

Hydrolethalus syndrome is caused by a missense mutation in a novel gene *HYLS1*

Lisa Mee^{1,†}, Heli Honkala^{5,†}, Outi Kopra^{5,6}, Jouni Vesa¹, Saara Finnilä¹, Ilona Visapää^{1,5}, Tzu-Kang Sang², George R. Jackson^{2,3,4}, Riitta Salonen⁷, Marjo Kestilä⁵ and Leena Peltonen^{1,5,8,*}

¹Department of Human Genetics, ²Neurogenetics Program, Department of Neurology, ³Brain Research Institute and ⁴Center for Neurobehavioral Genetics, Neuropsychiatric Institute, David Geffen School of Medicine, University of California at Los Angeles, Los Angeles, CA, USA, ⁵Department of Molecular Medicine, National Public Health Institute, Helsinki, Finland, ⁶Neuroscience Center, University of Helsinki, Helsinki, Finland, ⁷Department of Medical Genetics, Family Federation of Finland, Helsinki, Finland and ⁸Department of Medical Genetics, University of Helsinki, Helsinki, Finland

Received January 14, 2005; Revised March 11, 2005; Accepted April 8, 2005

Hydrolethalus syndrome (HLS) is an autosomal recessive lethal malformation syndrome characterized by multiple developmental defects of fetus. We have earlier mapped and restricted the HLS region to a critical 1 cM interval on 11q23–25. The linkage disequilibrium (LD) and haplotype analyses of single nucleotide polymorphism (SNP) markers helped to further restrict the HLS locus to 476 kb between genes *PKNOX2* and *DDX25*. An HLS associated mutation was identified in a novel regional transcript (GenBank accession no. FLJ32915), referred to here as the *HYLS1* gene. The identified A to G transition results in a D211G change in the 299 amino acid polypeptide with unknown function. The *HYLS1* gene shows alternative splicing and the transcript is found in multiple tissues during fetal development. *In situ* hybridization shows spatial and temporal distributions of transcripts in good agreement with the tissue phenotype of HLS patients. Immunostaining of *in vitro* expressed polypeptides from wild-type (WT) cDNA revealed cytoplasmic staining, whereas mutant polypeptides became localized in distinct nuclear structures, implying a disturbed cellular localization of the mutant protein. The *Drosophila melanogaster* model confirmed these findings and provides evidence for the significance of the mutation both *in vitro* and *in vivo*.

INTRODUCTION

Hydrolethalus syndrome (HLS; www.ncbi.nlm.nih.gov/Omim, MIM 236680) is a lethal malformation syndrome leading to stillbirth or death shortly after birth. HLS is characterized by hydrocephaly with absent upper midline structures of the brain, micrognathia and polydactyly. Various other features such as cleft lip or palate, club feet, anomalies of the ears, eyes and nose, a keyhole-shaped defect in the occipital bone, abnormal genitalia as well as congenital heart and respiratory organ defects have also been observed in affected individuals (1). Nowadays, HLS fetuses are efficiently detected by ultrasound scan (2,3) by 13–15 weeks of gestation. Characterization of the underlying molecular defect is of wider general significance as it potentially exposes new molecular pathway(s)

essential for normal human development and could provide new insights into fetal malformation syndromes.

This disease is inherited in an autosomal recessive manner (4). HLS belongs to the Finnish disease heritage (5) with the incidence of at least 1:20 000 (4) and the distribution of families implies a founder effect suggesting one major founder mutation (6). HLS-type cases have also been reported outside Finland (7–9), and some of the publications have described a significantly milder form of HLS. For example, Aughton and Cassidy (10) presented a case that survived >5 months and de Ravel *et al.* (11) described a case that died at the age of 7 months. However, concerning the partially different pathological findings compared with the Finnish HLS fetuses, the possibility exists that the non-Finnish cases represent other fetal malformation syndromes resembling HLS.

