR of the non-coding ataxin 8 opposite strand (*ATXN8OS*) gene (Koob et al., 1999) that overlaps with the 5'UTR of the Kelch-like 1 (*KLHL1*) gene in the opposite direction. *ATXN8OS* might regulate *KLHL1*, which encodes an actin-binding protein, through an antisense mechanism (Nemes et al., 2000). Expansion in *ATXN8OS* is also proposed

to cause SCA8 by an RNA gain-of-function mechanism (Ranum and Day, 2004b). On the opposite strand the CAG expansion is also located in the newly discovered ataxin 8 (*ATXN8*) gene. The expansion in *ATXN8* produces a protein that contains only glutamine residues, and is suggested to cause the disease by a protein gain-of-function mechanism (Moseley et al., 2006). SCA10 is caused by a large expansion of an intronic ATTCT repeat in the spinocerebellar ataxia 10 (*ATXN10*) gene (Matsuura et al., 2000) and is probably mediated by an RNA-mediated gain-of-function mechanism (Lin and Ashizawa, 2005). The CAG repeat expansion in SCA12 is located either in the promoter region or 5'UTR of the protein phosphatase 2 (*PPP2R2B*) gene, depending on different transcript transcriptional start sites (Brook et al., 1992; Holmes et al., 1999). The expansion mutation effect is through increasing the gene expression and an RNA gain-of-function mechanism (Holmes et al., 1999; Holmes et al., 2001a; Holmes et al., 2003).

Creutzfeldt-Jakob disease (CJD) is caused by a 24 bp minisatellite repeat expansion in the prion protein (*PRNP*) gene causing octapeptide repeats in the prion protein (PrP) (Owen et al., 1989; Owen et al., 1990). The cellular isoform PrP^{C} is converted to an abnormal conformed scrapie isoform PrP^{Sc}, which is an infectious agent.

Dentatorubral-pallidoluysian atrophy (DRPLA) (see earlier, section 2) is caused by a CAG repeat in the coding region of the adenine nucleotide translocator atrophin 1 (*ATN1*) gene (Koide et al., 1994), forming polyglutamine stretches. ATN1 is involved in repression of transcription (Zhang et al., 2002).

Huntington's disease (HD) is caused by a CAG repeat within the coding region of the huntingtin (*HD*) gene, which is translated as a polyglutamine repeat in the protein product (Huntington's Disease Collaborative Research Group, 1993). The proposed mechanisms includes altered transcription, protein aggregation, altered protein interaction or function, proteolytic cleavage of HD or defects in neurotransmitter signalling (Lee and Kim, 2006). Huntington disease-like 2 (HDL2) is caused by a CTG expansion, which depending on the splicing is located either in a variably spliced exon of the junctophilin 3 (*JPH3*) gene leading to polyalanine or polyleucine stretches, or in the 3'UTR leading to an RNA-mediated gain-of-function mechanism (Holmes et al., 2001b). JPH3 is a plasma membrane/endoplasmic reticulum junction protein (Takeshima et al., 2000).

Spinal and bulbar muscular atrophy (SBMA) caused by a CAG expansion (La Spada et al., 1991) in the coding region leads to polyglutamine disease through accumulation of mutant androgen receptor (AR), a hormone-activated transcription factor (Chang et al., 1988).

A variable number of tandem repeat (VNTR) regulatory polymorphism, located in the promoter region of the insulin gene (*INS*), is associated with susceptibility to diabetes. The VNTR consists of 14-15 bp oligonucleotides related to the sequence ACAGGGGTGTGGGG (Owerbach and Gabbay, 1993). The number of repeats varies and form three classes: class I (26-63 repeats), class II (mean of 80 repeats) and class III (141-209 repeats). The class I homozygous genotype is associated with elevated risk of developing type 1 diabetes mellitus (DM), while the class III allele is associated with e.g. type 2 diabetes and larger size at birth (Bennett and Todd, 1996; Bennett et al., 1996; Dunger et al., 1998).

The premutation associated with e.g. HD and FRAXA refers to an intermediate number of repeats that on transmission can increase to disease length. Anticipation, the genetic phenomenon in which disease severity increases and/or age of onset of disease decreases from one generation to the next as the repeat expansion becomes longer, is associated with e.g. FRAXA, DM1, SBMA, HD and FRDA.

AIMS OF THE PRESENT STUDY

Prior to this study, the *CSTB* gene responsible for Unverricht-Lundborg disease had recently been identified (Pennacchio et al., 1996). With the aim of enhancing our understanding of the disease pathogenesis of EPM1 this study was undertaken to analyze the regulation and expression of the *CSTB* gene, and to initiate the characterization of the CSTB protein.

The specific aims of this study were to:

- Characterize the *CSTB* gene.
- Determine the spatial and temporal expression of the *Cstb* gene.
- Define the subcellular localization of the CSTB protein.
- Study the consequences of EPM1-associated mutations on *CSTB* gene and CSTB protein expression.

MATERIALS AND METHODS

1. Ethical considerations

EPM1 patients and carriers gave informed consent for performing genetic, molecular or biochemical analyses before the blood or skin samples were collected. The project has been approved by the Ethical Committee of the Department of Medical Genetics, University of Helsinki. Mice care and handling were conducted according to the National Research Council's guide for the care and use of laboratory animals.

2. EPM1 patients and control individuals

EPM1 patients, carriers and controls were chosen from among individuals included in the earlier phase of the project. Two novel EPM1 patients were referred to us for mutation analysis by clinicians. EPM1 patients' and mutation carriers' total RNA (I, IV), genomic DNA (IV) and/or proteins (III, IV) were extracted from blood, lymphoblastoid cells or fibroblasts as described in I, III and IV. Samples from unrelated Finnish and Centre d'Etude du Polymorphisme Humain (CEPH) (http://ccr.coriell.org/nigms/ceph/ceph.html) individuals were used as well.

3. Mouse embryos and tissues

For whole-mount RNA *in situ* hybridization NMRI mouse embryos at E7 and E8 were collected. Previous and novel cryosections of mouse embryos from various wild-type mice strains at E14 and E17, and mouse brains at P1, P5, P21 and adult were used.

4. DNA constructs

The constructs containing different length *CSTB* promoters, wild-type or mutant human *CSTB* and *CSTB2* or mouse *Cstb* and *Cstb2* in suitable plasmid vectors produced and used in the study are summarized in Table 4. The *CSTB2* cDNA in pcDNA3.1(+) vector (Invitrogen) was generated with primers 5'-CAA GAT GAT GTG CGG GG-3' and 5'-GCT CTG GTA GAC GGA GGA TG-3' and mutations c.10G>C (p.Gly4Arg) and c.149G>A (p.Gly50Glu) were introduced using the QuikChange Site-Directed Mutagenesis Kit (Stratagene). The 299 bp fragment of *Cstb2* in pBluescript SK(+) vector (Stratagene), used as probe in RNA *in situ* hybridization, was generated with primers 5'-tcc acc gcg gtg gcg gcc gcg gCA TGA ACT GGG GAC ATA G-3' and 5'-ccc ccc ctc gag gtc gaC TGG AGG AAG ACA

GAC AGA CAG A-3' (this thesis), containing *Not*I or *Sal*I restriction sites (underlined), respectively.

Table 4. Plasmids constructed.

NAME	PLASMID	PUBLICATION				
CSTB promoter 5873 bp	pbluescript SK(-)	I				
CSTB promoter -670 to -1 bp	pGL3-Basic	I				
CSTB promoter -228 to -1 bp	pGL3-Basic	I				
CSTB promoter -145 to -1 bp	pGL3-Basic	I				
CSTB promoter -670 to -1 bp + 19x 12-mer	pGL3-Basic	I				
CSTB promoter -670 to -514 bp and -411 to -1 bp	pGL3-Basic	I				
CSTB	pcDNA3.1(+)			III		
CSTB p.Gly4Arg	pcDNA3.1(+)			III		
CSTB p.delVal23_Lys56	pcDNA3.1(+)				IV	
CSTB p.Gly50Glu	pcDNA3.1(+)				IV	
CSTB p.Arg68X	pcDNA3.1(+)			III		
CSTB p.Gln71Pro	pcDNA3.1(+)			III		
CSTB p.Leu73fsX3	pcDNA3.1(+)			III		
CSTB2	pcDNA3.1(+)					this thesis
CSTB2 p.Gly4Arg	pcDNA3.1(+)					this thesis
CSTB2 p.Gly50Glu	pcDNA3.1(+)					this thesis
Cstb 544 bp RNA *in situ* hybridization probe	pbluescript SK(+)		II			
Cstb 173 bp RNA *in situ* hybridization probe	pbluescript SK(+)		II			
Cstb2 RNA *in situ* hybridization probe	pbluescript SK(+)					this thesis

5. Cell lines

Commercially available cell lines COS-1 (African green monkey kidney cells) (I, III, IV) and BHK-21 (baby hamster kidney cells) (III, IV, this thesis) from the American Type Culture Collection (ATCC), as well as human primary pectoral muscle myoblasts (kindly provided by Olli Carpén, University of Helsinki, Finland) (III) and differentiated myotubes (III) were used.

6. Primary antibodies

The monoclonal CSTB 2E7 antibody against purified human CSTB protein produced in study III and the commercial polyclonal CSTB antibody (Biogenesis) were used in various experiments (III, IV, this thesis). As organelle markers we used polyclonal antibodies against protein disulphide isomerase (PDI) from StressGen Biotechnologies and *trans*-Golgi network protein 46 (TGN46) from Serotec; mammalian secretory protein 13 (mSec13) (kindly provided by Wanjin Hong, The National University of Singapore, Republic of Singapore), mannose-6-phosphate receptor (MPR) (Kurt von Figura, University of Göttingen, Germany),

aspartylglucosaminidase (AGA) (Anu Jalanko, National Public Health Institute, Finland) and lysosome-associated membrane protein 1 (LAMP1, lgp120) (Peter van der Sluijs, University Medical Center Utrecht, the Netherlands). In addition, FITC-labelled Lentil lectin was from Sigma and LysoTracker® Green DND-26 from Molecular Probes.

7. Methods

The methods used in the published papers and in the unpublished experiments included in this thesis are summarized in Table 5.

Table 5. Methods used.

METHODS	PUBLICATION				
Recombinant DNA techniques (cloning)	I	II	III	IV	this thesis
DNA extraction	I	II	III	IV	this thesis
PCR (polymerase chain reaction)	I	II	III	IV	this thesis
DNA sequencing	I	II	III	IV	this thesis
Computer sequence analysis (BLAST, RepeatMasker, Primer 3, MatInspector, ASAP, AceView)*, (Sequencer 4.0, Primer Express)	I	II	III	IV	this thesis
Northern blot analysis	I	II		IV	
Cell culture	I		III	IV	this thesis
Transient DNA transfections	I		III	IV	this thesis
RNA isolation	I			IV	
Luciferase and β-galactosidase activity assay	I				
Electrophoretic mobility shift assay (EMSA)	I				
Radioactive RNA *in situ* hybridization		II			this thesis
Computer-based MCID image analysis		II			
DIG-labelled RNA *in situ* hybridization (in whole-mount embryos)		II			
Production of monoclonal antibodies		II	III		this thesis
Quantitative real-time RT-PCR		II		IV	
Semi-quantitative RT-PCR					this thesis
Site-directed mutagenesis			III		this thesis
In vitro translation (IVT)			III		this thesis
Immunoprecipitation			III		this thesis
Protein extraction			III	IV	this thesis
Western blot analysis			III	IV	this thesis
ECL (enhanced chemiluminiscence)			III	IV	this thesis
Immunofluorescence stainings			III	IV	this thesis
Immunofluorescence microscopy			III	IV	this thesis
Dark- and bright-field microscopy		II			
Patient mutation analysis				IV	

* BLAST (http://www.ncbi.nlm.nih.gov/BLAST/), RepeatMasker (http://www.repeatmasker.org/), Primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi), MatInspector (http://www.genomatix.de/matinspector.html), ASAP (http://www.bioinformatics.ucla.edu/ASAP/), AceView (http://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/).

RESULTS AND DISCUSSION

1. Novel patient mutations (IV)

Two novel EPM1 mutations, c.149G>A and c.168+1_18del, in the *CSTB* gene were identified in compound heterozygous form with the dodecamer repeat expansion in two newly clinically diagnosed patients (IV). None of the 166 controls tested carried these mutations. The expansion mutations in these patients were verified with the novel PCR protocol (IV, Fig. 2), which allows direct reliable detection of expanded *CSTB* alleles.

The novel missense c.149G>A mutation was identified in a Finnish EPM1 patient. It resulted in substitution of glycine by glutamic acid at amino acid position 50 (p.Gly50Glu) (IV, Fig. 3A). This mutation is situated in the highly conserved QVVAG sequence critical for papain binding (Stubbs et al., 1990). No mutation affecting this pentapeptide has previously been identified. The p.Gly50Glu mutation seems to be translated in patient cells (see below, section 4.2) and is likely to affect the stability or life span of the inhibitory effect of the protein, or the interactions with target proteins (Jerala et al., 1990; Auerswald et al., 1995). That EPM1 results from a mutation in this site implies the physiological importance of cathepsin binding.

The second novel mutation, observed in an Italian EPM1 patient, is a deletion of the first 18 nucleotides downstream from the 5' splice donor site of intron 2 (c.168+1_18del) (IV, Fig. 3B). On the mRNA level, this mutation results in aberrant splicing of *CSTB* with production of two different transcripts (IV, Fig. 3C). The shorter transcript has an in-frame deletion of exon 2 and predicts a deletion of a 34 amino acid region containing the QVVAG sequence, creating a p.delVal23_Lys56 polypeptide. The longer transcript contains 25 nucleotides from intron 2 resulting in a frameshift and predicting a premature stop codon (p.Val57GlufsX28). It remains, however, unknown whether these transcripts are translated or whether the gene products are degraded after transcription, e.g. due to nonsense-mediated-decay. The p.delVal23_Lys56 protein was not detected *in vitro* or *in vivo* (see below, section 4.2).

2. Characterization of the *CSTB* gene

2. 1. The *CSTB* promoter and its transcriptional elements (I)

To characterize the *CSTB* promoter the 5' genomic sequence of *CSTB* was extended by 3051 bp by primer walking from the previously published sequence (Pennacchio et al., 1996). An Alu-rich region at -1782 to -686 bp upstream of the translation initiation codon was identified and further characterization was limited downstream of it. The sequence analysis revealed many similarities with a typical housekeeping gene promoter with a high GC-content (75%), a high number of CpG dinucleotides and the presence of several Sp1 binding sites. The promoter region of the cystatin C gene (abbreviated *CST3*) shares these features (Abrahamson et al., 1990), whereas the *CSTA* promoter has a low GC-content (36%) and lacks putative Sp1 binding sites (Takahashi et al., 1998). All three promoters lack CAAT and TATA boxes. The nucleotide sequence of the putative *CSTB* promoter region was analyzed with the MatInspector program to identify putative transcriptional elements. In addition to the previously published five putative Sp1 sites (Pennacchio et al., 1996), four putative AP1 sites, an androgen response element (ARE) half site and a MyoD site were identified (I, Fig. 1A). The minimal functional promoter was mapped using promoter-reporter gene constructs containing three different genomic fragments from -670 (construct 1), -228 (construct 2) and -145 (construct 3) to -1 bp upstream of the translation initiation codon (Fig. 4; I, Fig. 1B).

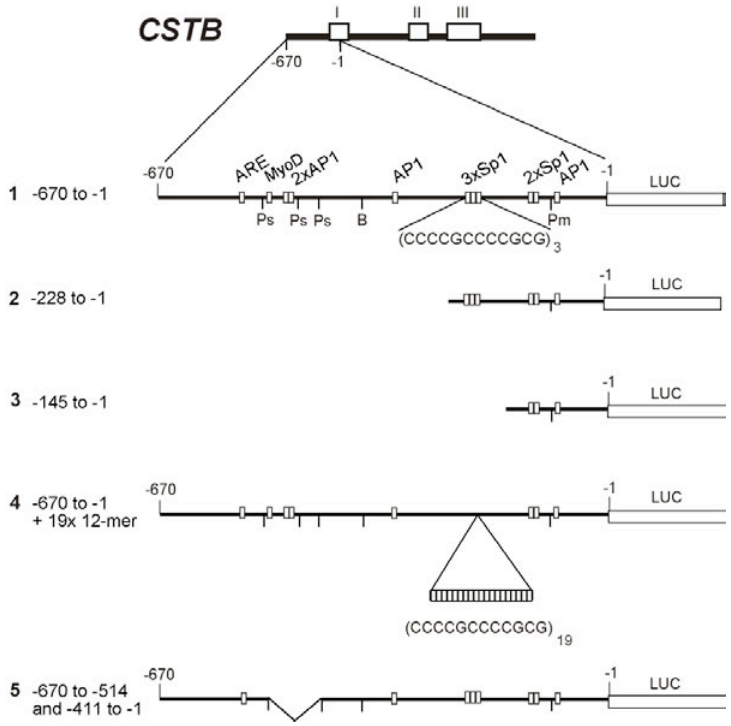


Figure 4. The *CSTB* gene and its putative promoter region. On the top is shown the genomic structure of the *CSTB* gene with its three exons. The enlarged view shows the *CSTB* promoter region with a putative androgen response element (ARE) half site, a potential MyoD element, four AP1 and five Sp1 sites. The three *Pst*I (Ps), a *Bssh*II (B) and a *Pml*I (Pm) restriction endonuclease cleavage sites are indicated below the black bar. Construct 1 from -670 to -1 bp, construct 2 from -228 to -1 bp and construct 3 from -145 to -1 bp were cloned into the pGL3-Basic vector containing the luciferase-reporter (LUC) gene. Construct 4 contains the region -670 to -1 bp and has 16 extra copies (19 copies in total) of the dodecamer repeat. Construct 5 contains regions -670 to -514 bp and -411 to -1 bp, and lacks the region containing two AP1 sites and the MyoD element. Constructs 1, 2 and 5 contain three dodecamer repeat copies. Picture from publication I.

The promoter activity of construct 3 containing two potential Sp1 sites and one AP1 site was over 40 times, of construct 2 with three additional Sp1 sites approximately 60 times, and of construct 1 with the addition of three AP1, one MyoD and one ARE sites about 300 times the activity of the empty luciferase construct (I, Fig. 2). The activity of the promoter was orientation-dependent, since constructs 1 (I, Fig. 2), 2 and 3 (I, data not shown) showed less activity in reverse than in forward orientation. Construct 5 without two AP1 sites and a MyoD element, lacking nucleotides -514 to -411 bp from construct 1, reduced the promoter activity by 30% (I, Fig. 2), suggesting that these sites are involved in the regulation of *CSTB* transcription. The sequence from -670 to -229 bp seems to contain strong activator elements as the activity of construct 1 was five times that of construct 2. These data suggested that the region between -670 and -1 bp from the translation initiation codon acts as the promoter of the *CSTB* gene. Studying a 3.2 kb promoter construct in four different cell lines (SK-N-BE, CHP [both derived from human neuroblastomas], HeLa

[derived from cervical cancer cells] and COS-7), Lalioti et al. (1999) detected at least a 100-fold increase in promoter activities when compared to a construct without a promoter. Our higher, 300-fold, increase in promoter activity could be explained by the very different lengths of promoter constructs and/or the different cell lines used (in our assays COS-1 cells).

DNA-protein interaction assays with four AP1 and five Sp1 sites as well as an ARE half site were performed by electrophoretic mobility shift assays (EMSAs). The interaction of the MyoD site was not studied. The tandem AP1 site at -480 to -469 bp (I, Fig. 3), as well as the AP1 sites at -316 to -313 bp (I, Fig. 3) and -83 to -79 bp (I, data not shown) were capable of forming complexes with c-Jun. The Sp1 sites at -107 to -86 bp, -120 to -99 bp and -210 to -175 bp formed complexes with purified human Sp1 protein (I, Fig. 4). Sp1 proteins bound particularly strongly to the oligonucleotide of Sp1 site at -210 to -175 bp containing three copies of the Sp1 element (I, Fig. 4), which are part of the sequence in the dodecamer repeats. Similarly, the *CST3* gene promoter region at -257 to -211 bp contains three Sp1 binding sites that are thought to play a major role in transcription (Olafsson, 1995). The *CST3* promoter contains 11 possible Sp1 binding sites, but only one AP1 site (-970 to -960 bp) (Olafsson, 1995). Binding of androgen receptor (AR) to the ARE half site appeared to be relatively weak as it was only recognized in the presence of an AR stabilizing antibody (I, Fig. 5). A similar ARE sequence is found in the mouse *Cst3* gene (Huh et al., 1995), but not in the human *CST3* or *CSTA* genes (Abrahamson et al., 1990; Takahashi et al., 1998). Two Sp1/GC, one AP2 and one AP1 transcription binding sites are found in human, mouse and rat *CSTB* promoter regions (Lalioti et al., 1999).

In conclusion, the region between -670 to -1 bp from the translation initiation codon is suggested to induce the basal activity of *CSTB* expression and all transcription binding sites studied were found to be capable of binding *in vitro* to sequence-specific factors.

2. 2. Expression of the *Cstb* and *CSTB* genes

Previously, the *CSTB* gene had been found to be ubiquitously expressed in adult human and mouse tissues (Pennacchio et al., 1996; Pennacchio and Myers, 1996). The clinical manifestations of EPM1 seem to be limited to the CNS. In order to assess the potential cellular specificity of *CSTB* expression, the spatial and temporal expression

of *Cstb* in mouse embryos and adult mouse tissues, with special emphasis on the brain, was characterized in more detail.