*To whom correspondence should be addressed at: Biomedicum Helsinki, Haartmaninkatu 8, 00290 Helsinki, Finland. Tel: +358 947448393; Fax: +358 947448480; Email: leena.peltonen@ktl.fi

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

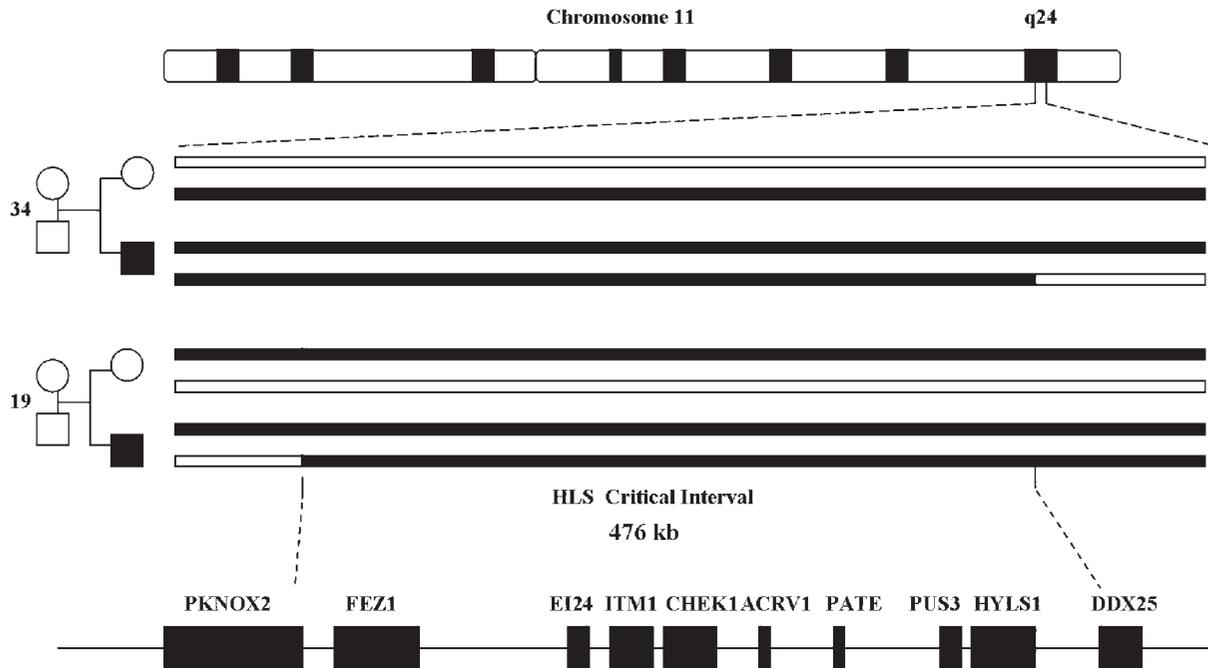


Figure 1. Fine resolution and physical map of HLS locus on 11q24. Recombinations in the affected children of families 19 and 34 restrict the HLS locus to a critical interval of 476 kb between genes *PKNOX2* (rs1630030) and *DDX25* (rs660619). Affected haplotypes are represented as shaded bars. The location of nine known genes (*PKNOX2*, *FEZ1*, *EI24*, *ITM1*, *CHEK1*, *ACRV1*, *PATE*, *PUS3* and *DDX25*) and one predicted transcript (FLJ32915; *HYLS1*) sequenced for mutation detection are represented as shaded boxes.

We earlier mapped the HLS locus to 11q23–25 by genome wide scan in eight Finnish HLS families. Through linkage disequilibrium (LD) and haplotype analyses, this region was restricted from an initial 8.5 cM interval to the 1 cM region between markers D11S933 and D11S934 (6). Here, we have defined the critical HLS region to a 476 kb interval using cases from a total of 20 Finnish HLS families. Sequence analysis of all the genes and transcripts on the critical DNA region revealed a point mutation in a novel gene resulting in a change of a well-conserved residue from Asp to Gly in a predicted polypeptide of unknown function. Although *HYLS1* does not possess any matching domains to any known polypeptides, its wide tissue expression pattern, well agreeing with the clinical phenotype of HLS, suggests a gene essential for early development of the human fetus.