2.2.1. Tissue expression (II, IV)

The tissue expression of *Cstb* in various mouse fetal and adult tissues was first studied with Northern blot analysis and quantitative real-time PCR. With both methods, a strikingly high level of *Cstb* expression was detected in E7 embryos (II, Fig. 1A, B). In real-time PCR analysis the level of *Cstb* expression in E7 embryos was approximately twenty times higher than the mean expression in other developmental stages studied (II, Fig. 1B). In Northern blot analysis highest expression in adult tissues was detected in thymus (II, Fig. 1A). In the quantitative real-time analysis, in which the commercial panel we used lacked thymus, highest expression was detected in kidney, liver and lung (II, Fig. 1B). Interestingly, the expression in brain was among the lowest (II, Fig. 1A, B). In a parallel study, *CSTB* expression in human fetal and adult tissues was quantitated using real-time PCR. *CSTB* was ubiquitously expressed, being highest in lung and relatively low in brain (IV, Fig. 4A), a pattern very similar to that seen in mouse (II).

2.2.2. Spatial expression in mouse embryos and postnatal brain (II, unpublished)

The strikingly high levels of *Cstb* in the E7 embryos in the Northern blot and quantitative real-time PCR analyses prompted us to study the spatial expression of *Cstb* during early mouse embryogenesis by whole-mount RNA *in situ* hybridization. These studies were continued using cryosections of embryos during late embryogenesis as well as of postnatal mouse brains, where the EPM1 symptoms manifest.

The expression of *Cstb* was intense and uniform in the entire E7 and E8 mouse embryos, but was not detected in extraembryonal tissues (II, Fig. 2A, C). Interestingly, cathepsins are expressed in the extraembryonal tissues at this stage (Sol-Church et al., 1999). Highest Ctsl activity is detected in visceral yolk sac and highest *Ctsl* mRNA levels in placenta of E9.5 embryos (Sol-Church et al., 1999). Expression of *Ctsb* mRNA in placenta increases gradually from E7.5 to E17.5 embryos (Sol-Church et al., 1999). In E9.5 and E10.5 embryos *Cstb* expression has been reported in brain, spinal chord, eye, otocyst, dorsal root ganglia, branchial arches, heart, gut,

kidney, limbs and skin (Reymond et al., 2002). *Cstb* has been included in a cluster of genes which by EST mining are expressed in the total fetus (Gitton et al., 2002).

The *Cstb* localization in later embryogenesis was studied using radioactive RNA *in situ* hybridization in sectioned E14 and E17 embryos. In E14 and E17 embryos quantification with a computer-based MCID image analysis system revealed strongest *Cstb* expression in thymus (II, Fig. 3A, C; Fig. 5B), concordant with high thymic expression detected in Northern blot analysis (II, Fig. 1A). In line with these findings, relatively strong expression of *Cstb* in thymus at this stage has earlier been reported (Reymond et al., 2002). Interestingly, the targets of CSTB, cathepsins, have been implicated to have a role in immune response through processing functions in mammalian antigen-presenting cells (reviewed in Zavasnik-Bergant and Turk, 2006). Moreover, *Tt-cysb*, the leech *Theromyzon tessulatum* homologue of mammalian *CSTB*, has been shown to become up-regulted after immunostimulation with a bacterial challenge, implying a role for Tt-cysb in leech innate immunity (Lefebvre et al., 2004).

Cstb expression in E14 and E17 embryos was detected in several tissues in addition to thymus with a restricted cellular expression pattern. *Cstb* was expressed in several epithelial layers such as in the nasopharyngeal epithelium (Fig. 5C; II, Fig. 3D), pharynx and tracheal epithelium (II, data not shown), epithelium lining the intestine (Fig. 5F; II, Fig. 3H) and epithelium of the skin (II, data not shown). Concordant with these findings, Reymond et al. (2002) have earlier reported *Cstb* expression in skin and intestine in E14.5 embryos. Interestingly, Cathepsin L-like cysteine protease (CPL-1) in *Caenorhabditis elegans* is highly expressed in gut cells during late embryonic development (Hashmi et al., 2002). The intestinal mucosa is known to play an important role in mucosal and systemic immune response (Muller et al., 2005), and it is possible that *Cstb* is also linked to this kind of immune response through regulation of cathepsin activity. *Cstb* was expressed in the neopallial cortex (future cerebral cortex) (Fig. 5A; II, Fig. 3A, B). Expression was observed in certain parts of the lower incisor tooth in outer enamel epithelium (Fig. 5D; II, Fig. 3E) and in maxilla (Fig. 5D; II, Fig. 3E) and mandible (II, Fig. 3F), as well as in vertebra bone forming cells (Fig. 5E; II, Fig. 3G). *Cstb* expression in developing bone (Fig. 5D, E; II, Fig. 3E-G) is in line with recent findings of CSTB having regulatory functions in

osteoclasts, where it could inhibit bone resorption by down-regulating intracellular CTSK activity (Laitala-Leinonen et al., 2005).

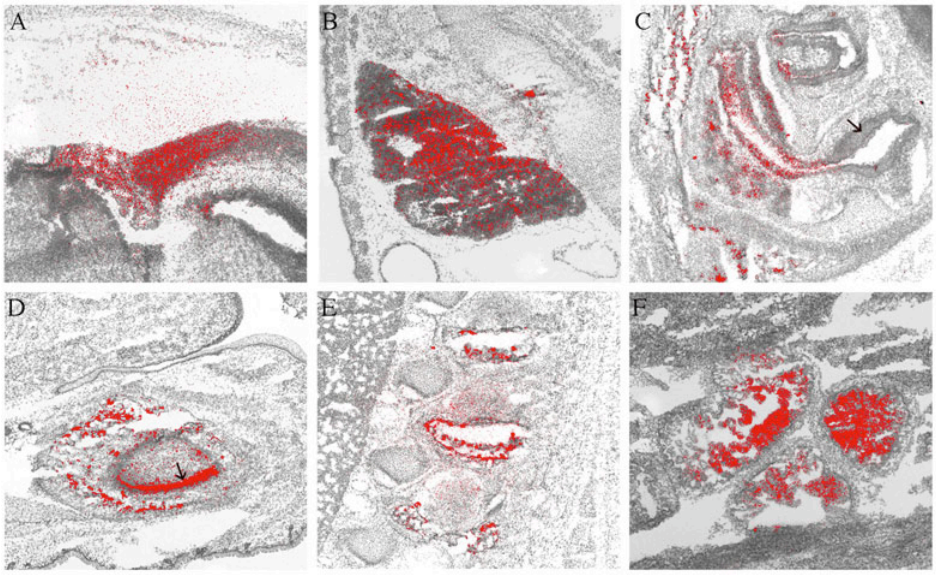


Figure 5. *Cstb* expression in E17 mouse embryos. *Cstb* hybridization signals visualized in neopallial cortex (future cerebral cortex) (A), thymus gland (B), nasopharyngeal epithelium, while the olfactory epithelium is negative (indicated by an arrow) (C), primordium of lower incisor tooth in outer enemal epithelium (shown by an arrow) (D), mandible (D), vertebra bone forming cells (E) and epithelium of the intestine (F). The dark-field images, that were inverted and stained red, are combined with the bright-field images presenting hematoxylin staining. Magnification is 50X. Picture modified from publication II.

In postnatal mouse brain the pattern of the *Cstb* expression was restricted. In P1 and P5 mouse brain sections *Cstb* was expressed in cortex, caudate putamen, hippocampus and thalamus (II, Fig. 4A, B; Fig. 6A). Expression, as determined by relative optical density levels, was higher at P1 than at P5 (II, Fig. 4A). The cortical expression seen in P1 and P5 mouse brains (II, Fig. 4A, B; Fig. 6A) is in line with the *Cstb*-deficient mice having cortical neuronal atrophy, neuronal apoptosis and gliosis (Shannon et al., 2002). In hippocampus, *Cstb* expression was seen in the CA1, CA2 and CA3 pyramidal cell layers and in the dentate gyrus granule cell layer (Fig. 6A-C; II, Fig. 4A-D), compatible with *Cstb*-deficient mice having neuronal apoptosis and gliosis in the hippocampal formation (Shannon et al., 2002). In line with our results are findings of *Cstb* mRNA expression in the rat, where expression was highest in the hippocampal formation and reticular thalamic nucleus, and moderate in the amygdala, thalamus, hypothalamus and cortical areas (D'Amato et al., 2000).

In the cerebellum *Cstb* expression was most prominent in the Purkinje cell layer in P21 (Fig. 6D; II, Fig. 4E, F) and adult mice brains (II, data not shown). *Cstb* expression in granule and molecular cell layers seems to be much weaker (Fig. 6D; II,

Fig. 4E, F), though the overall signal in the cerebellum could not be quantitated with the MCID image analysis system (II, data not shown). Concordant with the high expression in Purkinje cells, EPM1 patients have been reported to have loss of Purkinje cells, which might, however, been attributed to the use of phenytoin (Haltia et al., 1969; Koskiniemi et al., 1974a; Eldridge et al., 1983). This view is supported by the findings in *Cstb*-deficient mice that have not been exposed to medication and that display only occassional loss and ballooning of Purkinje cells, indicating that *Cstb*-deficiency *per se* can cause loss of these cells, although to a lesser extent (Shannon et al., 2002). In light of high *Cstb* expression in Purkinje cells the striking granule cell loss seen in the *Cstb*-deficient mice could be secondary to the dysfunction of Purkinje cells (Riccio et al., 2005).

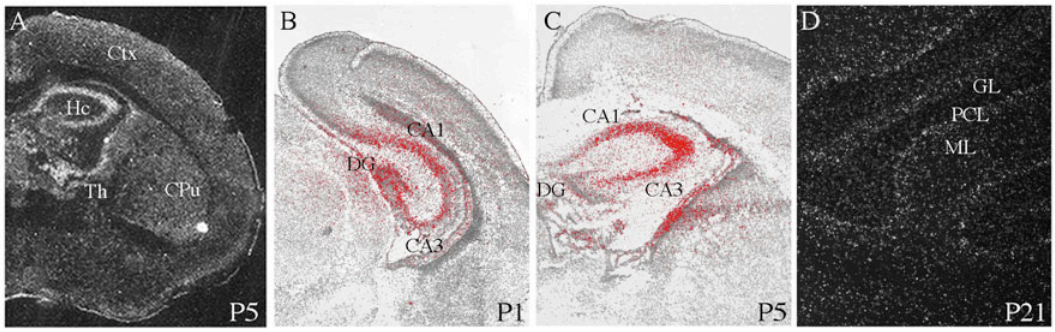


Figure 6. *Cstb* expression in postnatal mouse brain. (A) In frontal P5 brain section *Cstb* expression is seen in cortex (Ctx), caudate putamen (CPu), hippocampus (Hc) and thalamus (Th). In fronal P1 (B) and P5 (C) sections *Cstb* is detected in hippocampal regions CA1, CA3 and dentate gyrus (DG). (D) In sagittal P21 section expression is prominent in the Purkinje cell layer (PCL) of cerebellum. GL, granular layer. ML, molecular layer. Magnification is 25X in A, 50X in B, C and 100X in D. Modified from publication II.

In conclusion, *Cstb* expression seems to be under tight temporal and spatial regulation. *Cstb* is widely expressed throughout embryonic development, with highest expression in E7 embryos as well as in the embryonic and adult thymus, and epithelial layers of several tissues. In the CNS *Cstb* expression is highly restricted in concordance with the clinical presentation and limited neuronal pathology in humans and mice. Outside the CNS CSTB inhibitory activity seems to be functionally redundant, as there are no consistent developmental or non-neuronal phenotypes in EPM1 patients and *Cstb*-deficient mice.

We aimed to investigate the tissue distribution of the Cstb protein using immunohistochemical analysis. These studies were impossible to conduct, since despite several attempts we were unable to produce a specific monoclonal or polyclonal antibody recognizing the mouse Cstb protein. The commercially available

polyclonal antibody against human CSTB that has been used in several studies (Di Giaimo et al., 2002; Riccio et al., 2005) to recognize mouse and rat Cstb turned out, in our hands, to be unspecific in immunohistochemical stainings of mouse tissues and immunofluorescence analysis of mouse cells, as identical staining patterns were detected in both wild-type and *Cstb*-deficient mouse brain sections or primary neuronal cell cultures (data not shown). In Western blot staining this antibody seemed to be specific, since the Cstb protein was not detected from *Cstb*-deficient mouse brain extracts (data not shown). However, in addition to the Cstb-specific band, the antibody recognized several larger molecular mass proteins from wild-type and *Cstb*-deficient mouse brain extracts (data not shown). The analysis of Cstb protein expression in tissues still awaits a specific, working antibody.

2. 3. Splice variants and their expression (IV, unpublished)

Approximately 60% of human genes and at least 41% of mouse genes have alternatively spliced transcripts, resulting in the large number of proteins with different functions or tissue-specificities (Lander et al., 2001a; Okazaki et al., 2002). The number of tissue-specific splice variants is highest in brain, where they are involved in the development and differentiation of the CNS, and in controlling neuronal activity (Grabowski and Black, 2001; Xu et al., 2002). Alternative splicing of *CSTB* was suggested in the databases (ASAP and AceView). Human *CSTB* and mouse *Cstb* splice variants were isolated by *in silico* analysis and with RT-PCR, and were further characterized with various methods.

In addition to the normal *CSTB* splice form, a *CSTB2* splice variant was assembled, as well as *CSTB3*, *CSTB4* and *CSTB5* splice variants were identified (IV). Full assembly of the latter three variants was impossible, since amplification from the 5' end favored the *CSTB* isoform (IV). Moreover, expression of the *CSTB3*, *CSTB4* and *CSTB5* transcripts was restricted to colon, intestine and prostate (IV, data not shown), and therefore were thought to have no significance for EPM1 pathogenesis. The *CSTB2* transcript was ubiquitously expressed, also in brain. In this transcript intron 2 sequences are retained, and the putative CSTB2 protein is encoded from exons 1 and 2 and from intron 2 sequences containing a stop codon after 78 bp, creating a putative 83 amino acid protein. In Northern analysis, *CSTB2* was expressed at very low levels in all adult and fetal tissues studied (data not shown). In quantitative real-time PCR, human *CSTB2* was expressed in all fetal and adult tissues

examined, being highest in lung and thymus in fetal, and in lung in adult stages (IV, Fig. 4B). The proportion of *CSTB2* ranged from about 0.4% to approximately 5% of the expression of all *CSTB* transcripts, depending on the tissue examined (IV, Fig. 4C). The mouse *Cstb2* was PCR-amplified from mouse brain and testis with an open reading frame of 192 bp predicting a 64 amino acid Cstb2 protein (IV). *Cstb2* was expressed ubiquitously at low levels in all developmental stages and adult mouse tissues examined (Fig. 7). RNA *in situ* hybridization did not detect *Cstb2* mRNA expression in mice embryos, or in postnatal or adult mouse brain slices (data not shown), implying very low expression. These data show that *CSTB2* and *Cstb2* are both expressed ubiquitously at low levels.

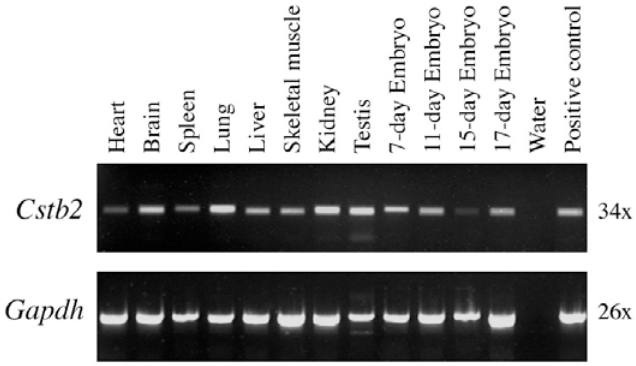


Figure 7. Semi-quantitative expression analysis of *Cstb2* in a Clontech mouse MTC panel, containing mouse adult tissues and whole mouse embryos at E7, E11, E15 and E17. The results at 34 cycles of amplification are shown. Mouse glyceraldehyde-3-phosphate dehydrogenase (*Gapdh*) gene expression at 26 cycles is shown as a control in the lower panel. *Cstb2* expression was detected in all adult mouse tissues and embryonal stages studies (unpublished).

3. Characterization of the CSTB protein (III)

3. 1. Production of a CSTB-specific monoclonal antibody

A monoclonal antibody against CSTB purified from human muscle and spleen was raised in study III. The 2E7 antibody recognized a protein of approximately 12 kDa, the calculated molecular weight of CSTB, in Western blots of COS-1 cell lysates and control lymphoblastoid cells (Fig. 8B; III, Fig. 1). The 2E7 antibody did not crossreact with the homologous CSTA protein (for protein sequence similarities with CSTB see Fig. 8A) in Western analysis (Fig. 8B; III, Fig. 1) or in ELISA (III, data not shown). Epitope mapping (III, data not shown) revealed amino acids 31 to 39 of the CSTB sequence as the epitope for the 2E7 antibody (Fig. 8A; III). The sequence in the corresponding region of CSTA is different (with the exception of one amino acid, Fig. 8A; III), giving proof of the specificity of the 2E7 antibody. In addition, in

Western blot analysis of lymphoblastoid cells of EPM1 patients homozygous for the repeat expansion mutation or compound heterozygous for the repeat expansion and c.202C>T nonsense mutations no CSTB protein was detected (Fig. 8B; III, Fig. 1), further proving the specificity of the 2E7 antibody.

A

```
MMCGAPSATQ PATAETQHIA DQVRSQLEEK ENKKFPVFKA VSFKSQVVAG TNYFIKVHVG
:  :  :    ::: : : :  : :  ::::: :       : :  : ::::: ::: :::  :
MIPGGLSEAK PATPEIQEIV DKVKPQLEEK TNETYGKLEA VQYKTQVVAG TNYYIKVRAG

DEDFVHLRVF QSLPHENKPL TLSNYQTNKA KHDELTYF    CSTB
:    :: ::  :::  :  :  :  ::  :  : :::: :
DNKYMHLKVF KSLPGQNEDL VLTGYQVDKN KDDELTGF    CSTA
```

B

Figure 8. Specificity of the monoclonal CSTB antibody. (A) CSTB is closely related to CSTA and the amino acid sequence identity between CSTB and CSTA is 53%. The epitope recognized by the monoclonal CSTB antibody 2E7 is underlined. (B) The 2E7 antibody detected CSTB in COS-1 cells (1) and control lymphoblastoid cells (2). CSTB protein was not detected in lymphoblastoid cells of an EPM1 patient homozygous for the repeat expansion mutation (3) or of a patient compound heterozygous for the repeat expansion and c.202C>T (p.Arg68X) mutations (4). The antibody recognized purified CSTB (5), but not purified CSTA (6). Picture B modified from publication III.

3. 2. Cellular localization of endogenous CSTB protein

CSTB has been thought to be a soluble cytoplasmic protein (Stubbs et al., 1990). Recent evidence had shown that within cells CSTB is localized in the nucleus and in cytoplasmic granular structures, representing either the Golgi complex or lysosomes (Calkins et al., 1998; D'Amato et al., 2000; Riccio et al., 2001). We further aimed to determine the cellular compartments where endogenous CSTB resides using the novel monoclonal 2E7 antibody and organelle marker antibodies in immunofluorescence analysis.

In addition to cytoplasm, endogenous CSTB protein was found to reside in two intracellular compartments in COS-1 cells and human primary myoblasts (Fig. 9A-D; III, Fig. 2). To optimally visualize CSTB in different cellular pools detergents with different permeabilization properties were used. Triton X-100 was used to visualize both nuclear and cytoplasmic pools, and saponin the cytoplasmic pool, as it does not permeabilize the nuclear envelope. In COS-1 cells and proliferating myoblasts, CSTB localized prominently to the nucleus (Fig. 9A; III, Fig. 2A, B). In

addition, CSTB was visible diffusely and in punctate structures in the cytoplasm (Fig. 9C; III, Fig. 2D, E). In differentiated myotubes, CSTB was localized mainly in the cytoplasm and excluded from the nucleus (Fig. 9B; III, Fig. 2C). In COS-1 cells and myoblasts numerous punctate cytoplasmic structures were observed (Fig. 9C; III, Fig. 2D, E), while in myotubes the number of these structures was decreased (Fig. 9D; III, Fig. 2F). On the basis of these results, it seems that the subcellular distribution of CSTB is dependent on the differentiation status of the cell. This is in agreement with previous findings in which CSTB was found in the nucleus in proliferating neuronal cells and in the cytoplasm in differentiated human neuroblastoma and rat primary cerebellar granule cells (Riccio et al., 2001).

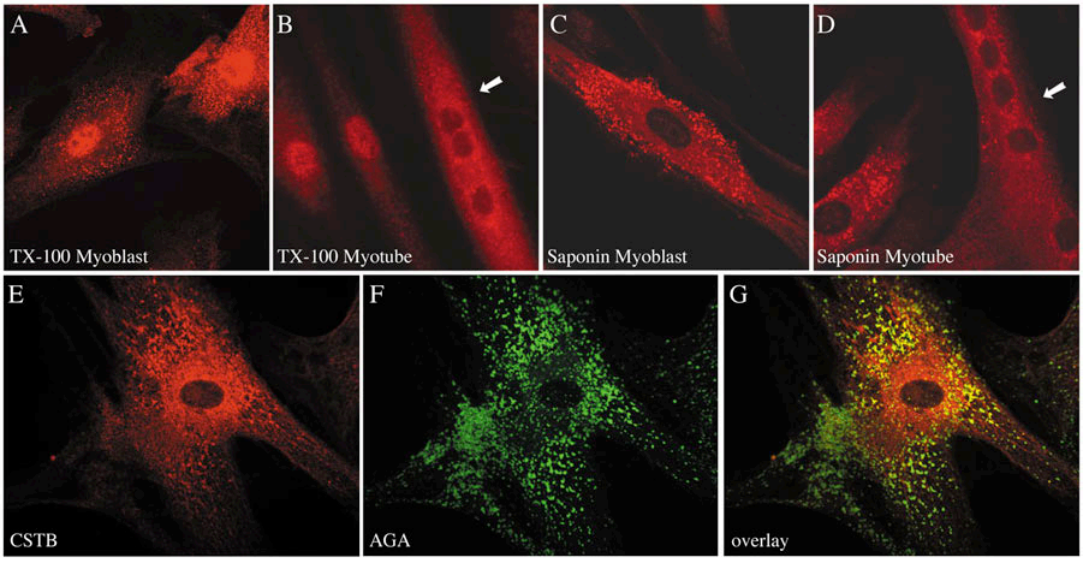


Figure 9. Subcellular localization of endogenous CSTB protein. Cells were permeabilized with TX-100 (A, B) or saponin (does not permeabilize the nuclear envelope) (C-G). In myoblasts CSTB localizes to the nucleus, cytoplasm and punctate cytoplasmic structures (A, C). In differentiated myotubes (shown by arrows) CSTB is localized to the cytoplasm and excluded from the nucleus and punctate cytoplasmic structures (B, D). CSTB-positive punctate cytoplasmic structures colocalized with the lysosomal marker AGA (E-G). Picture modified from publication III.

To identify the CSTB-positive punctate cytoplasmic structures different organelle markers were used. No overlap was observed with CSTB and markers for the endoplasmic reticulum (PDI), ER exit site (mSec13), Golgi (Lentil lectin), *trans*-Golgi network (TGN46) or late endosomes (MPR) (III, Fig. 3A-E). Instead, the CSTB protein displayed significant overlap with the lysosomal enzyme, aspartylglucosaminidase (AGA) (Fig. 9E-G; III, Fig. 3F). Many but not all AGA-positive structures stained with the CSTB antibody, indicating that CSTB associates only with a subset of lysosomes. In COS-1 cells similar colocalization of endogenous CSTB was observed, as CSTB colocalized with the AGA marker but not with other

cellular markers (III, data not shown). The lysosomal association of CSTB is in line with the current view of CSTB having a protective role as a cytoplasmic inhibitor of cysteine proteases released from lysosomes, and this view was further supported by the decreased cerebellar granule cell apoptosis detected in *Cstb*- and *Ctsb*-double-deficient mice (Houseweart et al., 2003a). CSTB could also be involved in signaling events involved in e.g. apoptosis with downstream effects on transcriptional regulation, because it was found to localize both in the cytoplasm and nucleus in addition to lysosomes.

3. 3. Biochemical properties of the CSTB protein

Whether the CSTB protein localizes inside or outside of lysosomes, and whether it is transported out of the cells, were investigated with immunoprecipitation of *in vitro* translated CSTB or endogenous CSTB from COS-1 cells.

In vitro translated CSTB was not enriched in microsomes, as was shown when the membranes were added co- or posttranslationally, but instead remained largely in the supernatants (III, Fig. 4A). This indicates that CSTB was unable to translocate into microsomal membranes. As a control, we used the soluble lysosomal enzyme (AGA) and an integral membrane protein (p58/ERGIC-53) that, as expected, translocated through or integrated in microsomal membranes, respectively (III, data not shown). In Western blot analysis CSTB could be immunoprecipitated from the COS-1 cell lysate but not from the culture medium (III, Fig. 4B), indicating that endogenous CSTB is not secreted in significant amounts in this cell type. CSTB was not detected in the cellular compartments of the early secretory pathway even when overexpressed (III, data not shown). These results, together with the fact that CSTB lacks an ER signal sequence, suggest that CSTB is attached to the outer side of the lysosomal membrane rather than residing within the lysosome. This hypothesis should be confirmed experimentally, for example by immunoelectronmicroscopy. However, it is tempting to speculate that the localization of CSTB to the lysosomal membrane is mediated through binding to cathepsins that are leaking out of the lysosomes. This speculation gets further support from the analysis of missense mutant CSTB proteins (see below, section 4.2) that are predicted to disturb cathepsin binding and that lose the lysosomal association.

4. *CSTB* gene and CSTB protein expression in EPM1 patients

4. 1. Effect of EPM1 mutations on *CSTB* gene expression (I, IV, unpublished)

To study the effect of the dodecamer repeat expansion on the expression of *CSTB*, three different approaches were used. First, we evaluated the promoter activity *in vitro* using promoter constructs containing wild-type and expanded alleles. Second, using Northern blot analysis we studied the mRNA levels in lymphoblastoid cells of individuals with "premutation" repeat expansions and of EPM1 patients homozygous for the full length repeat expansion. Third, using quantitative real-time PCR analysis we measured the mRNA levels in lymphoblastoid cells and fibroblasts of patients homozygous for the repeat expansion as well as compound heterozygous for the repeat expansion and some other EPM1 mutation.

In the luciferase reporter gene assays the addition of 16 dodecamer repeats to the luciferase-CSTB construct 1 (Fig. 4, section 2.1), creating construct 4, lowered its activity 10-fold in COS-1 cells (I, Fig. 2). This indicates that the repeat expansion significantly reduces the normal transcription *in vitro*. We were unable to obtain the "disease-length" expansions in the promoter reporter constructs, due to the unstable nature of the dodecamer repeat expansion (Lalioti et al., 1997b; Virtaneva et al., 1997). We could generate only a "premutation" length expansion construct containing 19 repeats (in EPM1 patients more that 30 copies have been reported), however, already this construct showed a clear reduction of luciferase expression. In another study, Lalioti et al. (1999) created a longer 3.2 kb promoter construct containing 50 dodecamer copies. They noted that cloning of the expanded promoters in commercially available luciferase plasmids resulted in instability, but the expansions were stable in a pbluescript vector. Interestingly, using this construct Lalioti et al. (1999) found only a two- to fourfold reduction of luciferase activity in SK-N-BE and HeLa cells. That we detected a 10-fold reduction in luciferase expression could be partly explained by our significantly shorter construct rather than by the difference in the expansion lengths, as it would be quite unexpected to see higher transcription with longer repeat expansions. They also showed reduced promoter activity due to the expansion in some cell types (Lalioti et al., 1999). The luciferase activity was either only slightly decreased or increased in CHP and COS-7 cells. Concordant with their previous finding on *CSTB* mRNA expression (Lalioti et al., 1997b) (discussed below), they suggested that the regulation of *CSTB* expression may be cell specific.

In Northern blot analysis CEPH individuals with "premutation" length expansions (Lalioti et al., 1997b) of 12 to 17 copies in both alleles had clearly reduced *CSTB* mRNA levels compared to controls, as did carriers with one normal and one expanded allele (I, Fig. 6). We further analyzed with quantitative real-time PCR the expression levels in these CEPH individuals. *CSTB* levels were consistently reduced in individuals with "premutation" expansions in both alleles, but not in individuals with one normal and one "premutation" allele (data not shown). These individuals have not been reported to have symptoms compatible with EPM1, suggesting that the development of the clinical signs of EPM1 depend on a threshold of *CSTB* expression. However, it should be noted that several of these Venezuelan CEPH individuals have a diagnosis of Huntington's disease, characterized by chorea and dementia, which could "mask" signs of EPM1 if not specifically looked for. It is also possible that the EPM1 symptoms could have a later onset or be milder than with the full length expansion.

The effect of the EPM1 mutations on *CSTB* mRNA expression levels was further studied with quantitative real-time RT-PCR from lymphoblastoid cell RNA of patients homozygous or compound heterozygous for the repeat expansion. In homozygous patients *CSTB* expression was markedly reduced, the mean expression (± standard deviation) being 8.8% ± 1.0% of the controls (IV, Fig. 5A). *CSTB* expression was also similarly reduced compared to controls in fibroblasts of two homozygous patients (IV, data not shown). The heterozygous expansion mutation carriers had 60.6% ± 12.7% of the *CSTB* expression in controls (IV, Fig. 5A). Compound heterozygotes for the expansion and the c.67-1G>C splice site mutations the level of *CSTB* was 17.9% ± 3.8% that of the controls (IV, Fig. 5A). The RT-PCR primers were designed so that the mutant c.67-1G>C mRNA was not amplified and the higher expression levels compared to those in patients with homozygous expansion mutations suggest that the splicing site is leaking (IV). Such partial penetrance mutations have been reported earlier (Svenson et al., 2001), and that possibility in relation to this mutation has also been discussed (Bespalova et al., 1997b). In a patient compound heterozygous for the expansion and the c.202C>T nonsense mutations, and another for the expansion and the novel c.149G>A missense mutations, the *CSTB* mRNA levels were 35.4% ± 4.5% and 15.6% ± 3.8% of controls, respectively (IV, Fig. 5A).

Lalioti et al. (1997b) have shown, using an RNase protection assay, that *CSTB* was markedly reduced in blood leukocytes (10-20%) whereas it was either normal or only slightly reduced in lymphoblastoid cells or fibroblasts of EPM1 patients carrying the expansion. These results are discrepant with our quantitative real-time PCR (IV) and the Northern blot results (Pennacchio et al., 1996; Lafrenière et al., 1997; Lalioti et al., 1997b; I), unequivocally showing that all mutations analyzed down-regulate the expression of *CSTB* mRNAs. Both our *in vitro* promoter study (I) and quantitative real-time PCR (IV) as well as the RNase protection assay in blood leukocytes (Lalioti et al., 1997b) imply that *CSTB* mRNA expression is approximately 10% of normal in patients homozygous for the expansion mutation. The contradictory findings of the RNase protection assay in other cell lines are intriguing (Lalioti et al., 1997b), but may be simply due to a more complicated experimental set-up in the assay.

In the *CSTB* promoter region three Sp1 binding sites, part of the dodecamer repeats, interacted strongly *in vitro* with the transcriptional Sp1 activator proteins (I, Fig. 4). Even if each dodecamer unit contains one Sp1 binding site and an expansion of these repeats results in an increase of these putatively activating sites, the consequence of the expansion is down-regulation of transcription (Lalioti et al., 1999; I). The explanation of this paradox might be that the changes in the chromatin structure result in disturbed transcriptional machinery or the expansion might be more prone to bind inhibitory proteins such as Sp3, also known to bind to GC-motifs (Hagen et al., 1994; see section 6 in the REVIEW OF THE LITERATURE). On the contrary, Lalioti et al. (1999) speculated that the different spacing of transcription elements could cause the reduction of *CSTB* transcription and tested their hypothesis by substituting the repeat with heterologous fragments of bacterial and mammalian sequences of similar sizes and showed similar reduction to that observed with the dodecamer repeat expansion. The dodecamer repeat could also be methylated, which has not been directly determined, but the *CSTB* promoter region is not known to be significantly hypermethylated (Lalioti et al., 1997b; Virtaneva et al., unpublished data). Our primary hypothesis of altered chromatin structure due to the repeat expansion as the underlying mechanism for altered *CSTB* expression got further support from the studies of Saha and Usdin (2001), who demonstrated that the dodecamer repeats form a variety of secondary structures and are unique in that tetraplexes are the only structures likely to form under physiological conditions (Fig.

2, section 3.3.1 in the REVIEW OF THE LITERATURE). Such stable secondary structures have been shown to affect transcription and even translation in association with trinucleotide repeat expansions (Feng et al., 1995; Parsons et al., 1998). Therefore, it is also likely that these structures are involved in affecting the transcription of *CSTB*.

4. 2. Altered localization of mutant CSTB proteins (III, IV, unpublished)

CSTB mRNA expression was known to be reduced due to the repeat expansion mutation (Pennacchio et al., 1996; Bespalova et al., 1997a; Bespalova et al., 1997b; Lafrenière et al., 1997; Lalioti et al., 1997b; I; IV). We further investigated the consequence of four previously reported and two novel EPM1 mutations on the localization and stability of the CSTB proteins using *in vitro* translation (IVT), immunofluorescence staining of exogenously expressed proteins and Western blot analysis of endogenous CSTB in patient cells.

First we investigated whether all of the mutant CSTB proteins were recognized with the novel monoclonal 2E7 antibody and/or with the polyclonal CSTB antibody. For that purpose we performed radioactive IVT using expression vectors of the mutant proteins and immunoprecipitation with the 2E7 antibody or polyclonal CSTB antibody. The wild-type, and all three missense mutant proteins, p.Gly4Arg, p.Gly50Glu and p.Gln71Pro, were immunoprecipitated with the monoclonal 2E7 antibody (Fig. 10). In contrast, the truncated p.Arg68X and p.Leu73fsX3 proteins could not be immunoprecipitated with the 2E7 antibody (Fig. 10) even though they contained the epitope recognized by the antibody (Fig. 8, section 3.1), indicating a change in the conformation of the epitope. The polyclonal CSTB antibody immunoprecipitated all *in vitro* translated CSTB proteins (data not shown), except the p.delVal23_Lys56 protein, which was not detected even though loaded directly from the IVT reaction to the gel (data not shown).

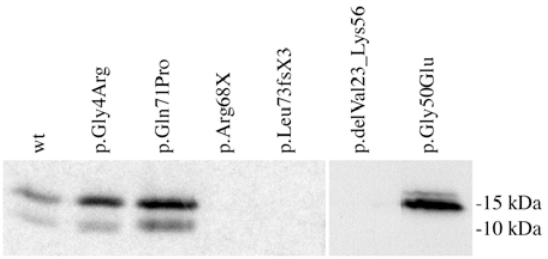


Figure 10. *In vitro* translated [^{35}S] methionine labelled wild-type (wt), p.Gly4Arg, p.Gln71Pro and p.Gly50Glu CSTB proteins were immunoprecipitated with the 2E7 antibody. The gradient 8-16% SDS-PAGE gel (Bio-Rad) resolving proteins from 6 to 70 kDa was used to detect also the predicted truncated (p.Arg68X, p.Leu73fsX3 and p.delVal23_Lys56) 7 kDa proteins (unpublished). The reason for CSTB proteins (III, Fig. 4A, B) migrating sometimes as doublets in the SDS-PAGE is not known.

We then studied the subcellular localization of overexpressed wild-type and mutant CSTB proteins in baby hamster kidney (BHK) cells, since no significant crossreactivity with the hamster protein was observed with the CSTB antibodies used. Lysosomes were visualized in the first study with LysoTracker Green (III) and in the second study with polyclonal lpg tail 120 antibody recognizing the lysosomal LAMP-1 protein in hamster cells (IV). The overexpressed wild-type CSTB displayed similar distribution to that of endogenous CSTB, as it was detected in the nucleus, cytoplasm and lysosomes with both the monoclonal 2E7 (Fig. 11A; III, Fig. 5A; IV, Fig. 6A) and polyclonal CSTB antibodies (Fig. 13A, section 5). The p.Gly4Arg (Fig. 11B; III, Fig. 5B), p.Gln71Pro (Fig. 11C; III, Fig. 5C) and p.Gly50Glu (Fig. 11D; IV, Fig. 6B) missense mutant proteins were diffusely localized to the cytoplasm and nucleus, but no punctate structures overlapping with the lysosomal markers were observed. The truncated p.Leu73fsX3 protein was seen diffusely in the cytoplasm and nucleus with no colocalization with the lysosomal marker (III, Fig. 5D). The p.Arg68X protein was only detectable when expressed in the presence of lactacystin, a proteasomal inhibitor, indicating rapid degradation of the protein (III, Fig. 5E). In conclusion, all mutant CSTB proteins studied failed to associate with lysosomes, implying that the association of CSTB with lysosomes is of physiological importance and that loss of this association contributes to the molecular pathogenesis of EPM1. It is tempting to speculate that the lysosomal association of CSTB is mediated through cathepsin interactions as none of the mutant proteins predicted to alter cathepsin binding associate with lysosomes. All the missense mutations are known to be located in the CSTB protein in such places that mutations are suspected to alter the papain binding affinity (discussed in section 3.3.2 in the REVIEW OF THE LITERATURE, and in section 1 in RESULTS AND DISCUSSION).

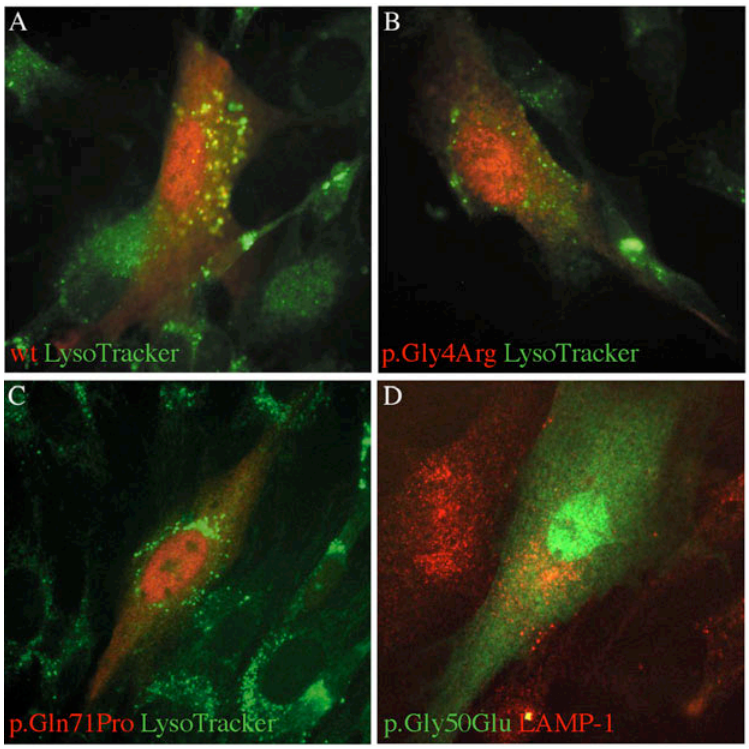


Figure 11. Subcellular localization of overexpressed wild-type (wt) and mutant CSTB proteins. (A) Transiently expressed wild-type CSTB localizes similarly to that of endogenous CSTB, to the nucleus, cytoplasm and lysosomes. The p.Gly4Arg (B), p.Gln71Pro (C) and p.Gly50Glu (D) mutant proteins are localized to the nucleus and cytoplasm, but no overlap with the lysosomal markers LysoTracker or LAMP-1 is detected. Pictures modified from publications III and IV.

Finally, we investigated using Western blot analysis whether the mutant CSTB proteins are present in lymphoblastoid cells of EPM1 patients. In the first study the 2E7 antibody was used in Western blot analysis (Fig. 8B, section 3.1; III, Fig. 1). In the second study we used the polyclonal CSTB antibody (IV, Fig. 5B), since it was noted that in immunofluorescence staining (III, data not shown) and immunoprecipitations (Fig. 10) that the 2E7 antibody did not recognize the truncated CSTB proteins. In an EPM1 patient homozygous for the expansion mutation the CSTB protein amount was significantly decreased (Fig. 8B; III, Fig. 1; IV, Fig. 5B), concordant with the reduction of mRNA expression (Pennacchio et al., 1996; Bespalova et al., 1997a; Bespalova et al., 1997b; Lafrenière et al., 1997; Lalioti et al., 1997b; I; IV). In a patient compound heterozygous for the repeat expansion mutation and c.202C>T mutation, predicting a truncated p.Arg68X protein, the protein amount was clearly decreased (Fig. 8B; III, Fig. 1; IV, Fig. 5B). Interestingly, the mRNA levels (35.4% ± 4.5%) detected in this patient were among the highest in patients (IV, Fig. 5A), while the protein amount (IV, Fig. 5B) was among the lowest, giving further proof that the p.Arg68X protein is also rapidly degraded *in vivo*. Using a circular

dichroism spectra, the p.Arg68X protein has been shown not to be folded *in vitro* like the wild-type CSTB protein (Rabzelj et al., 2005). In addition, the unfolded p.Arg68X protein was highly prone to aggregate and form amyloid fibrils (Rabzelj et al., 2005), laying the basis for the suggestion that fibril formation might contribute to the pathogenesis of EPM1 (Ceru et al., 2005). This hypothesis is discrepant with our *in vitro* and *in vivo* results (Fig. 8B, section 3.1; III, Fig. 1; III, Fig. 5E; IV, Fig. 5B), and it is likely that the p.Arg68X protein is degraded before it aggregates or forms amyloid fibrils. In contrast, the p.Gly4Arg protein was shown to be folded like the wild-type CSTB protein (Rabzelj et al., 2005). Unfortunately we did not have an opportunity to measure the mRNA or protein levels of this Moroccan EPM1 patient since no sample was available. In the patient compound heterozygous for the repeat expansion and the c.149G>A mutation, resulting in a p.Gly50Glu protein, a slight decrease in the amount of the CSTB protein, attributable to the reduced expression from the expansion allele, was observed (IV, Fig. 5B). The amino acid changes in the QVVAG pentapeptide have been speculated to affect the stability or life span of the inhibitory effect of the protein, or the interactions with their target proteins (Jerala et al., 1990; Auerswald et al., 1995). The p.Gly50Glu protein stability seems not to be affected (Fig. 11D; IV, Fig. 5B; IV, Fig. 6B), while the interactions with the target proteins might be affected since at least association with lysosomes has been lost (Fig. 11D; IV, Fig. 6B). In the EPM1 patient compound heterozygous for the repeat expansion and the c.67-1G>C mutation, resulting in an in-frame deletion of the entire exon 2 and a 34 amino acid shorter p.delVal23_Lys56 protein, no protein was detected (IV, Fig. 5B). The p.delVal23_Lys56 protein was also not detected when loaded directly from the IVT reaction to the gel (data not shown) or in immunofluorescence staining even in the presence of lactacystin, the proteasomal inhibitor (IV, data not shown). These findings indicate that the p.delVal23_Lys56 protein is highly unstable.