RESULTS

Finemapping of the HLS locus

To further restrict the critical HLS region of 1.4 Mb (6), we analyzed 16 Finnish HLS families and four unrelated Finnish HLS individuals using eight markers over the critical region (D11S1752, D11S933, D11S4158, D11S1896, D11S934, D11S4110, D11S990 and D11S975). On the basis of ancestral haplotype and high LD observed between markers D11S4158 and D11S975, the previously assigned 1.4 Mb HLS region between markers D11S933 and D11S934 was further restricted to a 904 kb interval. A physical contig over the critical interval was electronically assembled using bacterial artificial chromosome (BAC)

clones retrieved from publicly available databases (NCBI, UCSC). In the 904 kb critical region, we identified nine genes of known function (*PKNOX2*, *FEZ1*, *EI24*, *ITM1*, *CHEK1*, *ACRV1*, *PATE*, *PUS3* and *DDX25*) and one hypothetical transcript (FLJ32915; *HYLS1*) for mutation analysis (Fig. 1). Hypothetical transcripts were identified by comparing expressed sequence tag (EST) clusters with the genomic sequence. All coding regions and flanking intronic sequences were sequenced from genomic DNA of two controls and two affected individuals, homozygotes for the characteristic haplotype of disease alleles.

All family material was genotyped for each informative single nucleotide polymorphism (SNP) identified by sequencing. The observed ancestral recombinations in families 19, 34, 36, 40 and 43 restricted the HLS region to a 476 kb critical interval between the *PKNOX2* and the *DDX25* genes (Fig. 1 and Table 1). To further restrict this region, SNPs from eight regional candidate genes were monitored for LD in disease alleles (Table 2). One major haplotype, D-A-1-T-A-1-A-10-1-G (Table 1), was observed in all analyzed HLS affected chromosomes. For LD, one SNP within a GenBank predicted transcript FLJ32915 showed a highly significant P -value of $10^{-\infty}$. At this SNP, all HLS patients were homozygous for the G allele, whereas all parental, non-transmitted chromosomes carried an A allele at this position. Also, four patients without family members were homozygous for this SNP. Interestingly, father of the family 22 was homozygous for all the informative markers over the entire critical region, with the exception of the A to G SNP located in the coding region of FLJ32915 for which he was heterozygous (Table 1). This SNP thus represented the only nucleotide

Table 1. Haplotypes of the hydrolethalus-carrying chromosomes

Family	D11S1752	D11S933	D11S4158	PKNOX2.1	PKNOX2.2	NA FEZ1	rs97570 FEZ1	EI24	rs11220159 ITM1	rs3731395 CHEK1	rs3731480 CHEK1	NA ACRV1	HYLS1	PUS3	mutation HYLS1	rs660619 DDY25	rs683155 DDY25	D11S975	D11S990	D11S1896	D11S934
19P	6	10	8	1	1	D	A	1	T	A	1	A	10	1	G	11	T	5	10	7	17
19M	6	6	4	2	2	D	A	1	T	A	1	A	10	1	G	11	T	5	10	7	7
22P	6	6	4	2	2	D	A	1	T	A	1	A	10	1	G	11	T	5	10	7	16
22M	6	6	4	2	2	D	A	1	T	A	1	A	10	1	A	11	T	5	10	7	16
22M	4	6	4	2	2	D	A	1	T	A	1	A	10	1	G	11	T	5	10	7	16
34P	6	6	4	2	2	D	A	1	T	A	1	A	10	1	G	10	C	10	6	7	4
34M	6	6	4	2	2	D	A	1	T	A	1	A	10	1	G	11	T	5	10	7	16
36P	6	6	4	2	2	D	A	1	T	A	1	A	10	1	G	11	T	5	10	7	7
36M	6	4	4	1	1	D	A	1	T	A	1	A	10	1	G	11	T	5	10	7	16
40P	6	6	4	2	2	D	A	1	T	A	1	A	10	1	G	11	T	5	10	7	16
40M	6	6	4	2	2	D	A	1	T	A	1	A	10	1	G	10	C	nd	6	7	15
43P	6	6	4	2	2	D	A	1	T	A	1	A	10	1	G	11	T	5	10	7	16
43M	6	6	4	2	2	D	A	1	T	A	1	A	10	1	G	10	C				6

5' UTR (D = G del.)

exon 2

exon 12.3' intron

5' UTR

3' UTR:GAAA ins. (1 = ins.)

exon 3:5' intron

T-stretch exon 4

exon 6

ins./del.intron 4

exon 10

Horizontal shading indicates alleles of the most common haplotype; vertical shading indicates the HLS-causing mutation. P, paternal allele; M, maternal allele; PKNOX2.1: 1, GTG; 2, ACA; PKNOX2.2:1, CG; 2, TG; EI24: 1, AGG; PUS3:1, CAGC, nd, not determined; UTR, untranslated region; del., deletion; ins., insertion.

change completely co-segregating with the disease in the pedigrees implying that it represents the disease-causing mutation.