5. Characterization of the CSTB2 protein (IV, unpublished)

The CSTB2 protein, putatively translated from the *CSTB2* splice variant, is predicted to be a 9 kDa protein of 83 amino acids, including 57 identical amino acids to CSTB followed by 26 amino acids encoded by the sequence of intron 2 (IV). *CSTB2* was expressed at very low levels compared to *CSTB* in human tissues (IV, discussed in

section 2.3). The *CSTB2* mRNA levels in lymphoblastoid cells of EPM1 patients and controls could not be reliably measured due to inter-assay variation (IV). Instead, in the fibroblasts of two patients homozygous for the repeat expansion mutation similarly reduced levels of *CSTB* and *CSTB2* compared to controls were detected, implying that the expression is regulated by the same promoter (IV, data not shown). The biological properties of CSTB2 were investigated to evaluate the potential role of this splice variant in EPM1 pathogenesis.

We first tried to investigate whether *CSTB2* is translated *in vivo*. CSTB2 protein was not visible, due to the expected small amount compared to CSTB, in Western blot analysis with the 2E7 antibody or polyclonal CSTB antibody (data not shown). Furthermore, despite repeated attempts a specific antibody for CSTB2, not recognizing CSTB, could not be produced. Thus the presence of endogenous CSTB2 could not be verified, and we therefore decided to study the stability of *in vitro* translated CSTB2. *In vitro* translated wild-type, p.Gly4Arg and p.Gly50Glu CSTB2 proteins were stable and could be immunoprecipitated with the polyclonal CSTB antibody (Fig. 12), but not with the 2E7 antibody (data not shown).

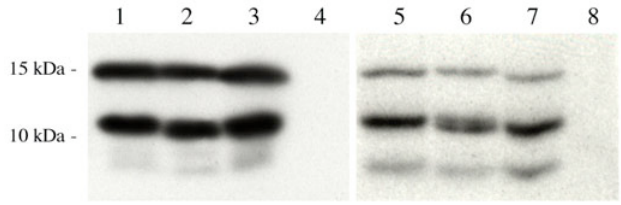


Figure 12. *In vitro* translations of CSTB2. SDS-PAGE gels of *in vitro* translated wild-type (1), p.Gly4Arg (2) and p.Gly50Glu (3) CSTB2 proteins. IVT followed by immunoprecipitation using the polyclonal CSTB antibody: wild-type (5), p.Gly4Arg (6) and p.Gly50Glu (7) proteins are immunoprecipitated in equal amounts. An empty pcDNA3.1(+) expression vector was used as a negative control (4, 8) (unpublished). The reason for CSTB (Fig. 10; III, Fig. 4A, B) and CSTB2 proteins migrating as doublets in the SDS-PAGE is not known.

We then studied the subcellular localization of overexpressed CSTB2 using immunofluorescence analysis with the polyclonal CSTB antibody. First, the overexpressed CSTB protein was shown to localize with the polyclonal CSTB antibody (Fig. 13A), similarly to with the monoclonal 2E7 antibody (Fig. 11A; III, Fig. 5A; IV, Fig. 6A), in the nucleus, cytoplasm and lysosomes. Wild-type (Fig. 13B) and p.Gly4Arg (Fig. 13C) CSTB2 proteins localized evenly to cytoplasm and nucleus, while no lysosomal punctate structures were visible. The effect of the p.Gly50Glu mutation on CSTB2 protein localization was dramatic as it aggregated to the

cytoplasm or nucleus (Fig. 13D, E). Colocalization with several aggresomal markers was attempted but no overlap could be visualized (data not shown).

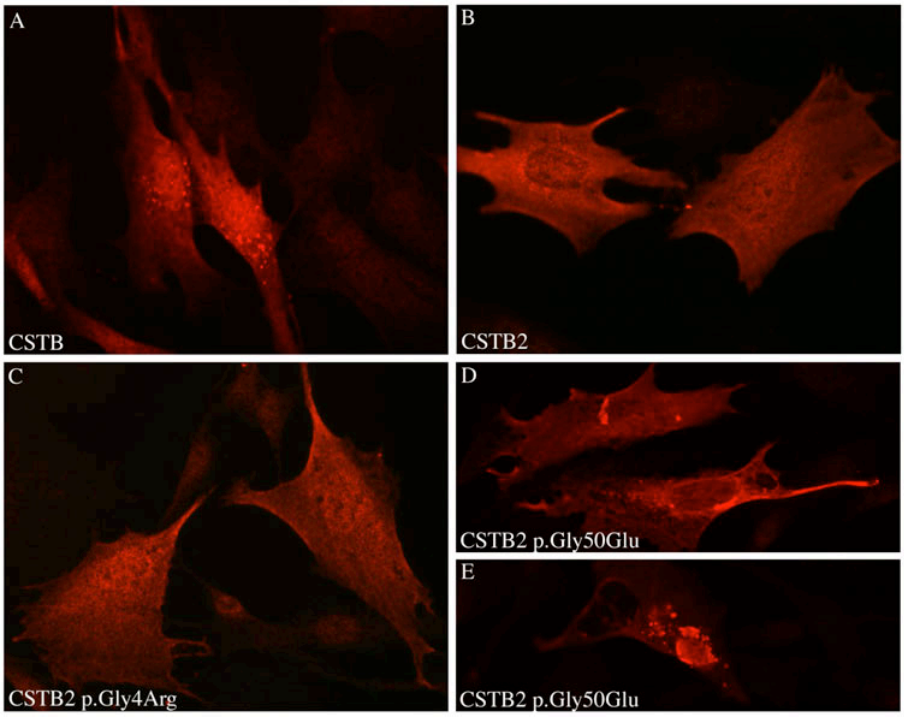


Figure 13. Immunofluorescence analysis of the cellular distribution of wild-type CSTB and CSTB2, and mutant CSTB2 proteins in transiently transfected BHK cells immunostained with the polyclonal CSTB antibody. (A) CSTB localizes to nucleus, lysosomes and cytoplasm. Wild-type (B) and p.Gly4Arg (C) CSTB2 proteins localize evenly to cytoplasm and nucleus. p.Gly50Glu mutant CSTB2 protein aggregates to cytoplasm (D) or nucleus (E) (unpublished).

In conclusion, the existence of endogenous CSTB2 protein could not be proven, while the overexpressed CSTB2 protein was stable and localized to cytoplasm and nucleus, but not to lysosomes. Whether CSTB2 is translated *in vivo* remains unknown. The *CSTB2* transcript is observed in multiple expressed sequence tags (ESTs) from cDNA libraries implying that it has physiological relevance. Moreover, it is also present in mouse EST databases. *CSTB2* is spliced following the GT-AG rule and is conserved among humans and mice. The mRNA expression of *CSTB2* and *Cstb2* was detected at very low levels with various methods. *CSTB2* could have a regulatory function on *CSTB* expression at the post-transcriptional level by inhibiting the splicing of the pre-mRNA or by representing a storage form of the immature transcript. It also remains possible that it is translated into a protein that has distinct functions from that of CSTB. EPM1 patients also had similarly decreased expression of *CSTB2* than that of *CSTB*. The relevance of this ubiquitously expressed *CSTB2* transcript in normal physiological and EPM1 pathological conditions remains to be investigated.

CONCLUSIONS AND FUTURE PROSPECTS

The work aim within our research group is understanding the molecular pathogenetic mechanisms that underlie the progressive myoclonus epilepsy of Unverricht-Lundborg type (EPM1) caused by mutations in the *CSTB* gene. CSTB, a cysteine protease inhibitor, has been considered to inhibit inappropriate proteolysis due to cathepsins that leak out of the lysosomes, but its physiological function is still unknown. In this study, the *CSTB* gene was characterized, and expression and localization studies of CSTB mRNA and protein were initiated.

The CSTB promoter was shown to have a typical housekeeping gene promoter sequence and to contain several active transcriptional binding sites. However, *Cstb* was shown to be widely expressed throughout mouse embryonic development albeit with tight temporal regulation and a restricted tissue and cellular expression pattern. *Cstb* expression was highest in E7 embryos and in epithelial layers of several tissues outside the nervous system even though no developmental or consistent non-neuronal phenotypes have been detected in EPM1 patients or the *Cstb*-deficient mouse model. This suggests that the Cstb inhibitory activity is functionally redundant during development and in most tissues, compatible with Cstb being a member of a large family of cysteine protease inhibitors. The high *Cstb* expression in thymus is intriguing in light of the suggested function of cathepsins in immune response and the proposed role of the leech CSTB homologue in innate immunity. The limited cellular distribution of *Cstb* in the CNS is in accordance with the clinical presentation and limited neuronal pathology detected in EPM1 patients and the *Cstb*-deficient mouse. It is crucial to continue the investigations of Cstb protein expression in the CNS, the major organ affected in EPM1, as well as its neuronal metabolism. The *Cstb*-deficient mouse provides an excellent model for studying the function of *Cstb* in *in vivo* situations. To continue the investigations a specific antibody for the mouse Cstb protein is needed.

The lysosomal localization of CSTB demonstrated in this study is compatible with the current view of CSTB having a protective role as a constitutive cytoplasmic inhibitor for cysteine proteases released from lysosomes. Whether the lysosomal association is due to cathepsin interaction, which would be suggested based on our localization studies with mutant CSTB proteins, remains to be further studied. This hypothesis is supported by the decreased cerebellar granule cell apoptosis detected in

mice doubly deficient for *Cstb* and *Ctsb*. The ataxia and seizure phenotypes in these mice were not dimished, suggesting that CSTB might have functions other than inhibition of cathepsins. CSTB localized also to cytoplasm and nucleus, suggesting that it could also function for example in signaling events involved in apoptosis with downstream effects on transcriptional regulation. CSTB localization was also dependent on the differentiation status of the cells, as the subcellular distribution of the CSTB was altered upon cell differentiation. The distribution of CSTB should be determined in several cell types under different conditions, especially in neuronal cells.

The repeat expansion decreased the *CSTB* promoter activity *in vitro*, concordantly significantly reduced amounts of CSTB mRNA and protein were detected in patient cells. In line with this, as the expansion mutation significantly reduces the amount of CSTB protein capable of inhibiting cathepsins, cathepsins B, L and S activities are increased. In addition to the repeat expansion mutation, the other EPM1 mutations were also shown to decrease *in vivo* the CSTB mRNA and/or protein levels. All mutant proteins lost lysosomal association, highlighting the importance of this association for the normal physiological and pathological function of CSTB. It remains speculative if this association is a consequence of altered cathepsin inhibitory function.

The *CSTB* gene was shown to be alternatively spliced, but the physiological relevance of this remains unknown.

More studies have to be conducted to clarify the cellular function of CSTB and the related molecular mechanism of neuronal survival and function, as well as the molecular pathogenesis of EPM1. Understanding the normal physiological and pathogenetic mechanism of CSTB will be essential for the discovery of new drugs and better treatment methods for EPM1 in the future.

ACKNOWLEDGEMENTS

This study was carried out at the Folkhälsan Institute of Genetics, Department of Medical Genetics and Neuroscience Center, University of Helsinki. The research project was part of the Centre of Excellence in Disease Genetics (2000-2005) and the Centre of Excellence in Complex Disease Genetics (2006-2011) of the Academy of Finland. This study was financially supported by the Academy of Finland (projects 43029, 44870 and 50011), the European Commission Fifth Framework Programme "Cystatin B in Epilepsy" (project QLG3-CT-2000-01405), the Sigrid Jusélius Foundation, the Folkhälsan Research Foundation and was part of the Life 2000 consortium on Functional Genomics of Finnish disease genes. I had the opportunity to be a student at the Helsinki Biomedical Graduate School (matching funded, 2002-2005). I wish to express my sincere gratitude to all those who made this thesis possible, especially to:

The former and present heads of the Institutes, *Albert de la Chapelle*, *Juha Kere*, *Leena Peltonen-Palotie*, *Pertti Aula*, *Anna-Elina Lehesjoki*, *Kristiina Aittomäki*, *Päivi Peltomäki* and *Heikki Rauvala*, for providing me with excellent working facilities.

Anna-Elina Lehesjoki, my supervisor, for guiding me through this thesis project, for her never-ending support and encouragement as well as for her interest and kindness.

My co-authors, *Tarja Joensuu*, *Ulla Lahtinen* and *Kimmo Virtaneva*, for teaching me a lot and for their huge contribution. *Nina Aula* for always being willing to help and for her easy-going attitude. *Outi Kopra* for helping especially with MCID image analysis. *Saara Tegelberg*, with whom I have made many things in the lab, experienced ups and downs, and enjoyed our time together. *Paula Hakala* for excellent technical assistance, it has been a pleasure working with you. *Mervi Kuronen* for all Taqman runs and for her bright attitude. *Laura Huopaniemi* for sharing her knowledge about real-time PCR with Mervi and Tarja. *Teija-Tuulia Toivonen* for skillful cell culturing and for helping whenever needed. *Jorma Palvimo* for his valuable contribution in the promoter study and for always being willing to advise. *Ekkehard Weber* for producing monoclonal 2E7 antibody, without it the protein study would have been impossible. *Gerit Theil* for skillful contribution in the antibody production and for nice conversations. *Riitta Rinne* and *Pekka Saukko*, for purifying human CSTB protein and for nice collaborations. *Mart Saarma, Antti Aalto*, *Marjo Heikura* and *Deying Guo,* for valuable collaboration with the radioactive RNA *in situ* hybridization study. *Kirsi Sainio* for her huge experienced contribution in the RNA *in situ* hybridization study. *Virpi Syvälahti* and *Jetta Kelppe* for technical assistance in whole-mount *in situ* hybridization. *Satu Kuure* for Ventana trials and for helping me with microscopes etc. in the Sariola lab. *Anna Vihola* for advice on immunohistochemistry and for trying to cut cryosections. *Gerrit-Jan de Haan*, *Dick Lindhout* and *Dicky Halley*, for providing the novel missense mutation. *Eva Reinmaa* for being one of us in the EPM1 group. I give the EPM1-project willingly to the next generation, since there are enthusiastic Ph.D. students *Saara Tegelberg* and *Mervi Kuronen* (with the help of others) to continue the project.

Jaana Tyynelä-Vesterinen and *Ismo Ulmanen*, the official reviewers of this thesis, for their valuable comments. *Anu Jalanko* and *Anu Wartiovaara*, for being my thesis committee members.

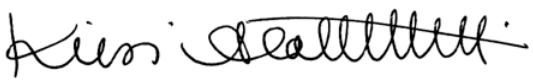
Jodie Painter for the language revision of this thesis, and for her interest and encouragement.

All the present and former members of the Folkhälsan Research Institute for friendliness and for good company. *Riikka Hämäläinen*, *Eija Siintola* and *Saara Tegelberg*, for being good friends, for sharing the joys and sorrows of life. *Juha Kolehmainen* for sharing the office room during almost all these years and helping me whenever needed. *Liina Lonka* and *Jukka Kallijärvi*, for all scientific discussions and for being my friends. *Kati Donner* for her joyful attitude towards life and for nice conversations. *Anna-Kaisa Anttonen* and *Anna Vihola* for all our nice conversations and helpful advice. *Ann-Liz Träskelin* and *Tarja Joensuu* for their maternal caring attitudes. *Juha Isosomppi* for willingness to help me for finding solutions to problems. *Tarja Salonen* for skillfull comments and for her fresh attitude. *Liisa Myllykangas* for helping me with immunohistochemistry problems and for nice discussions. *Anne Saarinen* and *Vilma-Lotta Lehtokari* for refreshing my daily life. *Laura Huopaniemi* and *Anne Rantala* for tutoring me during the first year in Folkhälsan. *José Dieguez-Lucena*, *Kristiina Avela*, *Maria Aminoff-Backlund* and *Susanna Ranta* for educational discussions and for good spirit. *TJ* and *PPLU* for all nice home parties. *Peter Hackman* for his energetic attitude towards science. *Aila Riikonen* and *Jaana Welin-Haapamäki* for helping in so many practical matters and for arranging many nice parties. *Sinikka Lindh* for taking on a visit of an EPM1 patient. *Stephan Keskinen* for taking care of financial matters. *Merja Soininen*, *Hanna Olanne* and *Jaakko Sarparanta* for nice lunch company. *Herman's group*, especially *Satu Heinonen*, *Adile Samaletdin* and *Anja Koskela*, for nice conversations about life in general. All those whom I have travelled with in the Europe and the States for their nice company.

Pirkko-Liisa Saarikoski, my senior high school biology teacher, for fascinating lectures that inspired me about genetics.

My dear friend, *Noora Pyyry* for being there. *Noora* and *Joakim Pyyry* for reminding me that there is so much more in the world and life than science. *Carina Welling*, *Noora Reittu*, *Mira Lystimäki*, *Tytti Jaakkola*, *Mari Salunen*, *Marjo Sabelström-Ollila*, *Miia Hukari*, *Anu Heikkinen*, *Anne Kanerva*, *Maria Mecklin*, *Maarit Takatalo*, *Niina Idänheimo*, *Henna Tyynismaa* and *Riika Penttilä*, for being good friends. My goddaughter *Nea Salmijärvi* and her sister *Janina*, for being so lovely. My brothers, *Mikko* and *Sami Alakurtti*, for sharing our whole life, hopefully I have not given you too much pressure for your studies. *Sini Kinnunen* and *Sari Ellonen*, for refreshing our family as new members. I am looking forward to be an aunt and godmother for the newcomer in January. *Tommi Kokonaho* for supporting my "gradu"-project and for our cheerful life together, which has just begun. My parents, *Annikki* and *Matti Alakurtti*, to whom I have dedicated this thesis, for all the love and support that you have given me, your trust and encouragement have made everything possible. With all of you I have shared so many unforgettable moments, which have brought much joy and happiness to my life.

Helsinki, October 2006



REFERENCES

Abrahamson, M., Barrett, A.J., Salvesen, G. and Grubb, A.: Isolation of six cysteine proteinase inhibitors from human urine. Their physicochemical and enzyme kinetic properties and concentrations in biological fluids. J Biol Chem 261 (1986) 11282-9.

Abrahamson, M., Grubb, A., Olafsson, I. and Lundwall, A.: Molecular cloning and sequence analysis of cDNA coding for the precursor of the human cysteine proteinase inhibitor cystatin C. FEBS Lett 216 (1987) 229-33.

Abrahamson, M., Olafsson, I., Palsdottir, A., Ulvsback, M., Lundwall, A., Jensson, O. and Grubb, A.: Structure and expression of the human cystatin C gene. Biochem J 268 (1990) 287-94.

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., George, R.A., Lewis, S.E., Richards, S., Ashburner, M., Henderson, S.N., Sutton, G.G., Wortman, J.R., Yandell, M.D., Zhang, Q., Chen, L.X., Brandon, R.C., Rogers, Y.H., Blazej, R.G., Champe, M., Pfeiffer, B.D., Wan, K.H., Doyle, C., Baxter, E.G., Helt, G., Nelson, C.R., Gabor, G.L., Abril, J.F., Agbayani, A., An, H.J., Andrews-Pfannkoch, C., Baldwin, D., Ballew, R.M., Basu, A., Baxendale, J., Bayraktaroglu, L., Beasley, E.M., Beeson, K.Y., Benos, P.V., Berman, B.P., Bhandari, D., Bolshakov, S., Borkova, D., Botchan, M.R., Bouck, J., Brokstein, P., Brottier, P., Burtis, K.C., Busam, D.A., Butler, H., Cadieu, E., Center, A., Chandra, I., Cherry, J.M., Cawley, S., Dahlke, C., Davenport, L.B., Davies, P., de Pablos, B., Delcher, A., Deng, Z., Mays, A.D., Dew, I., Dietz, S.M., Dodson, K., Doup, L.E., Downes, M., Dugan-Rocha, S., Dunkov, B.C., Dunn, P., Durbin, K.J., Evangelista, C.C., Ferraz, C., Ferriera, S., Fleischmann, W., Fosler, C., Gabrielian, A.E., Garg, N.S., Gelbart, W.M., Glasser, K., Glodek, A., Gong, F., Gorrell, J.H., Gu, Z., Guan, P., Harris, M., Harris, N.L., Harvey, D., Heiman, T.J., Hernandez, J.R., Houck, J., Hostin, D., Houston, K.A., Howland, T.J., Wei, M.H., Ibegwam, C., et al.: The genome sequence of Drosophila melanogaster. Science 287 (2000) 2185-95.

Anderson, D.J., Groves, A., Lo, L., Ma, Q., Rao, M., Shah, N.M. and Sommer, L.: Cell lineage determination and the control of neuronal identity in the neural crest. Cold Spring Harb Symp Quant Biol 62 (1997) 493-504.

Andren, L., Dymling, J.F., Hogeman, K.E. and Wendeberg, B.: Osteopetrosis acro-osteolytica. A syndrome of osteopetrosis, acro-osteolysis and open sutures of the skull. Acta Chir Scand 124 (1962) 496-7.

Angel, P. and Karin, M.: The role of Jun, Fos and the AP-1 complex in cell-proliferation and transformation. Biochim Biophys Acta 1072 (1991) 129-57.