To determine the population frequency of this variant, 908 control chromosomes collected from various regions of Finland and a control panel of 98 chromosomes of mixed European descent were genotyped. Among the chromosomes originating from the late settlement central and eastern Finland, 2.5% (14/556) carried this change, whereas 1.1% (4/352) chromosomes collected from the early settlement western Finland carried the change and no carriers were identified in the European control panel.

HYLS1

The *HYLS1* gene consists of six exons spanning a genomic region of ~17 kb having several alternative transcripts. *HYLS1a* (NCBI: AK127394) appears to be the longest isoform of this cDNA (2064 bp) consisting of exons 2, 4 and 6 (Fig. 2A). The length of exon 2 varies; for example, the isoform CR625776 is 1803 bp long. An open reading frame of 897 nucleotides is fully covered by the last exon (Fig. 2B) and is represented as an FLJ32915 variant in the NCBI's AceView (www.ncbi.nlm.nih.gov/IEB/Research/AceView/index.html).

The A to G transition is located in exon 6 resulting in a D211G change in the polypeptide of 299 amino acids. On the basis of the observed EST homologies, this isoform can be identified in uterus, brain, breast, skin, kidney, liver, lung, fetus, placenta, colon and pooled germ cell tumors (obtained by BLAST analysis of *HYLS1a* to ESTs from public databases). Variant *HYLS1b* consists of exons 2 and 6 (Fig. 2B) and is supported by clones BQ686872 and CB141176 from pancreatic carcinoma and liver, respectively.

When studying the cDNA from the fibroblast cell lines of patients and controls, we found these two isoforms. In addition, three other isoforms representing alternative splicing patterns have been published in Celera and NCBI databases, according to thorough database searches. However, none of these three isoforms could be experimentally verified with RT-PCR from the RNA of fetal samples. Each of the transcripts has the same conserved translated region coded by exon 6 as *HYLS1a* and *HYLS1b* (Fig. 2A).

Expression analysis of the human *HYLS1* transcript

To study the expression pattern of each *HYLS1* isoform, RT-PCR analyses of a human fetal cDNA panel were performed

Table 2. LD analyses of SNPs in eight regional candidate genes in hydrolethalus and parental control chromosomes

Locus	Percentage of hydrolethalus chromosomes ($n = 44$)/percentage of control chromosomes ($n = 34$) ^a								χ^2 LRT ^b	<i>P</i> -value	λ^c
	Allele 1	Allele 2	Allele 3	Allele 4	Allele 5	Allele 6	Allele 7	Allele 8			
D11S4158	70.5/21.2	2.3/3	6.8/9.1	9.1/39.4	11.4/27.3	–	–	–	16.3	2.8×10^{-5}	0.611
<i>PKNOX2</i>	87.5/0	2.5/40.6	5/0	0/12.5	5/15.6	0/31.2	–	–	61.2	3×10^{-15}	0.872
<i>FEZ1</i>	90.9/26.5	9.1/73.5	–	–	–	–	–	–	34.3	2.4×10^{-9}	0.872
<i>EI24</i>	95.2/13.3	2.4/56.7	2.4/0	0/30	–	–	–	–	57.6	1.7×10^{-14}	0.971
<i>ITM1</i>	92.5/3.3	7.5/96.7	–	–	–	–	–	–	60.6	4×10^{-15}	0.920
<i>CHEK1</i>	90.9/11.8	9.1/88.2	–	–	–	–	–	–	51.4	3.9×10^{-13}	0.893
<i>ACRV1</i>	92.9/6.3	7.1/93.7	–	–	–	–	–	–	60.2	5×10^{-15}	0.921
<i>HYLS1</i>	100/0	0/100	–	–	–	–	–	–	98.4	$10^{-\infty}$	1.000
<i>DDX25</i>	100/26.9	0/73.1	–	–	–	–	–	–	40.6	9.7×10^{-11}	1.000
D11S975	0/0	80/32.3	2.5/22.6	0/0	0/6.5	2.5/12.9	15/12.9	0/12.9	23.8	5.9×10^{-5}	0.693

Bold indicates highest association to hydrolethalus syndrome.