Annegers, J.F., Rocca, W.A. and Hauser, W.A.: Causes of epilepsy: contributions of the Rochester epidemiology project. Mayo Clin Proc 71 (1996) 570-5.

Auerswald, E.A., Nagler, D.K., Assfalg-Machleidt, I., Stubbs, M.T., Machleidt, W. and Fritz, H.: Hairpin loop mutations of chicken cystatin have different effects on the inhibition of cathepsin B, cathepsin L and papain. FEBS Lett 361 (1995) 179-84.

Awano, T., Katz, M.L., O'Brien D, P., Sohar, I., Lobel, P., Coates, J.R., Khan, S., Johnson, G.C., Giger, U. and Johnson, G.S.: A frame shift mutation in canine TPP1 (the ortholog of human CLN2) in a juvenile Dachshund with neuronal ceroid lipofuscinosis. Mol Genet Metab (2006a).

Awano, T., Katz, M.L., O'Brien, D.P., Taylor, J.F., Evans, J., Khan, S., Sohar, I., Lobel, P. and Johnson, G.S.: A mutation in the cathepsin D gene (CTSD) in American Bulldogs with neuronal ceroid lipofuscinosis. Mol Genet Metab 87 (2006b) 341-8.

Barrett, A.J.: The cystatins: a diverse superfamily of cysteine peptidase inhibitors. Biomed Biochim Acta 45 (1986) 1363-74.

Barrett, A.J.: The cystatins: a new class of peptidase inhibitors. Trends Biochem Sci 12 (1987) 193-196.

Barrett, A.J., Davies, M.E. and Grubb, A.: The place of human gamma-trace (cystatin C) amongst the cysteine proteinase inhibitors. Biochem Biophys Res Commun 120 (1984) 631-6.

Barrett, A.J., Fritz, H., Grubb, A., Isemura, S., Jarvinen, M., Katunuma, N., Machleidt, W., Muller-Esterl, W., Sasaki, M. and Turk, V.: Nomenclature and classification of the proteins homologous with the cysteine-proteinase inhibitor chicken cystatin. Biochem J 236 (1986a) 312.

Barrett, A.J., Rawlings, N.D., Davies, M.E., Machleidt, W., Salvesen, G. and Turk, V.: Inhibitors of cysteine proteases. In Proteinease inhibitors., ed. Barrett, A. J. and Salvesen, G. pp. 515-569. Elsevier, Amsterdam, The Netherlands, 1986b.

Bate, L. and Gardiner, M.: Molecular genetics of human epilepsies. Expert Rev Mol Med 1999 (1999) 1-22.

Bennett, S.T. and Todd, J.A.: Human type 1 diabetes and the insulin gene: principles of mapping polygenes. Annu Rev Genet 30 (1996) 343-70.

Bennett, S.T., Wilson, A.J., Cucca, F., Nerup, J., Pociot, F., McKinney, P.A., Barnett, A.H., Bain, S.C. and Todd, J.A.: IDDM2-VNTR-encoded susceptibility to type 1 diabetes: dominant protection and parental transmission of alleles of the insulin gene-linked minisatellite locus. J Autoimmun 9 (1996) 415-21.

Berkovic, S.F., Andermann, F., Carpenter, S. and Wolfe, L.S.: Progressive myoclonus epilepsies: specific causes and diagnosis. N Engl J Med 315 (1986) 296-305.

Berkovic, S.F., Mazarib, A., Walid, S., Neufeld, M.Y., Manelis, J., Nevo, Y., Korczyn, A.D., Yin, J., Xiong, L., Pandolfo, M., Mulley, J.C. and Wallace, R.H.: A new clinical and molecular form of Unverricht-Lundborg disease localized by homozygosity mapping. Brain (2005).

Bertrand, N., Castro, D.S. and Guillemot, F.: Proneural genes and the specification of neural cell types. Nat Rev Neurosci 3 (2002) 517-30.

Bespalova, I.N., Adkins, S., Pranzatelli, M. and Burmeister, M.: Novel cystatin B mutation and diagnostic PCR assay in an Unverricht-Lundborg progressive myoclonus epilepsy patient. Am J Med Genet 74 (1997a) 467-71.

Bespalova, I.N., Pranzatelli, M. and Burmeister, M.: G to C transversion at a splice acceptor site causes exon skipping in the cystatin B gene. Mutat Res 382 (1997b) 67-74.

Bindoff, L.A., Desnuelle, C., Birch-Machin, M.A., Pellissier, J.F., Serratrice, G., Dravet, C., Bureau, M., Howell, N. and Turnbull, D.M.: Multiple defects of the mitochondrial respiratory chain in a mitochondrial encephalopathy (MERRF): a clinical, biochemical and molecular study. J Neurol Sci 102 (1991) 17-24.

Blackwell, T.K. and Weintraub, H.: Differences and similarities in DNA-binding preferences of MyoD and E2A protein complexes revealed by binding site selection. Science 250 (1990) 1104-10.

Bonten, E., van der Spoel, A., Fornerod, M., Grosveld, G. and d'Azzo, A.: Characterization of human lysosomal neuraminidase defines the molecular

basis of the metabolic storage disorder sialidosis. Genes Dev 10 (1996) 3156-69.

Brömme, D., Okamoto, K., Wang, B.B. and Biroc, S.: Human cathepsin O2, a matrix protein-degrading cysteine protease expressed in osteoclasts. Functional expression of human cathepsin O2 in Spodoptera frugiperda and characterization of the enzyme. J Biol Chem 271 (1996) 2126-32.

Brömme, D., Rinne, R. and Kirschke, H.: Tight-binding inhibition of cathepsin S by cystatins. Biomed Biochim Acta 50 (1991) 631-5.

Bronson, R.T., Donahue, L.R., Johnson, K.R., Tanner, A., Lane, P.W. and Faust, J.R.: Neuronal ceroid lipofuscinosis (nclf), a new disorder of the mouse linked to chromosome 9. Am J Med Genet 77 (1998) 289-97.

Brook, J.D., McCurrach, M.E., Harley, H.G., Buckler, A.J., Church, D., Aburatani, H., Hunter, K., Stanton, V.P., Thirion, J.P., Hudson, T. and et al.: Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein kinase family member. Cell 68 (1992) 799-808.

Burgess-Beusse, B., Farrell, C., Gaszner, M., Litt, M., Mutskov, V., Recillas-Targa, F., Simpson, M., West, A. and Felsenfeld, G.: The insulation of genes from external enhancers and silencing chromatin. Proc Natl Acad Sci U S A 99 Suppl 4 (2002) 16433-7.

Butler, J.E. and Kadonaga, J.T.: Enhancer-promoter specificity mediated by DPE or TATA core promoter motifs. Genes Dev 15 (2001) 2515-9.

Calkins, C.C., Sameni, M., Koblinski, J., Sloane, B.F. and Moin, K.: Differential localization of cysteine protease inhibitors and a target cysteine protease, cathepsin B, by immuno-confocal microscopy. J Histochem Cytochem 46 (1998) 745-51.

Campuzano, V., Montermini, L., Lutz, Y., Cova, L., Hindelang, C., Jiralerspong, S., Trottier, Y., Kish, S.J., Faucheux, B., Trouillas, P., Authier, F.J., Durr, A., Mandel, J.L., Vescovi, A., Pandolfo, M. and Koenig, M.: Frataxin is reduced in Friedreich ataxia patients and is associated with mitochondrial membranes. Hum Mol Genet 6 (1997) 1771-80.

Campuzano, V., Montermini, L., Molto, M.D., Pianese, L., Cossee, M., Cavalcanti, F., Monros, E., Rodius, F., Duclos, F., Monticelli, A., Zara, F., Canizares, J., Koutnikova, H., Bidichandani, S.I., Gellera, C., Brice, A., Trouillas, P., De Michele, G., Filla, A., De Frutos, R., Palau, F., Patel, P.I., Di Donato, S., Mandel, J.L., Cocozza, S., Koenig, M. and Pandolfo, M.: Friedreich's ataxia: autosomal recessive disease caused by an intronic GAA triplet repeat expansion. Science 271 (1996) 1423-7.

Carpenter, S., Karpati, G., Andermann, F., Jacob, J.C. and Andermann, E.: Lafora's disease: peroxisomal storage in skeletal muscle. Neurology 24 (1974) 531-8.

C. elegans Sequencing Consortium: Genome sequence of the nematode C. elegans: a platform for investigating biology. Science 282 (1998) 2012-8.

Ceru, S., Rabzelj, S., Kopitar-Jerala, N., Turk, V. and Zerovnik, E.: Protein aggregation as a possible cause for pathology in a subset of familial Unverricht-Lundborg disease. Med Hypotheses 64 (2005) 955-9.

Chan, E.M., Ackerley, C.A., Lohi, H., Ianzano, L., Cortez, M.A., Shannon, P., Scherer, S.W. and Minassian, B.A.: Laforin preferentially binds the neurotoxic starch-like polyglucosans, which form in its absence in progressive myoclonus epilepsy. Hum Mol Genet 13 (2004a) 1117-29.

Chan, E.M., Andrade, D.M., Franceschetti, S. and Minassian, B.: Progressive myoclonus epilepsies: EPM1, EPM2A, EPM2B. Adv Neurol 95 (2005) 47-57.

Chan, E.M., Omer, S., Ahmed, M., Bridges, L.R., Bennett, C., Scherer, S.W. and Minassian, B.A.: Progressive myoclonus epilepsy with polyglucosans (Lafora disease): evidence for a third locus. Neurology 63 (2004b) 565-7.

Chan, E.M., Young, E.J., Ianzano, L., Munteanu, I., Zhao, X., Christopoulos, C.C., Avanzini, G., Elia, M., Ackerley, C.A., Jovic, N.J., Bohlega, S., Andermann, E., Rouleau, G.A., Delgado-Escueta, A.V., Minassian, B.A. and Scherer, S.W.: Mutations in NHLRC1 cause progressive myoclonus epilepsy. Nat Genet 35 (2003) 125-7.

Chang, C.S., Kokontis, J. and Liao, S.T.: Structural analysis of complementary DNA and amino acid sequences of human and rat androgen receptors. Proc Natl Acad Sci U S A 85 (1988) 7211-5.

Chimpanzee Sequencing and Analysis Consortium: Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 437 (2005) 69-87.

Consensus Group Marseille: Classification of progressive myoclonus epilepsies and related disorders. Ann Neurol 28 (1990) 113-6.

Corbin, F., Bouillon, M., Fortin, A., Morin, S., Rousseau, F. and Khandjian, E.W.: The fragile X mental retardation protein is associated with poly(A)+ mRNA in actively translating polyribosomes. Hum Mol Genet 6 (1997) 1465-72.

Cotman, S.L., Vrbanac, V., Lebel, L.A., Lee, R.L., Johnson, K.A., Donahue, L.R., Teed, A.M., Antonellis, K., Bronson, R.T., Lerner, T.J. and MacDonald, M.E.: Cln3(Deltaex7/8) knock-in mice with the common JNCL mutation exhibit progressive neurologic disease that begins before birth. Hum Mol Genet 11 (2002) 2709-21.

D'Amato, E., Kokaia, Z., Nanobashvili, A., Reeben, M., Lehesjoki, A.E., Saarma, M. and Lindvall, O.: Seizures induce widespread upregulation of cystatin B, the gene mutated in progressive myoclonus epilepsy, in rat forebrain neurons. Eur J Neurosci 12 (2000) 1687-95.

de Geest, N., Bonten, E., Mann, L., de Sousa-Hitzler, J., Hahn, C. and d'Azzo, A.: Systemic and neurologic abnormalities distinguish the lysosomal disorders sialidosis and galactosialidosis in mice. Hum Mol Genet 11 (2002) 1455-64.

de Haan, G.J., Halley, D.J., Doelman, J.C., Geesink, H.H., Augustijn, P.B., Jager-Jongkind, A.D., Majoie, M., Bader, A.J., Leliefeld-Ten Doeschate, L.A., Deelen, W.H., Bertram, E., Lehesjoki, A.E. and Lindhout, D.: Univerricht-Lundborg disease: underdiagnosed in the Netherlands. Epilepsia 45 (2004) 1061-3.

Deng, A., Irizarry, M.C., Nitsch, R.M., Growdon, J.H. and Rebeck, G.W.: Elevation of cystatin C in susceptible neurons in Alzheimer's disease. Am J Pathol 159 (2001) 1061-8.

Deussing, J., Roth, W., Saftig, P., Peters, C., Ploegh, H.L. and Villadangos, J.A.: Cathepsins B and D are dispensable for major histocompatibility complex class II-mediated antigen presentation. Proc Natl Acad Sci U S A 95 (1998) 4516-21.

Di Giaimo, R., Riccio, M., Santi, S., Galeotti, C., Ambrosetti, D.C. and Melli, M.: New insights into the molecular basis of progressive myoclonus epilepsy: a multiprotein complex with cystatin B. Hum Mol Genet 11 (2002) 2941-50.

Dunger, D.B., Ong, K.K., Huxtable, S.J., Sherriff, A., Woods, K.A., Ahmed, M.L., Golding, J., Pembrey, M.E., Ring, S., Bennett, S.T. and Todd, J.A.:

Association of the INS VNTR with size at birth. ALSPAC Study Team. Avon Longitudinal Study of Pregnancy and Childhood. Nat Genet 19 (1998) 98-100.

Dynan, W.S. and Tjian, R.: Isolation of transcription factors that discriminate between different promoters recognized by RNA polymerase II. Cell 32 (1983a) 669-80.

Dynan, W.S. and Tjian, R.: The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. Cell 35 (1983b) 79-87.

Eldridge, R., Iivanainen, M., Stern, R., Koerber, T. and Wilder, B.J.: "Baltic" myoclonus epilepsy: hereditary disorder of childhood made worse by phenytoin. Lancet 2 (1983) 838-42.

Engel, J., Jr.: A proposed diagnostic scheme for people with epileptic seizures and with epilepsy: report of the ILAE Task Force on Classification and Terminology. Epilepsia 42 (2001) 796-803.

Evans, R.M.: The steroid and thyroid hormone receptor superfamily. Science 240 (1988) 889-95.

Felbor, U., Kessler, B., Mothes, W., Goebel, H.H., Ploegh, H.L., Bronson, R.T. and Olsen, B.R.: Neuronal loss and brain atrophy in mice lacking cathepsins B and L. Proc Natl Acad Sci U S A 99 (2002) 7883-8.

Feng, Y., Zhang, F., Lokey, L.K., Chastain, J.L., Lakkis, L., Eberhart, D. and Warren, S.T.: Translational suppression by trinucleotide repeat expansion at FMR1. Science 268 (1995) 731-4.

Flanagan, P.M., Kelleher, R.J., 3rd, Sayre, M.H., Tschochner, H. and Kornberg, R.D.: A mediator required for activation of RNA polymerase II transcription in vitro. Nature 350 (1991) 436-8.

Flink, I.L. and Morkin, E.: Alternatively processed isoforms of cellular nucleic acid-binding protein interact with a suppressor region of the human beta-myosin heavy chain gene. J Biol Chem 270 (1995) 6959-65.

Fossum, K. and Whitaker, J.R.: Ficin and papain inhibitor from chicken egg white. Arch Biochem Biophys 125 (1968) 367-75.

Fukuhara, N., Tokiguchi, S., Shirakawa, K. and Tsubaki, T.: Myoclonus epilepsy associated with ragged-red fibres (mitochondrial abnormalities): disease entity or a syndrome? Light-and electron-microscopic studies of two cases and review of literature. J Neurol Sci 47 (1980) 117-33.

Ganesh, S., Agarwala, K.L., Ueda, K., Akagi, T., Shoda, K., Usui, T., Hashikawa, T., Osada, H., Delgado-Escueta, A.V. and Yamakawa, K.: Laforin, defective in the progressive myoclonus epilepsy of Lafora type, is a dual-specificity phosphatase associated with polyribosomes. Hum Mol Genet 9 (2000) 2251-61.

Ganesh, S., Delgado-Escueta, A.V., Sakamoto, T., Avila, M.R., Machado-Salas, J., Hoshii, Y., Akagi, T., Gomi, H., Suzuki, T., Amano, K., Agarwala, K.L., Hasegawa, Y., Bai, D.S., Ishihara, T., Hashikawa, T., Itohara, S., Cornford, E.M., Niki, H. and Yamakawa, K.: Targeted disruption of the Epm2a gene causes formation of Lafora inclusion bodies, neurodegeneration, ataxia, myoclonus epilepsy and impaired behavioral response in mice. Hum Mol Genet 11 (2002) 1251-62.

Gao, H., Boustany, R.M., Espinola, J.A., Cotman, S.L., Srinidhi, L., Antonellis, K.A., Gillis, T., Qin, X., Liu, S., Donahue, L.R., Bronson, R.T., Faust, J.R., Stout, D., Haines, J.L., Lerner, T.J. and MacDonald, M.E.: Mutations in a novel CLN6-encoded transmembrane protein cause variant neuronal ceroid lipofuscinosis in man and mouse. Am J Hum Genet 70 (2002) 324-35.

Gardiner, M., Sandford, A., Deadman, M., Poulton, J., Cookson, W., Reeders, S., Jokiaho, I., Peltonen, L., Eiberg, H. and Julier, C.: Batten disease (Spielmeyer-Vogt disease, juvenile onset neuronal ceroid-lipofuscinosis) gene (CLN3) maps to human chromosome 16. Genomics 8 (1990) 387-90.

Gecz, J., Bielby, S., Sutherland, G.R. and Mulley, J.C.: Gene structure and subcellular localization of FMR2, a member of a new family of putative transcription activators. Genomics 44 (1997) 201-13.

Gecz, J., Gedeon, A.K., Sutherland, G.R. and Mulley, J.C.: Identification of the gene FMR2, associated with FRAXE mental retardation. Nat Genet 13 (1996) 105-8.

Gelb, B.D., Shi, G.P., Chapman, H.A. and Desnick, R.J.: Pycnodysostosis, a lysosomal disease caused by cathepsin K deficiency. Science 273 (1996) 1236-8.

Ghiso, J., Jensson, O. and Frangione, B.: Amyloid fibrils in hereditary cerebral hemorrhage with amyloidosis of Icelandic type is a variant of gamma-trace basic protein (cystatin C). Proc Natl Acad Sci U S A 83 (1986) 2974-8.

Gibbs, R.A., Weinstock, G.M., Metzker, M.L., Muzny, D.M., Sodergren, E.J., Scherer, S., Scott, G., Steffen, D., Worley, K.C., Burch, P.E., Okwuonu, G., Hines, S., Lewis, L., DeRamo, C., Delgado, O., Dugan-Rocha, S., Miner, G., Morgan, M., Hawes, A., Gill, R., Celera, Holt, R.A., Adams, M.D., Amanatides, P.G., Baden-Tillson, H., Barnstead, M., Chin, S., Evans, C.A., Ferriera, S., Fosler, C., Glodek, A., Gu, Z., Jennings, D., Kraft, C.L., Nguyen, T., Pfannkoch, C.M., Sitter, C., Sutton, G.G., Venter, J.C., Woodage, T., Smith, D., Lee, H.M., Gustafson, E., Cahill, P., Kana, A., Doucette-Stamm, L., Weinstock, K., Fechtel, K., Weiss, R.B., Dunn, D.M., Green, E.D., Blakesley, R.W., Bouffard, G.G., De Jong, P.J., Osoegawa, K., Zhu, B., Marra, M., Schein, J., Bosdet, I., Fjell, C., Jones, S., Krzywinski, M., Mathewson, C., Siddiqui, A., Wye, N., McPherson, J., Zhao, S., Fraser, C.M., Shetty, J., Shatsman, S., Geer, K., Chen, Y., Abramzon, S., Nierman, W.C., Havlak, P.H., Chen, R., Durbin, K.J., Egan, A., Ren, Y., Song, X.Z., Li, B., Liu, Y., Qin, X., Cawley, S., Cooney, A.J., D'Souza, L.M., Martin, K., Wu, J.Q., Gonzalez-Garay, M.L., Jackson, A.R., Kalafus, K.J., McLeod, M.P., Milosavljevic, A., Virk, D., Volkov, A., Wheeler, D.A., Zhang, Z., Bailey, J.A., Eichler, E.E., Tuzun, E., et al.: Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature 428 (2004) 493-521.

Gitton, Y., Dahmane, N., Baik, S., Ruiz i Altaba, A., Neidhardt, L., Scholze, M., Herrmann, B.G., Kahlem, P., Benkahla, A., Schrinner, S., Yildirimman, R., Herwig, R., Lehrach, H. and Yaspo, M.L.: A gene expression map of human chromosome 21 orthologues in the mouse. Nature 420 (2002) 586-90.

Glass, C.K.: Differential recognition of target genes by nuclear receptor monomers, dimers, and heterodimers. Endocr Rev 15 (1994) 391-407.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H. and Oliver, S.G.: Life with 6000 genes. Science 274 (1996) 546, 563-7.

Grabowski, P.J. and Black, D.L.: Alternative RNA splicing in the nervous system. Prog Neurobiol 65 (2001) 289-308.

Green, E.L.: The genetics of a new hair deficiency, furless. J Hered 45 (1954) 115-118.

Green, G.D., Kembhavi, A.A., Davies, M.E. and Barrett, A.J.: Cystatin-like cysteine proteinase inhibitors from human liver. Biochem J 218 (1984) 939-46.

Gronostajski, R.M.: Roles of the NFI/CTF gene family in transcription and development. Gene 249 (2000) 31-45.

Grubb, A.O., Weiber, H. and Lofberg, H.: The gamma-trace concentration of normal human seminal plasma is thirty-six times that of normal human blood plasma. Scand J Clin Lab Invest 43 (1983) 421-5.

Gu, Y., Shen, Y., Gibbs, R.A. and Nelson, D.L.: Identification of FMR2, a novel gene associated with the FRAXE CCG repeat and CpG island. Nat Genet 13 (1996) 109-13.

Gupta, P., Soyombo, A.A., Atashband, A., Wisniewski, K.E., Shelton, J.M., Richardson, J.A., Hammer, R.E. and Hofmann, S.L.: Disruption of PPT1 or PPT2 causes neuronal ceroid lipofuscinosis in knockout mice. Proc Natl Acad Sci U S A 98 (2001) 13566-71.

Hagen, G., Muller, S., Beato, M. and Suske, G.: Cloning by recognition site screening of two novel GT box binding proteins: a family of Sp1 related genes. Nucleic Acids Res 20 (1992) 5519-25.

Hagen, G., Muller, S., Beato, M. and Suske, G.: Sp1-mediated transcriptional activation is repressed by Sp3. Embo J 13 (1994) 3843-51.

Hagerman, P.J. and Hagerman, R.J.: The fragile-X premutation: a maturing perspective. Am J Hum Genet 74 (2004) 805-16.

Halangk, W., Lerch, M.M., Brandt-Nedelev, B., Roth, W., Ruthenbuerger, M., Reinheckel, T., Domschke, W., Lippert, H., Peters, C. and Deussing, J.: Role of cathepsin B in intracellular trypsinogen activation and the onset of acute pancreatitis. J Clin Invest 106 (2000) 773-81.

Haltia, M.: The neuronal ceroid-lipofuscinoses. J Neuropathol Exp Neurol 62 (2003) 1-13.

Haltia, M., Kristensson, K. and Sourander, P.: Neuropathological studies in three Scandinavian cases of progressive myoclonus epilepsy. Acta Neurol Scand 45 (1969) 63-77.

Hashmi, S., Britton, C., Liu, J., Guiliano, D.B., Oksov, Y. and Lustigman, S.: Cathepsin L is essential for embryogenesis and development of Caenorhabditis elegans. J Biol Chem 277 (2002) 3477-86.

Huntington's Disease Collaborative Research Group: A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. Cell 72 (1993) 971-83.

Harriman, D.G., Millar, J.H. and Stevenson, A.C.: Progressive familial myoclonic epilepsy in three families: its clinical features and pathological basis. Brain 78 (1955) 325-49.

Holmes, S.E., Hearn, E.O., Ross, C.A. and Margolis, R.L.: SCA12: an unusual mutation leads to an unusual spinocerebellar ataxia. Brain Res Bull 56 (2001a) 397-403.

Holmes, S.E., O'Hearn, E. and Margolis, R.L.: Why is SCA12 different from other SCAs? Cytogenet Genome Res 100 (2003) 189-97.

Holmes, S.E., O'Hearn, E., Rosenblatt, A., Callahan, C., Hwang, H.S., Ingersoll-Ashworth, R.G., Fleisher, A., Stevanin, G., Brice, A., Potter, N.T., Ross, C.A. and Margolis, R.L.: A repeat expansion in the gene encoding junctophilin-3 is associated with Huntington disease-like 2. Nat Genet 29 (2001b) 377-8.

Holmes, S.E., O'Hearn, E.E., McInnis, M.G., Gorelick-Feldman, D.A., Kleiderlein, J.J., Callahan, C., Kwak, N.G., Ingersoll-Ashworth, R.G., Sherr, M., Sumner,

A.J., Sharp, A.H., Ananth, U., Seltzer, W.K., Boss, M.A., Vieria-Saecker, A.M., Epplen, J.T., Riess, O., Ross, C.A. and Margolis, R.L.: Expansion of a novel CAG trinucleotide repeat in the 5' region of PPP2R2B is associated with SCA12. Nat Genet 23 (1999) 391-2.

Horiuchi, H., Osawa, M., Furutani, R., Morita, M., Tian, W., Awatsu, Y., Shimazaki, H. and Umetsu, K.: Polymerase chain reaction-based analysis using deaminated DNA of dodecamer expansions in CSTB, associated with Unverricht-Lundborg myoclonus epilepsy. Genet Test 9 (2005) 328-33.

Houseweart, M.K., Pennacchio, L.A., Vilaythong, A., Peters, C., Noebels, J.L. and Myers, R.M.: Cathepsin B but not cathepsins L or S contributes to the pathogenesis of Unverricht-Lundborg progressive myoclonus epilepsy (EPM1). J Neurobiol 56 (2003a) 315-27.

Houseweart, M.K., Vilaythong, A., Yin, X.M., Turk, B., Noebels, J.L. and Myers, R.M.: Apoptosis caused by cathepsins does not require Bid signaling in an in vivo model of progressive myoclonus epilepsy (EPM1). Cell Death Differ 10 (2003b) 1329-35.

Huh, C., Nagle, J.W., Kozak, C.A., Abrahamson, M. and Karlsson, S.: Structural organization, expression and chromosomal mapping of the mouse cystatin-C-encoding gene (Cst3). Gene 152 (1995) 221-6.

Huh, C.G., Hakansson, K., Nathanson, C.M., Thorgeirsson, U.P., Jonsson, N., Grubb, A., Abrahamson, M. and Karlsson, S.: Decreased metastatic spread in mice homozygous for a null allele of the cystatin C protease inhibitor gene. Mol Pathol 52 (1999) 332-40.

Ianzano, L., Zhang, J., Chan, E.M., Zhao, X.C., Lohi, H., Scherer, S.W. and Minassian, B.A.: Lafora progressive Myoclonus Epilepsy mutation database-EPM2A and NHLRC1 (EMP2B) genes. Hum Mutat 26 (2005) 397.

Iivanainen, M. and Himberg, J.J.: Valproate and clonazepam in the treatment of severe progressive myoclonus epilepsy. Arch Neurol 39 (1982) 236-8.

Ikeuchi, T., Igarashi, S., Takiyama, Y., Onodera, O., Oyake, M., Takano, H., Koide, R., Tanaka, H. and Tsuji, S.: Non-Mendelian transmission in dentatorubral-pallidoluysian atrophy and Machado-Joseph disease: the mutant allele is preferentially transmitted in male meiosis. Am J Hum Genet 58 (1996) 730-3.

International Batten Disease Consortium: Isolation of a novel gene underlying Batten disease, CLN3. Cell 82 (1995) 949-57.

International Human Genome Sequencing Consortium: Finishing the euchromatic sequence of the human genome. Nature 431 (2004) 931-45.

Jalanko, A., Vesa, J., Manninen, T., von Schantz, C., Minye, H., Fabritius, A.L., Salonen, T., Rapola, J., Gentile, M., Kopra, O. and Peltonen, L.: Mice with Ppt1Deltaex4 mutation replicate the INCL phenotype and show an inflammation-associated loss of interneurons. Neurobiol Dis 18 (2005) 226-41.

Järvelä, I., Schleutker, J., Haataja, L., Santavuori, P., Puhakka, L., Manninen, T., Palotie, A., Sandkuijl, L.A., Renlund, M., White, R. and et al.: Infantile form of neuronal ceroid lipofuscinosis (CLN1) maps to the short arm of chromosome 1. Genomics 9 (1991) 170-3.

Järvinen, M. and Rinne, A.: Human spleen cysteineproteinase inhibitor. Purification, fractionation into isoelectric variants and some properties of the variants. Biochim Biophys Acta 708 (1982) 210-7.

Jerala, R., Trstenjak-Prebanda, M., Kroon-Zitko, L., Lenarcic, B. and Turk, V.: Mutations in the QVVAG region of the cysteine proteinase inhibitor stefin B. Biol Chem Hoppe Seyler 371 Suppl (1990) 157-60.

Jin, P. and Warren, S.T.: Understanding the molecular basis of fragile X syndrome. Hum Mol Genet 9 (2000) 901-8.

Kagitani-Shimono, K., Imai, K., Okamoto, N., Ono, J. and Okada, S.: Unverricht-Lundborg disease with cystatin B gene abnormalities. Pediatr Neurol 26 (2002) 55-60.

Karin, M., Liu, Z. and Zandi, E.: AP-1 function and regulation. Curr Opin Cell Biol 9 (1997) 240-6.

Kasper, D., Planells-Cases, R., Fuhrmann, J.C., Scheel, O., Zeitz, O., Ruether, K., Schmitt, A., Poet, M., Steinfeld, R., Schweizer, M., Kornak, U. and Jentsch, T.J.: Loss of the chloride channel ClC-7 leads to lysosomal storage disease and neurodegeneration. Embo J 24 (2005) 1079-91.

Katz, M.L., Khan, S., Awano, T., Shahid, S.A., Siakotos, A.N. and Johnson, G.S.: A mutation in the CLN8 gene in English Setter dogs with neuronal ceroid-lipofuscinosis. Biochem Biophys Res Commun 327 (2005) 541-7.

Katz, M.L., Shibuya, H., Liu, P.C., Kaur, S., Gao, C.L. and Johnson, G.S.: A mouse gene knockout model for juvenile ceroid-lipofuscinosis (Batten disease). J Neurosci Res 57 (1999) 551-6.

Kirschke, H. and Wiederanders, B.: Cathepsin S and related lysosomal endopeptidases. Methods Enzymol 244 (1994) 500-11.

Kirschke, H., Wiederanders, B., Bromme, D. and Rinne, A.: Cathepsin S from bovine spleen. Purification, distribution, intracellular localization and action on proteins. Biochem J 264 (1989) 467-73.

Kiviranta, R., Morko, J., Alatalo, S.L., NicAmhlaoibh, R., Risteli, J., Laitala-Leinonen, T. and Vuorio, E.: Impaired bone resorption in cathepsin K-deficient mice is partially compensated for by enhanced osteoclastogenesis and increased expression of other proteases via an increased RANKL/OPG ratio. Bone 36 (2005) 159-72.

Klockars, T., Savukoski, M., Isosomppi, J., Laan, M., Järvelä, I., Petrukhin, K., Palotie, A. and Peltonen, L.: Efficient construction of a physical map by fiber-FISH of the CLN5 region: refined assignment and long-range contig covering the critical region on 13q22. Genomics 35 (1996) 71-8.

Knight, S.J., Flannery, A.V., Hirst, M.C., Campbell, L., Christodoulou, Z., Phelps, S.R., Pointon, J., Middleton-Price, H.R., Barnicoat, A., Pembrey, M.E. and et al.: Trinucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. Cell 74 (1993) 127-34.

Koide, R., Ikeuchi, T., Onodera, O., Tanaka, H., Igarashi, S., Endo, K., Takahashi, H., Kondo, R., Ishikawa, A., Hayashi, T. and et al.: Unstable expansion of CAG repeat in hereditary dentatorubral-pallidoluysian atrophy (DRPLA). Nat Genet 6 (1994) 9-13.

Koob, M.D., Moseley, M.L., Schut, L.J., Benzow, K.A., Bird, T.D., Day, J.W. and Ranum, L.P.: An untranslated CTG expansion causes a novel form of spinocerebellar ataxia (SCA8). Nat Genet 21 (1999) 379-84.

Kopra, O., Vesa, J., von Schantz, C., Manninen, T., Minye, H., Fabritius, A.L., Rapola, J., van Diggelen, O.P., Saarela, J., Jalanko, A. and Peltonen, L.: A mouse model for Finnish variant late infantile neuronal ceroid lipofuscinosis, CLN5, reveals neuropathology associated with early aging. Hum Mol Genet 13 (2004) 2893-906.

Koskiniemi, M.: Psychological findings in progressive myoclonus epilepsy without Lafora bodies. Epilepsia 15 (1974) 537-45.

Koskiniemi, M.: Baltic myoclonus. Ann Clin Res 19 (1987) 311-2.

Koskiniemi, M., Donner, M., Majuri, H., Haltia, M. and Norio, R.: Progressive myoclonus epilepsy. A clinical and histopathological study. Acta Neurol Scand 50 (1974a) 307-32.

Koskiniemi, M., Toivakka, E. and Donner, M.: Progressive myoclonus epilepsy. Electroencephalographical findings. Acta Neurol Scand 50 (1974b) 333-59.

Koskiniemi, M., Van Vleymen, B., Häkämies, L., Lamusuo, S. and Taalas, J.: Piracetam relieves symptoms in progressive myoclonus epilepsy: a multicentre, randomised, double blind, crossover study comparing the efficacy and safety of three dosages of oral piracetam with placebo. J Neurol Neurosurg Psychiatry 64 (1998) 344-8.

Koskiniemi, M.L.: Baltic myoclonus. Adv Neurol 43 (1986) 57-64.

Kremer, E.J., Pritchard, M., Lynch, M., Yu, S., Holman, K., Baker, E., Warren, S.T., Schlessinger, D., Sutherland, G.R. and Richards, R.I.: Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. Science 252 (1991) 1711-4.

La Spada, A.R., Wilson, E.M., Lubahn, D.B., Harding, A.E. and Fischbeck, K.H.: Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. Nature 352 (1991) 77-9.

Lafora, G. and Gluck, B.: Beitrag zur histopathologie der myoklonischen epilepsie. Z Ges Neurol Phychiat 6 (1911) 1-14.

Lafrenière, R.G., Rochefort, D.L., Chrétien, N., Rommens, J.M., Cochius, J.I., Kälviainen, R., Nousiainen, U., Patry, G., Farrell, K., Söderfeldt, B., Federico, A., Hale, B.R., Cossio, O.H., Sørensen, T., Pouliot, M.A., Kmiec, T., Uldall, P., Janszky, J., Pranzatelli, M.R., Andermann, F., Andermann, E. and Rouleau, G.A.: Unstable insertion in the 5' flanking region of the cystatin B gene is the most common mutation in progressive myoclonus epilepsy type 1, EPM1. Nat Genet 15 (1997) 298-302.

Laitala-Leinonen, T., Rinne, R., Saukko, P., Väänänen, H.K. and Rinne, A.: Cystatin B as an intracellular modulator of bone resorption. Matrix Biol (2005).

Laitala-Leinonen, T., Rinne, R., Saukko, P., Väänänen, H.K. and Rinne, A.: Cystatin B as an intracellular modulator of bone resorption. Matrix Biol 25 (2006) 149-57.

Lalioti, M.D., Mirotsou, M., Buresi, C., Peitsch, M.C., Rossier, C., Ouazzani, R., Baldy-Moulinier, M., Bottani, A., Malafosse, A. and Antonarakis, S.E.: Identification of mutations in cystatin B, the gene responsible for the Unverricht-Lundborg type of progressive myoclonus epilepsy (EPM1). Am J Hum Genet 60 (1997a) 342-51.

Lalioti, M.D., Scott, H.S. and Antonarakis, S.E.: Altered spacing of promoter elements due to the dodecamer repeat expansion contributes to reduced expression of the cystatin B gene in EPM1. Hum Mol Genet 8 (1999) 1791-8.

Lalioti, M.D., Scott, H.S., Buresi, C., Rossier, C., Bottani, A., Morris, M.A., Malafosse, A. and Antonarakis, S.E.: Dodecamer repeat expansion in cystatin B gene in progressive myoclonus epilepsy. Nature 386 (1997b) 847-51.

Lalioti, M.D., Scott, H.S., Genton, P., Grid, D., Ouazzani, R., M'Rabet, A., Ibrahim, S., Gouider, R., Dravet, C., Chkili, T., Bottani, A., Buresi, C., Malafosse, A. and Antonarakis, S.E.: A PCR amplification method reveals instability of the dodecamer repeat in progressive myoclonus epilepsy (EPM1) and no

correlation between the size of the repeat and age at onset. Am J Hum Genet 62 (1998) 842-7.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., et al.: Initial sequencing and analysis of the human genome. Nature 409 (2001a) 860-921.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J.P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J.C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R.H., Wilson, R.K., Hillier, L.W., McPherson, J.D., Marra, M.A., Mardis, E.R., Fulton, L.A., Chinwalla, A.T., Pepin, K.H., Gish, W.R., Chissoe, S.L., Wendl, M.C., Delehaunty, K.D., Miner, T.L., Delehaunty, A., Kramer, J.B., Cook, L.L., Fulton, R.S., Johnson, D.L., Minx, P.J., Clifton, S.W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J.F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., et al.: Initial sequencing and analysis of the human genome. Nature 409 (2001b) 860-921.

Larson, G.P., Ding, S., Lafrenière, R.G., Rouleau, G.A. and Krontiris, T.G.: Instability of the EPM1 minisatellite. Hum Mol Genet 8 (1999) 1985-8.

Lassar, A.B., Paterson, B.M. and Weintraub, H.: Transfection of a DNA locus that mediates the conversion of 10T1/2 fibroblasts to myoblasts. Cell 47 (1986) 649-56.

Lee, J.E.: NeuroD and neurogenesis. Dev Neurosci 19 (1997) 27-32.

Lee, S.T. and Kim, M.: Aging and neurodegeneration. Molecular mechanisms of neuronal loss in Huntington's disease. Mech Ageing Dev 127 (2006) 432-5.

Lefebvre, C., Cocquerelle, C., Vandenbulcke, F., Hot, D., Huot, L., Lemoine, Y. and Salzet, M.: Transcriptomic analysis in the leech Theromyzon tessulatum: involvement of cystatin B in innate immunity. Biochem J 380 (2004) 617-25.

Lehesjoki, A.E.: Clinical features and genetics of Unverricht-Lundborg disease. Adv Neurol 89 (2002) 193-7.

Lehesjoki, A.E.: Molecular background of progressive myoclonus epilepsy. Embo J 22 (2003) 3473-8.

Lehesjoki, A.E. and Koskiniemi, M.: Clinical features and genetics of progressive myoclonus epilepsy of the Univerricht-Lundborg type. Ann Med 30 (1998) 474-80.

Lehesjoki, A.E. and Koskiniemi, M.: Progressive myoclonus epilepsy of Unverricht-Lundborg type. Epilepsia 40 Suppl 3 (1999) 23-8.

Lehesjoki, A.E., Koskiniemi, M., Norio, R., Tirrito, S., Sistonen, P., Lander, E. and de la Chapelle, A.: Localization of the EPM1 gene for progressive myoclonus epilepsy on chromosome 21: linkage disequilibrium allows high resolution mapping. Hum Mol Genet 2 (1993) 1229-34.

Lehesjoki, A.E., Koskiniemi, M., Pandolfo, M., Antonelli, A., Kyllerman, M., Wahlstrom, J., Nergardh, A., Burmeister, M., Sistonen, P., Norio, R. and et al.: Linkage studies in progressive myoclonus epilepsy: Unverricht-Lundborg and Lafora's diseases. Neurology 42 (1992) 1545-50.

Lehesjoki, A.E., Koskiniemi, M., Sistonen, P., Miao, J., Hastbacka, J., Norio, R. and de la Chapelle, A.: Localization of a gene for progressive myoclonus epilepsy to chromosome 21q22. Proc Natl Acad Sci U S A 88 (1991) 3696-9.

Lehesjoki, A.E., Tassinari, C.A., Avanzini, G., Michelucci, R., Franceschetti, S., Antonelli, A., Rubboli, G. and de la Chapelle, A.: PME of Unverricht-Lundborg type in the Mediterranean region: linkage and linkage disequilibrium confirm the assignment to the EPM1 locus. Hum Genet 93 (1994) 668-74.

Lenarcic, B., Krizaj, I., Zunec, P. and Turk, V.: Differences in specificity for the interactions of stefins A, B and D with cysteine proteinases. FEBS Lett 395 (1996) 113-8.

Leonardi, A., Turk, B. and Turk, V.: Inhibition of bovine cathepsins L and S by stefins and cystatins. Biol Chem Hoppe Seyler 377 (1996) 319-21.

Levine, M. and Tjian, R.: Transcription regulation and animal diversity. Nature 424 (2003) 147-51.

Levy, E., Sastre, M., Kumar, A., Gallo, G., Piccardo, P., Ghetti, B. and Tagliavini, F.: Codeposition of cystatin C with amyloid-beta protein in the brain of Alzheimer disease patients. J Neuropathol Exp Neurol 60 (2001) 94-104.

Lieberman, B.A., Bona, B.J., Edwards, D.P. and Nordeen, S.K.: The constitution of a progesterone response element. Mol Endocrinol 7 (1993) 515-27.

Lieuallen, K., Pennacchio, L.A., Park, M., Myers, R.M. and Lennon, G.G.: Cystatin B-deficient mice have increased expression of apoptosis and glial activation genes. Hum Mol Genet 10 (2001) 1867-71.

Lin, X. and Ashizawa, T.: Recent progress in spinocerebellar ataxia type-10 (SCA10). Cerebellum 4 (2005) 37-42.

Lindblad-Toh, K., Wade, C.M., Mikkelsen, T.S., Karlsson, E.K., Jaffe, D.B., Kamal, M., Clamp, M., Chang, J.L., Kulbokas, E.J., 3rd, Zody, M.C., Mauceli, E., Xie, X., Breen, M., Wayne, R.K., Ostrander, E.A., Ponting, C.P., Galibert, F., Smith, D.R., DeJong, P.J., Kirkness, E., Alvarez, P., Biagi, T., Brockman, W., Butler, J., Chin, C.W., Cook, A., Cuff, J., Daly, M.J., DeCaprio, D., Gnerre,

S., Grabherr, M., Kellis, M., Kleber, M., Bardeleben, C., Goodstadt, L., Heger, A., Hitte, C., Kim, L., Koepfli, K.P., Parker, H.G., Pollinger, J.P., Searle, S.M., Sutter, N.B., Thomas, R., Webber, C., Baldwin, J., Abebe, A., Abouelleil, A., Aftuck, L., Ait-Zahra, M., Aldredge, T., Allen, N., An, P., Anderson, S., Antoine, C., Arachchi, H., Aslam, A., Ayotte, L., Bachantsang, P., Barry, A., Bayul, T., Benamara, M., Berlin, A., Bessette, D., Blitshteyn, B., Bloom, T., Blye, J., Boguslavskiy, L., Bonnet, C., Boukhgalter, B., Brown, A., Cahill, P., Calixte, N., Camarata, J., Cheshatsang, Y., Chu, J., Citroen, M., Collymore, A., Cooke, P., Dawoe, T., Daza, R., Decktor, K., DeGray, S., Dhargay, N., Dooley, K., Dooley, K., Dorje, P., Dorjee, K., Dorris, L., Duffey, N., Dupes, A., Egbiremolen, O., Elong, R., Falk, J., Farina, A., Faro, S., Ferguson, D., Ferreira, P., Fisher, S., FitzGerald, M., et al.: Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438 (2005) 803-19.

Linnevers, C., Smeekens, S.P. and Bromme, D.: Human cathepsin W, a putative cysteine protease predominantly expressed in CD8+ T-lymphocytes. FEBS Lett 405 (1997) 253-9.

Liquori, C.L., Ricker, K., Moseley, M.L., Jacobsen, J.F., Kress, W., Naylor, S.L., Day, J.W. and Ranum, L.P.: Myotonic dystrophy type 2 caused by a CCTG expansion in intron 1 of ZNF9. Science 293 (2001) 864-7.

Lohi, H., Young, E.J., Fitzmaurice, S.N., Rusbridge, C., Chan, E.M., Vervoort, M., Turnbull, J., Zhao, X.C., Ianzano, L., Paterson, A.D., Sutter, N.B., Ostrander, E.A., Andre, C., Shelton, G.D., Ackerley, C.A., Scherer, S.W. and Minassian, B.A.: Expanded repeat in canine epilepsy. Science 307 (2005) 81.

Lombes, M., Binart, N., Oblin, M.E., Joulin, V. and Baulieu, E.E.: Characterization of the interaction of the human mineralocorticosteroid receptor with hormone response elements. Biochem J 292 (Pt 2) (1993) 577-83.

Lubahn, D.B., Joseph, D.R., Sullivan, P.M., Willard, H.F., French, F.S. and Wilson, E.M.: Cloning of human androgen receptor complementary DNA and localization to the X chromosome. Science 240 (1988) 327-30.

Lundborg, H.: Die progressive Myoklonusepilepsie (Unverricht's Myoklonie). Uppsala: Almqvist and Wiksell (pub.) 8 (1903) 567-70.

Lundborg, H.: Der Erbgang der progressiven Myoklonusepilepsie. (Myoklonie-Epilepsie, Unverricht's familiaere Myoklonie). Z Ges Neurol Psychiat 9 (1912) 353-8.

Lundborg, H.: Medizinisch-biologische Familienforschungen innerhalb eines 2232 koepfigen Bauerngeschlechtes in Schweden. Jena: Fischer (pub.) (1913).

Löfberg, H. and Grubb, A.O.: Quantitation of gamma-trace in human biological fluids: indications for production in the central nervous system. Scand J Clin Lab Invest 39 (1979) 619-26.

Machleidt, W., Thiele, U., Assfalg-Machleidt, I., Forger, D. and Auerswald, E.A.: Molecular mechanism of inhibition of cysteine proteinases by their protein inhibitors: kinetic studies with natural and recombinant variants of cystatins and stefins. Biomed Biochim Acta 50 (1991) 613-20.

Malafosse, A., Lehesjoki, A.E., Genton, P., Labauge, P., Durand, G., Tassinari, C.A., Dravet, C., Michelucci, R. and de la Chapelle, A.: Identical genetic locus for Baltic and Mediterranean myoclonus. Lancet 339 (1992) 1080-1.

Mancuso, M., Filosto, M., Mootha, V.K., Rocchi, A., Pistolesi, S., Murri, L., DiMauro, S. and Siciliano, G.: A novel mitochondrial tRNAPhe mutation causes MERRF syndrome. Neurology 62 (2004) 2119-21.

Maroteaux, P. and Lamy, M.: La pycnodysostose. Presse Med 70 (1962) 999-1002.

Martin, J.R., Craven, C.J., Jerala, R., Kroon-Zitko, L., Zerovnik, E., Turk, V. and Waltho, J.P.: The three-dimensional solution structure of human stefin A. J Mol Biol 246 (1995) 331-43.

Matsuura, T., Yamagata, T., Burgess, D.L., Rasmussen, A., Grewal, R.P., Watase, K., Khajavi, M., McCall, A.E., Davis, C.F., Zu, L., Achari, M., Pulst, S.M., Alonso, E., Noebels, J.L., Nelson, D.L., Zoghbi, H.Y. and Ashizawa, T.: Large expansion of the ATTCT pentanucleotide repeat in spinocerebellar ataxia type 10. Nat Genet 26 (2000) 191-4.

Mazarib, A., Xiong, L., Neufeld, M.Y., Birnbaum, M., Korczyn, A.D., Pandolfo, M. and Berkovic, S.F.: Unverricht-Lundborg disease in a five-generation Arab family: instability of dodecamer repeats. Neurology 57 (2001) 1050-4.

Melone, M.A., Tessa, A., Petrini, S., Lus, G., Sampaolo, S., di Fede, G., Santorelli, F.M. and Cotrufo, R.: Revelation of a new mitochondrial DNA mutation (G12147A) in a MELAS/MERFF phenotype. Arch Neurol 61 (2004) 269-72.

Merz, G.S., Benedikz, E., Schwenk, V., Johansen, T.E., Vogel, L.K., Rushbrook, J.I. and Wisniewski, H.M.: Human cystatin C forms an inactive dimer during intracellular trafficking in transfected CHO cells. J Cell Physiol 173 (1997) 423-32.

Minassian, B.A., Andrade, D.M., Ianzano, L., Young, E.J., Chan, E., Ackerley, C.A. and Scherer, S.W.: Laforin is a cell membrane and endoplasmic reticulum-associated protein tyrosine phosphatase. Ann Neurol 49 (2001) 271-5.

Minassian, B.A., Ianzano, L., Meloche, M., Andermann, E., Rouleau, G.A., Delgado-Escueta, A.V. and Scherer, S.W.: Mutation spectrum and predicted function of laforin in Lafora's progressive myoclonus epilepsy. Neurology 55 (2000) 341-6.

Minassian, B.A., Lee, J.R., Herbrick, J.A., Huizenga, J., Soder, S., Mungall, A.J., Dunham, I., Gardner, R., Fong, C.Y., Carpenter, S., Jardim, L., Satishchandra, P., Andermann, E., Snead, O.C., 3rd, Lopes-Cendes, I., Tsui, L.C., Delgado-Escueta, A.V., Rouleau, G.A. and Scherer, S.W.: Mutations in a gene encoding a novel protein tyrosine phosphatase cause progressive myoclonus epilepsy. Nat Genet 20 (1998) 171-4.

Mitchison, H.M., Bernard, D.J., Greene, N.D., Cooper, J.D., Junaid, M.A., Pullarkat, R.K., de Vos, N., Breuning, M.H., Owens, J.W., Mobley, W.C., Gardiner, R.M., Lake, B.D., Taschner, P.E. and Nussbaum, R.L.: Targeted disruption of the Cln3 gene provides a mouse model for Batten disease. The Batten Mouse Model Consortium [corrected]. Neurobiol Dis 6 (1999) 321-34.

Molkentin, J.D. and Olson, E.N.: Defining the regulatory networks for muscle development. Curr Opin Genet Dev 6 (1996) 445-53.

Moseley, M.L., Zu, T., Ikeda, Y., Gao, W., Mosemiller, A.K., Daughters, R.S., Chen, G., Weatherspoon, M.R., Clark, H.B., Ebner, T.J., Day, J.W. and Ranum, L.P.: Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. Nat Genet 38 (2006) 758-69.

Moulard, B., Genton, P., Grid, D., Jeanpierre, M., Ouazzani, R., Mrabet, A., Morris, M., LeGuern, E., Dravet, C., Mauguiere, F., Utermann, B., Baldy-Moulinier, M., Belaidi, H., Bertran, F., Biraben, A., Ali Cherif, A., Chkili, T., Crespel, A., Darcel, F., Dulac, O., Geny, C., Humbert-Claude, V., Kassiotis, P., Buresi, C. and Malafosse, A.: Haplotype study of West European and North African

Unverricht-Lundborg chromosomes: evidence for a few founder mutations. Hum Genet 111 (2002) 255-62.

Muller, C.A., Autenrieth, I.B. and Peschel, A.: Innate defenses of the intestinal epithelial barrier. Cell Mol Life Sci 62 (2005) 1297-1307.

Mulley, J.C., Scheffer, I.E., Harkin, L.A., Berkovic, S.F. and Dibbens, L.M.: Susceptibility genes for complex epilepsy. Hum Mol Genet 14 Spec No. 2 (2005) R243-9.

Murre, C., McCaw, P.S., Vaessin, H., Caudy, M., Jan, L.Y., Jan, Y.N., Cabrera, C.V., Buskin, J.N., Hauschka, S.D., Lassar, A.B. and et al.: Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence. Cell 58 (1989) 537-44.

Myers, E.W., Sutton, G.G., Delcher, A.L., Dew, I.M., Fasulo, D.P., Flanigan, M.J., Kravitz, S.A., Mobarry, C.M., Reinert, K.H., Remington, K.A., Anson, E.L., Bolanos, R.A., Chou, H.H., Jordan, C.M., Halpern, A.L., Lonardi, S., Beasley, E.M., Brandon, R.C., Chen, L., Dunn, P.J., Lai, Z., Liang, Y., Nusskern, D.R., Zhan, M., Zhang, Q., Zheng, X., Rubin, G.M., Adams, M.D. and Venter, J.C.: A whole-genome assembly of Drosophila. Science 287 (2000) 2196-204.

Naito, H. and Oyanagi, S.: Familial myoclonus epilepsy and choreoathetosis: hereditary dentatorubral-pallidoluysian atrophy. Neurology 32 (1982) 798-807.

Nakagawa, T.Y., Brissette, W.H., Lira, P.D., Griffiths, R.J., Petrushova, N., Stock, J., McNeish, J.D., Eastman, S.E., Howard, E.D., Clarke, S.R., Rosloniec, E.F., Elliott, E.A. and Rudensky, A.Y.: Impaired invariant chain degradation and antigen presentation and diminished collagen-induced arthritis in cathepsin S null mice. Immunity 10 (1999) 207-17.

Nakamura, M., Nakano, S., Goto, Y., Ozawa, M., Nagahama, Y., Fukuyama, H., Akiguchi, I., Kaji, R. and Kimura, J.: A novel point mutation in the mitochondrial tRNA(Ser(UCN)) gene detected in a family with MERRF/MELAS overlap syndrome. Biochem Biophys Res Commun 214 (1995) 86-93.

Nemes, J.P., Benzow, K.A., Moseley, M.L., Ranum, L.P. and Koob, M.D.: The SCA8 transcript is an antisense RNA to a brain-specific transcript encoding a novel actin-binding protein (KLHL1). Hum Mol Genet 9 (2000) 1543-51.

Nemes, Z. and Steinert, P.M.: Bricks and mortar of the epidermal barrier. Exp Mol Med 31 (1999) 5-19.

Nordeen, S.K., Suh, B.J., Kuhnel, B. and Hutchison, C.D.: Structural determinants of a glucocorticoid receptor recognition element. Mol Endocrinol 4 (1990) 1866-73.

Norio, R.: The Finnish Disease Heritage III: the individual diseases. Hum Genet 112 (2003) 470-526.

Norio, R. and Koskiniemi, M.: Progressive myoclonus epilepsy: genetic and nosological aspects with special reference to 107 Finnish patients. Clin Genet 15 (1979) 382-98.

Norio, R., Nevanlinna, H.R. and Perheentupa, J.: Hereditary diseases in Finland; rare flora in rare soul. Ann Clin Res 5 (1973) 109-41.

North, M.J., Robertson, C.D. and Coombs, G.H.: The specificity of trichomonad cysteine proteinases analysed using fluorogenic substrates and specific inhibitors. Mol Biochem Parasitol 39 (1990) 183-93.

O'Brien, J.S.: Neuraminidase deficiency in the cherry red spot-myoclonus syndrome. Biochem Biophys Res Commun 79 (1977) 1136-41.

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., et al.: Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. Nature 420 (2002) 563-73.

Olafsson, I.: The human cystatin C gene promoter: functional analysis and identification of heterogeneous mRNA. Scand J Clin Lab Invest 55 (1995) 597-607.

Olsson, T., Nygren, J., Hakansson, K., Lundblad, C., Grubb, A., Smith, M.L. and Wieloch, T.: Gene deletion of cystatin C aggravates brain damage following focal ischemia but mitigates the neuronal injury after global ischemia in the mouse. Neuroscience 128 (2004) 65-71.

Oostra, B.A. and Willemsen, R.: A fragile balance: FMR1 expression levels. Hum Mol Genet 12 Spec No 2 (2003) R249-57.

Orphanides, G., Lagrange, T. and Reinberg, D.: The general transcription factors of RNA polymerase II. Genes Dev 10 (1996) 2657-83.

Orr, H.T. and Zoghbi, H.Y.: SCA1 molecular genetics: a history of a 13 year collaboration against glutamines. Hum Mol Genet 10 (2001) 2307-11.

Owen, F., Poulter, M., Lofthouse, R., Collinge, J., Crow, T.J., Risby, D., Baker, H.F., Ridley, R.M., Hsiao, K. and Prusiner, S.B.: Insertion in prion protein gene in familial Creutzfeldt-Jakob disease. Lancet 1 (1989) 51-2.

Owen, F., Poulter, M., Shah, T., Collinge, J., Lofthouse, R., Baker, H., Ridley, R., McVey, J. and Crow, T.J.: An in-frame insertion in the prion protein gene in familial Creutzfeldt-Jakob disease. Brain Res Mol Brain Res 7 (1990) 273-6.

Owerbach, D. and Gabbay, K.H.: Localization of a type I diabetes susceptibility locus to the variable tandem repeat region flanking the insulin gene. Diabetes 42 (1993) 1708-14.

Palsdottir, A., Abrahamson, M., Thorsteinsson, L., Arnason, A., Olafsson, I., Grubb, A. and Jensson, O.: Mutation in cystatin C gene causes hereditary brain haemorrhage. Lancet 2 (1988) 603-4.

Pandolfo, M.: Molecular genetics and pathogenesis of Friedreich ataxia. Neuromuscul Disord 8 (1998) 409-15.

Parmeggiani, A., Lehesjoki, A.E., Carelli, V., Posar, A., Santi, A., Santucci, M., Gobbi, G., Pini, A. and Rossi, P.G.: Familial Unverricht-Lundborg disease: a clinical, neurophysiologic, and genetic study. Epilepsia 38 (1997) 637-41.

Parsons, M.A., Sinden, R.R. and Izban, M.G.: Transcriptional properties of RNA polymerase II within triplet repeat-containing DNA from the human myotonic dystrophy and fragile X loci. J Biol Chem 273 (1998) 26998-7008.

Pataskar, S.S., Dash, D. and Brahmachari, S.K.: Intramolecular i-motif structure at acidic pH for progressive myoclonus epilepsy (EPM1) repeat d(CCCCGCCCCGCG)n. J Biomol Struct Dyn 19 (2001a) 307-13.

Pataskar, S.S., Dash, D. and Brahmachari, S.K.: Progressive myoclonus epilepsy [EPM1] repeat d(CCCCGCCCCGCG)n forms folded hairpin structures at physiological pH. J Biomol Struct Dyn 19 (2001b) 293-305.

Pennacchio, L.A., Bouley, D.M., Higgins, K.M., Scott, M.P., Noebels, J.L. and Myers, R.M.: Progressive ataxia, myoclonic epilepsy and cerebellar apoptosis in cystatin B-deficient mice. Nat Genet 20 (1998) 251-8.

Pennacchio, L.A., Lehesjoki, A.E., Stone, N.E., Willour, V.L., Virtaneva, K., Miao, J., D'Amato, E., Ramirez, L., Faham, M., Koskiniemi, M., Warrington, J.A., Norio, R., de la Chapelle, A., Cox, D.R. and Myers, R.M.: Mutations in the gene encoding cystatin B in progressive myoclonus epilepsy (EPM1). Science 271 (1996) 1731-4.

Pennacchio, L.A. and Myers, R.M.: Isolation and characterization of the mouse cystatin B gene. Genome Res 6 (1996) 1103-9.

Pennisi, E.: Why do humans have so few genes? Science 309 (2005) 80.

Philipsen, S. and Suske, G.: A tale of three fingers: the family of mammalian Sp/XKLF transcription factors. Nucleic Acids Res 27 (1999) 2991-3000.

Poet, M., Kornak, U., Schweizer, M., Zdebik, A.A., Scheel, O., Hoelter, S., Wurst, W., Schmitt, A., Fuhrmann, J.C., Planells-Cases, R., Mole, S.E., Hubner, C.A. and Jentsch, T.J.: Lysosomal storage disease upon disruption of the neuronal chloride transport protein ClC-6. Proc Natl Acad Sci U S A 103 (2006) 13854-9.

Pol, E. and Björk, I.: Importance of the second binding loop and the C-terminal end of cystatin B (stefin B) for inhibition of cysteine proteinases. Biochemistry 38 (1999) 10519-26.

Pol, E., Olsson, S.L., Estrada, S., Prasthofer, T.W. and Björk, I.: Characterization by spectroscopic, kinetic and equilibrium methods of the interaction between recombinant human cystatin A (stefin A) and cysteine proteinases. Biochem J 311 (Pt 1) (1995) 275-82.

Potts, W., Bowyer, J., Jones, H., Tucker, D., Freemont, A.J., Millest, A., Martin, C., Vernon, W., Neerunjun, D., Slynn, G., Harper, F. and Maciewicz, R.: Cathepsin L-deficient mice exhibit abnormal skin and bone development and show increased resistance to osteoporosis following ovariectomy. Int J Exp Pathol 85 (2004) 85-96.

Rabzelj, S., Turk, V. and Zerovnik, E.: In vitro study of stability and amyloid-fibril formation of two mutants of human stefin B (cystatin B) occurring in patients with EPM1. Protein Sci (2005).

Ranta, S., Topcu, M., Tegelberg, S., Tan, H., Ustubutun, A., Saatci, I., Dufke, A., Enders, H., Pohl, K., Alembik, Y., Mitchell, W.A., Mole, S.E. and Lehesjoki, A.E.: Variant late infantile neuronal ceroid lipofuscinosis in a subset of Turkish patients is allelic to Northern epilepsy. Hum Mutat 23 (2004) 300-5.

Ranta, S., Zhang, Y., Ross, B., Lonka, L., Takkunen, E., Messer, A., Sharp, J., Wheeler, R., Kusumi, K., Mole, S., Liu, W., Soares, M.B., Bonaldo, M.F., Hirvasniemi, A., de la Chapelle, A., Gilliam, T.C. and Lehesjoki, A.E.: The

neuronal ceroid lipofuscinoses in human EPMR and mnd mutant mice are associated with mutations in CLN8. Nat Genet 23 (1999) 233-6.
Ranum, L.P. and Cooper, T.A.: RNA-Mediated Neuromuscular Disorders. Annu Rev Neurosci (2006).
Ranum, L.P. and Day, J.W.: Myotonic dystrophy: RNA pathogenesis comes into focus. Am J Hum Genet 74 (2004a) 793-804.
Ranum, L.P. and Day, J.W.: Pathogenic RNA repeats: an expanding role in genetic disease. Trends Genet 20 (2004b) 506-12.
Rawlings, N.D. and Barrett, A.J.: Evolution of proteins of the cystatin superfamily. J Mol Evol 30 (1990) 60-71.
Reymond, A., Marigo, V., Yaylaoglu, M.B., Leoni, A., Ucla, C., Scamuffa, N., Caccioppoli, C., Dermitzakis, E.T., Lyle, R., Banfi, S., Eichele, G., Antonarakis, S.E. and Ballabio, A.: Human chromosome 21 gene expression atlas in the mouse. Nature 420 (2002) 582-6.
Riccio, M., Di Giaimo, R., Pianetti, S., Palmieri, P.P., Melli, M. and Santi, S.: Nuclear localization of cystatin B, the cathepsin inhibitor implicated in myoclonus epilepsy (EPM1). Exp Cell Res 262 (2001) 84-94.
Riccio, M., Santi, S., Dembic, M., Di Giaimo, R., Cipollini, E., Costantino-Ceccarini, E., Ambrosetti, D., Maraldi, N.M. and Melli, M.: Cell-specific expression of the epm1 (cystatin B) gene in developing rat cerebellum. Neurobiol Dis 20 (2005) 104-14.
Riese, R.J. and Chapman, H.A.: Cathepsins and compartmentalization in antigen presentation. Curr Opin Immunol 12 (2000) 107-13.
Rinne, A., Järvinen, M., Martikainen, J., Alavaikko, M. and Räsänen, O.: Über das Vorkommen des epidermalen SH-Protease-Inhibitors im lymphatischen Gewebe. Verb Anat Ges 75 (1981) 573-4.
Rinne, R., Saukko, P., Järvinen, M. and Lehesjoki, A.E.: Reduced cystatin B activity correlates with enhanced cathepsin activity in progressive myoclonus epilepsy. Ann Med 34 (2002) 380-5.
Ritonja, A., Machleidt, W. and Barrett, A.J.: Amino acid sequence of the intracellular cysteine proteinase inhibitor cystatin B from human liver. Biochem Biophys Res Commun 131 (1985) 1187-92.
Roche, P.J., Hoare, S.A. and Parker, M.G.: A consensus DNA-binding site for the androgen receptor. Mol Endocrinol 6 (1992) 2229-35.
Rosing, H.S., Hopkins, L.C., Wallace, D.C., Epstein, C.M. and Weidenheim, K.: Maternally inherited mitochondrial myopathy and myoclonic epilepsy. Ann Neurol 17 (1985) 228-37.
Roth, W., Deussing, J., Botchkarev, V.A., Pauly-Evers, M., Saftig, P., Hafner, A., Schmidt, P., Schmahl, W., Scherer, J., Anton-Lamprecht, I., Von Figura, K., Paus, R. and Peters, C.: Cathepsin L deficiency as molecular defect of furless: hyperproliferation of keratinocytes and pertubation of hair follicle cycling. Faseb J 14 (2000) 2075-86.
Rotig, A., de Lonlay, P., Chretien, D., Foury, F., Koenig, M., Sidi, D., Munnich, A. and Rustin, P.: Aconitase and mitochondrial iron-sulphur protein deficiency in Friedreich ataxia. Nat Genet 17 (1997) 215-7.
Saftig, P., Hetman, M., Schmahl, W., Weber, K., Heine, L., Mossmann, H., Koster, A., Hess, B., Evers, M., von Figura, K. and et al.: Mice deficient for the lysosomal proteinase cathepsin D exhibit progressive atrophy of the intestinal mucosa and profound destruction of lymphoid cells. Embo J 14 (1995) 3599-608.

Saftig, P., Hunziker, E., Everts, V., Jones, S., Boyde, A., Wehmeyer, O., Suter, A. and von Figura, K.: Functions of cathepsin K in bone resorption. Lessons from cathepsin K deficient mice. Adv Exp Med Biol 477 (2000) 293-303.

Saftig, P., Hunziker, E., Wehmeyer, O., Jones, S., Boyde, A., Rommerskirch, W., Moritz, J.D., Schu, P. and von Figura, K.: Impaired osteoclastic bone resorption leads to osteopetrosis in cathepsin-K-deficient mice. Proc Natl Acad Sci U S A 95 (1998) 13453-8.

Saha, T. and Usdin, K.: Tetraplex formation by the progressive myoclonus epilepsy type-1 repeat: implications for instability in the repeat expansion diseases. FEBS Lett 491 (2001) 184-7.

Santamaria, I., Velasco, G., Cazorla, M., Fueyo, A., Campo, E. and Lopez-Otin, C.: Cathepsin L2, a novel human cysteine proteinase produced by breast and colorectal carcinomas. Cancer Res 58 (1998) 1624-30.

Santavuori, P.: Neuronal ceroid-lipofuscinoses in childhood. Brain Dev 10 (1988) 80-3.

Savukoski, M., Klockars, T., Holmberg, V., Santavuori, P., Lander, E.S. and Peltonen, L.: CLN5, a novel gene encoding a putative transmembrane protein mutated in Finnish variant late infantile neuronal ceroid lipofuscinosis. Nat Genet 19 (1998) 286-8.

Schulz, A., Dhar, S., Rylova, S., Dbaibo, G., Alroy, J., Hagel, C., Artacho, I., Kohlschutter, A., Lin, S. and Boustany, R.M.: Impaired cell adhesion and apoptosis in a novel CLN9 Batten disease variant. Ann Neurol 56 (2004) 342-50.

Schulz, A., Mousallem, T., Venkataramani, M., Persaud-Sawin, D.A., Zucker, A., Luberto, C., Bielawska, A., Bielawski, J., Holthuis, J.C., Jazwinski, S.M., Kozhaya, L., Dbaibo, G.S. and Boustany, R.M.: The CLN9 protein, a regulator of dihydroceramide synthase. J Biol Chem 281 (2006) 2784-94.

Schuurmans, C. and Guillemot, F.: Molecular mechanisms underlying cell fate specification in the developing telencephalon. Curr Opin Neurobiol 12 (2002) 26-34.

Serratosa, J.M., Gomez-Garre, P., Gallardo, M.E., Anta, B., de Bernabe, D.B., Lindhout, D., Augustijn, P.B., Tassinari, C.A., Malafosse, R.M., Topcu, M., Grid, D., Dravet, C., Berkovic, S.F. and de Cordoba, S.R.: A novel protein tyrosine phosphatase gene is mutated in progressive myoclonus epilepsy of the Lafora type (EPM2). Hum Mol Genet 8 (1999) 345-52.

Seyrantepe, V., Poupetova, H., Froissart, R., Zabot, M.T., Maire, I. and Pshezhetsky, A.V.: Molecular pathology of NEU1 gene in sialidosis. Hum Mutat 22 (2003) 343-52.

Shannon, P., Pennacchio, L.A., Houseweart, M.K., Minassian, B.A. and Myers, R.M.: Neuropathological changes in a mouse model of progressive myoclonus epilepsy: cystatin B deficiency and Unverricht-Lundborg disease. J Neuropathol Exp Neurol 61 (2002) 1085-91.

Sharp, J.D., Wheeler, R.B., Lake, B.D., Savukoski, M., Jarvela, I.E., Peltonen, L., Gardiner, R.M. and Williams, R.E.: Loci for classical and a variant late infantile neuronal ceroid lipofuscinosis map to chromosomes 11p15 and 15q21-23. Hum Mol Genet 6 (1997) 591-5.

Shi, G.P., Sukhova, G.K., Kuzuya, M., Ye, Q., Du, J., Zhang, Y., Pan, J.H., Lu, M.L., Cheng, X.W., Iguchi, A., Perrey, S., Lee, A.M., Chapman, H.A. and Libby, P.: Deficiency of the cysteine protease cathepsin S impairs microvessel growth. Circ Res 92 (2003) 493-500.

Shi, G.P., Villadangos, J.A., Dranoff, G., Small, C., Gu, L., Haley, K.J., Riese, R., Ploegh, H.L. and Chapman, H.A.: Cathepsin S required for normal MHC class II peptide loading and germinal center development. Immunity 10 (1999) 197-206.

Shoffner, J.M., Lott, M.T., Lezza, A.M., Seibel, P., Ballinger, S.W. and Wallace, D.C.: Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial DNA tRNA(Lys) mutation. Cell 61 (1990) 931-7.

Siintola, E., Partanen, S., Stromme, P., Haapanen, A., Haltia, M., Maehlen, J., Lehesjoki, A.E. and Tyynelä, J.: Cathepsin D deficiency underlies congenital human neuronal ceroid-lipofuscinosis. Brain 129 (2006) 1438-45.

Siintola, E., Topcu, M., Kohlschutter, A., Salonen, T., Joensuu, T., Anttonen, A.K. and Lehesjoki, A.E.: Two novel CLN6 mutations in variant late-infantile neuronal ceroid lipofuscinosis patients of Turkish origin. Clin Genet 68 (2005) 167-73.

Sleat, D.E., Donnelly, R.J., Lackland, H., Liu, C.G., Sohar, I., Pullarkat, R.K. and Lobel, P.: Association of mutations in a lysosomal protein with classical late-infantile neuronal ceroid lipofuscinosis. Science 277 (1997) 1802-5.

Sleat, D.E., Wiseman, J.A., El-Banna, M., Kim, K.H., Mao, Q., Price, S., Macauley, S.L., Sidman, R.L., Shen, M.M., Zhao, Q., Passini, M.A., Davidson, B.L., Stewart, G.R. and Lobel, P.: A mouse model of classical late-infantile neuronal ceroid lipofuscinosis based on targeted disruption of the CLN2 gene results in a loss of tripeptidyl-peptidase I activity and progressive neurodegeneration. J Neurosci 24 (2004) 9117-26.

Sloane, B.F.: Cathepsin B and cystatins: evidence for a role in cancer progression. Semin Cancer Biol 1 (1990) 137-52.

Smale, S.T. and Kadonaga, J.T.: The RNA polymerase II core promoter. Annu Rev Biochem 72 (2003) 449-79.

Smith, J.K., Gonda, V.E. and Malamud, N.: Unusual form of cerebellar ataxia; combined dentato-rubral and pallido-Luysian degeneration. Neurology 8 (1958) 205-9.

Sol-Church, K., Shipley, J., Beckman, D.A. and Mason, R.W.: Expression of cysteine proteases in extraembryonic tissues during mouse embryogenesis. Arch Biochem Biophys 372 (1999) 375-81.

Somerville, E.R. and Olanow, C.W.: Valproic acid. Treatment of myoclonus in dyssynergia cerebellaris myoclonica. Arch Neurol 39 (1982) 527-8.

Steinfeld, R., Reinhardt, K., Schreiber, K., Hillebrand, M., Kraetzner, R., Bruck, W., Saftig, P. and Gartner, J.: Cathepsin d deficiency is associated with a human neurodegenerative disorder. Am J Hum Genet 78 (2006) 988-98.

Steinlein, O.K.: Genetic mechanisms that underlie epilepsy. Nat Rev Neurosci 5 (2004) 400-8.

Stobrawa, S.M., Breiderhoff, T., Takamori, S., Engel, D., Schweizer, M., Zdebik, A.A., Bosl, M.R., Ruether, K., Jahn, H., Draguhn, A., Jahn, R. and Jentsch, T.J.: Disruption of ClC-3, a chloride channel expressed on synaptic vesicles, leads to a loss of the hippocampus. Neuron 29 (2001) 185-96.

Stoka, V., Nycander, M., Lenarcic, B., Labriola, C., Cazzulo, J.J., Bjork, I. and Turk, V.: Inhibition of cruzipain, the major cysteine proteinase of the protozoan parasite, Trypanosoma cruzi, by proteinase inhibitors of the cystatin superfamily. FEBS Lett 370 (1995) 101-4.

Stoka, V., Turk, B., Schendel, S.L., Kim, T.H., Cirman, T., Snipas, S.J., Ellerby, L.M., Bredesen, D., Freeze, H., Abrahamson, M., Bromme, D., Krajewski, S.,

Reed, J.C., Yin, X.M., Turk, V. and Salvesen, G.S.: Lysosomal protease pathways to apoptosis. Cleavage of bid, not pro-caspases, is the most likely route. J Biol Chem 276 (2001) 3149-57.

Stubbs, M.T., Laber, B., Bode, W., Huber, R., Jerala, R., Lenarcic, B. and Turk, V.: The refined 2.4 A X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. Embo J 9 (1990) 1939-47.

Stypmann, J., Glaser, K., Roth, W., Tobin, D.J., Petermann, I., Matthias, R., Monnig, G., Haverkamp, W., Breithardt, G., Schmahl, W., Peters, C. and Reinheckel, T.: Dilated cardiomyopathy in mice deficient for the lysosomal cysteine peptidase cathepsin L. Proc Natl Acad Sci U S A 99 (2002) 6234-9.

Suske, G.: The Sp-family of transcription factors. Gene 238 (1999) 291-300.

Svenson, I.K., Ashley-Koch, A.E., Gaskell, P.C., Riney, T.J., Cumming, W.J., Kingston, H.M., Hogan, E.L., Boustany, R.M., Vance, J.M., Nance, M.A., Pericak-Vance, M.A. and Marchuk, D.A.: Identification and expression analysis of spastin gene mutations in hereditary spastic paraplegia. Am J Hum Genet 68 (2001) 1077-85.

Takahashi, H., Asano, K., Kinouchi, M., Ishida-Yamamoto, A., Wuepper, K.D. and Iizuka, H.: Structure and transcriptional regulation of the human cystatin A gene. The 12-O-tetradecanoylphorbol-13-acetate (TPA) responsive element-2 site (-272 to -278) on cystatin A gene is critical for TPA-dependent regulation. J Biol Chem 273 (1998) 17375-80.

Takeshima, H., Komazaki, S., Nishi, M., Iino, M. and Kangawa, K.: Junctophilins: a novel family of junctional membrane complex proteins. Mol Cell 6 (2000) 11-22.

Tang, C.H., Lee, J.W., Galvez, M.G., Robillard, L., Mole, S.E. and Chapman, H.A.: Murine cathepsin F deficiency causes neuronal lipofuscinosis and late-onset neurological disease. Mol Cell Biol 26 (2006) 2309-16.

Tsuji, S., Choudary, P.V., Martin, B.M., Stubblefield, B.K., Mayor, J.A., Barranger, J.A. and Ginns, E.I.: A mutation in the human glucocerebrosidase gene in neuronopathic Gaucher's disease. N Engl J Med 316 (1987) 570-5.

Turk, B., Turk, D. and Turk, V.: Lysosomal cysteine proteases: more than scavengers. Biochim Biophys Acta 1477 (2000) 98-111.

Tyynelä, J., Sohar, I., Sleat, D.E., Gin, R.M., Donnelly, R.J., Baumann, M., Haltia, M. and Lobel, P.: A mutation in the ovine cathepsin D gene causes a congenital lysosomal storage disease with profound neurodegeneration. Embo J 19 (2000) 2786-92.

Unverricht, H.: Die Myoclonie. Berlin: Franz Deuticke (pub.) (1891).

Unverricht, H.: Ueber familiaere Myoclonie. Dtsch Z Nervenheilk (1895) 32-67.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith, H.O., Yandell, M., Evans, C.A., Holt, R.A., Gocayne, J.D., Amanatides, P., Ballew, R.M., Huson, D.H., Wortman, J.R., Zhang, Q., Kodira, C.D., Zheng, X.H., Chen, L., Skupski, M., Subramanian, G., Thomas, P.D., Zhang, J., Gabor Miklos, G.L., Nelson, C., Broder, S., Clark, A.G., Nadeau, J., McKusick, V.A., Zinder, N., Levine, A.J., Roberts, R.J., Simon, M., Slayman, C., Hunkapiller, M., Bolanos, R., Delcher, A., Dew, I., Fasulo, D., Flanigan, M., Florea, L., Halpern, A., Hannenhalli, S., Kravitz, S., Levy, S., Mobarry, C., Reinert, K., Remington, K., Abu-Threideh, J., Beasley, E., Biddick, K., Bonazzi, V., Brandon, R., Cargill, M., Chandramouliswaran, I., Charlab, R., Chaturvedi, K., Deng, Z., Di Francesco, V., Dunn, P., Eilbeck, K.,

Evangelista, C., Gabrielian, A.E., Gan, W., Ge, W., Gong, F., Gu, Z., Guan, P., Heiman, T.J., Higgins, M.E., Ji, R.R., Ke, Z., Ketchum, K.A., Lai, Z., Lei, Y., Li, Z., Li, J., Liang, Y., Lin, X., Lu, F., Merkulov, G.V., Milshina, N., Moore, H.M., Naik, A.K., Narayan, V.A., Neelam, B., Nusskern, D., Rusch, D.B., Salzberg, S., Shao, W., Shue, B., Sun, J., Wang, Z., Wang, A., Wang, X., Wang, J., Wei, M., Wides, R., Xiao, C., Yan, C., et al.: The sequence of the human genome. Science 291 (2001) 1304-51.

Verkerk, A.J., Pieretti, M., Sutcliffe, J.S., Fu, Y.H., Kuhl, D.P., Pizzuti, A., Reiner, O., Richards, S., Victoria, M.F., Zhang, F.P. and et al.: Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell 65 (1991) 905-14.

Vesa, J., Hellsten, E., Verkruyse, L.A., Camp, L.A., Rapola, J., Santavuori, P., Hofmann, S.L. and Peltonen, L.: Mutations in the palmitoyl protein thioesterase gene causing infantile neuronal ceroid lipofuscinosis. Nature 376 (1995) 584-7.

Virtaneva, K., D'Amato, E., Miao, J., Koskiniemi, M., Norio, R., Avanzini, G., Franceschetti, S., Michelucci, R., Tassinari, C.A., Omer, S., Pennacchio, L.A., Myers, R.M., Dieguez-Lucena, J.L., Krahe, R., de la Chapelle, A. and Lehesjoki, A.E.: Unstable minisatellite expansion causing recessively inherited myoclonus epilepsy, EPM1. Nat Genet 15 (1997) 393-6.

Virtaneva, K., Miao, J., Träskelin, A.L., Stone, N., Warrington, J.A., Weissenbach, J., Myers, R.M., Cox, D.R., Sistonen, P., de la Chapelle, A. and et al.: Progressive myoclonus epilepsy EPM1 locus maps to a 175-kb interval in distal 21q. Am J Hum Genet 58 (1996) 1247-53.

Vogt, P.K. and Bos, T.J.: Jun: oncogene and transcription factor. Adv Cancer Res 55 (1990) 1-35.

Wagner, E.F.: AP-1--Introductory remarks. Oncogene 20 (2001) 2334-5.

Wallace, D.C., Zheng, X.X., Lott, M.T., Shoffner, J.M., Hodge, J.A., Kelley, R.I., Epstein, C.M. and Hopkins, L.C.: Familial mitochondrial encephalomyopathy (MERRF): genetic, pathophysiological, and biochemical characterization of a mitochondrial DNA disease. Cell 55 (1988) 601-10.

Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S.E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., Bork, P., Botcherby, M., Bray, N., Brent, M.R., Brown, D.G., Brown, S.D., Bult, C., Burton, J., Butler, J., Campbell, R.D., Carninci, P., Cawley, S., Chiaromonte, F., Chinwalla, A.T., Church, D.M., Clamp, M., Clee, C., Collins, F.S., Cook, L.L., Copley, R.R., Coulson, A., Couronne, O., Cuff, J., Curwen, V., Cutts, T., Daly, M., David, R., Davies, J., Delehaunty, K.D., Deri, J., Dermitzakis, E.T., Dewey, C., Dickens, N.J., Diekhans, M., Dodge, S., Dubchak, I., Dunn, D.M., Eddy, S.R., Elnitski, L., Emes, R.D., Eswara, P., Eyras, E., Felsenfeld, A., Fewell, G.A., Flicek, P., Foley, K., Frankel, W.N., Fulton, L.A., Fulton, R.S., Furey, T.S., Gage, D., Gibbs, R.A., Glusman, G., Gnerre, S., Goldman, N., Goodstadt, L., Grafham, D., Graves, T.A., Green, E.D., Gregory, S., Guigo, R., Guyer, M., Hardison, R.C., Haussler, D., Hayashizaki, Y., Hillier, L.W., Hinrichs, A., Hlavina, W., Holzer, T., Hsu, F., Hua, A., Hubbard, T., Hunt, A., Jackson, I., Jaffe, D.B., Johnson, L.S., Jones, M., Jones, T.A., Joy, A., Kamal, M., Karlsson, E.K., et al.: Initial sequencing and comparative analysis of the mouse genome. Nature 420 (2002) 520-62.

Watson, J.D. and Crick, F.H.: Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature 171 (1953) 737-8.

Weinhaeusel, A., Morris, M.A., Antonarakis, S.E. and Haas, O.A.: DNA deamination enables direct PCR amplification of the cystatin B (CSTB) gene-associated dodecamer repeat expansion in myoclonus epilepsy type Unverricht-Lundborg. Hum Mutat 22 (2003) 404-8.

Wheeler, R.B., Sharp, J.D., Schultz, R.A., Joslin, J.M., Williams, R.E. and Mole, S.E.: The gene mutated in variant late-infantile neuronal ceroid lipofuscinosis (CLN6) and in nclf mutant mice encodes a novel predicted transmembrane protein. Am J Hum Genet 70 (2002) 537-42.

Wijmenga, C., Hewitt, J.E., Sandkuijl, L.A., Clark, L.N., Wright, T.J., Dauwerse, H.G., Gruter, A.M., Hofker, M.H., Moerer, P., Williamson, R. and et al.: Chromosome 4q DNA rearrangements associated with facioscapulohumeral muscular dystrophy. Nat Genet 2 (1992) 26-30.

Wilson, R.B.: Frataxin and frataxin deficiency in Friedreich's ataxia. J Neurol Sci 207 (2003) 103-5.

Woychik, N.A. and Hampsey, M.: The RNA polymerase II machinery: structure illuminates function. Cell 108 (2002) 453-63.

Xu, Q., Modrek, B. and Lee, C.: Genome-wide detection of tissue-specific alternative splicing in the human transcriptome. Nucleic Acids Res 30 (2002) 3754-66.

Yoshikawa, M., Uchida, S., Ezaki, J., Rai, T., Hayama, A., Kobayashi, K., Kida, Y., Noda, M., Koike, M., Uchiyama, Y., Marumo, F., Kominami, E. and Sasaki, S.: CLC-3 deficiency leads to phenotypes similar to human neuronal ceroid lipofuscinosis. Genes Cells 7 (2002) 597-605.

Zavasnik-Bergant, T. and Turk, B.: Cysteine cathepsins in the immune response. Tissue Antigens 67 (2006) 349-55.

Zeman, W. and Albert, M.: On the Nature of the "Stored" Lipid Substances in Juvenile Amaurotic Idiocy (Batten-Spielmeyer-Vogt). Ann Histochim 22 (1963) 255-7.

Zhang, S., Xu, L., Lee, J. and Xu, T.: Drosophila atrophin homolog functions as a transcriptional corepressor in multiple developmental processes. Cell 108 (2002) 45-56.

Zhao, C. and Meng, A.: Sp1-like transcription factors are regulators of embryonic development in vertebrates. Dev Growth Differ 47 (2005) 201-11.