^aAs DNA samples of the other family members of four cases were not available, the control number of chromosomes is less than the number of HLS chromosomes.

^b χ^2 statistic of the likelihood-ratio test (23).

^cProportion of excess of a certain allele in the disease-carrying chromosome.

using combinations of primers crossing various *HYLS1* exon boundaries. Fetal cDNA was used as *HYLS1* must be critical for normal embryonic development. The observed PCR fragments (using primers in exons 2 and 6) corresponded to the expected sizes of transcripts *HYLS1a* and *HYLS1b*, the only *HYLS1* isoforms ubiquitously expressed in fetal tissue. RT-PCR analysis of the available *HYLS1* sequence was also performed on placenta, control and patient (A and B) fibroblast cDNAs. The expected size of the *HYLS1a* transcript was found in the patients' tissues and cells. Sequence analysis of the RT-PCR product did not reveal any additional nucleotide changes or splice variants in the affected individual.

To further monitor the changes in expression levels or splice variants of the *HYLS1* gene in affected individuals, northern blot analysis of control and patient (C) mRNA was performed. An ~1.7 kb transcript was detected in both control and patient fibroblasts with no significant difference in the steady state expression level. While monitoring for expression levels in different fetal tissues, the transcript levels appeared highest in brain, liver, lung and kidney.

Identification and analysis of the mouse *Hyls1* gene

To identify the *HYLS1* mouse ortholog, the human *HYLS1a* cDNA was used as a query sequence against NCBI and Celera databases. Mouse ESTs AK013896 (NCBI) and mCT127332 (Celera) were identified as *HYLS1* orthologs. As EST AK013896 perfectly aligned to the mouse genome sequence, this transcript was used here as the reference sequence. This mouse *Hyls1* transcript was confirmed by sequencing of RT-PCR products from mouse post-embryonic brain cDNA and by alignment to the human *HYLS1* transcript. The mouse transcript is predicted as 2.1 kb by Celera and was confirmed by ~2 kb transcript observed in northern blot analysis. *Hyls1* is composed of four exons with an ORF of 984 nucleotides producing a predicted polypeptide of 328 amino acids. Exon 4 is the only translated *Hyls1* exon and bears a high sequence similarity to exon 6 of *HYLS1*. No significant homologies with other *HYLS1* exons have been found. The ORF in human

and mouse are highly conserved with 84% sequence identity at the mRNA level and 83% at the amino acid level. The aspartic acid mutated in HLS patients is also conserved in the *Hyls1* transcript in addition to the flanking amino acid sequence: 28 residues upstream (following the first adjacent residue) and 14 residues downstream of Asp (Fig. 3). Further support for the *Hyls1* transcript to represent the human *HYLS1* ortholog emerges from the chromosomal position of the mouse sequence on chromosome 9, the region syntenic to the human 11q24 region. This aspartic acid is also conserved across other species as low as *Caenorhabditis elegans*. Designated as residue 211 in the human *HYLS1* polypeptide, this conserved aspartic acid appears as amino acid 231 in Figure 3 due to additional 5' protein sequence in other species as well as gaps in the sequence alignment.

Expression analysis of the mouse *Hyls1* transcript

RT-PCR analysis of post-embryonic mouse brain cDNAs revealed that this transcript is expressed from P1 to 6 months in mouse. Hybridization of the mouse embryo northern blot panel revealed an ~2 kb transcript in each embryonic stage, the highest expression level observed at E11.

To characterize the spatial expression pattern of *Hyls1* during development, we performed *in situ* hybridizations on whole embryo (E12.5 and E15.5) and adult mouse brain (3 months) sections. Figure 4A–F shows sagittal sections of an E15.5 mouse embryo. At this stage, the *Hyls1* expression can be observed in multiple tissues being particularly high in the developing central nervous system (CNS). Expression in the nervous system is also evident in the spinal cord and in the dorsal root ganglia. In the cephalic region of the E15.5 embryo, *Hyls1* is detected in the telencephalon, the mid-brain, the medulla, the choroid plexus and the ganglionic eminence. Expression was also seen in the neural layer of the retina (data not shown). *Hyls1* signal was detected in developing organs such as the cardiovascular and respiratory systems, especially in the epithelium of the lungs. Strong expression was detected also in stomach, testis, pancreas, intestine and kidney and in the cartilage tissue of developing

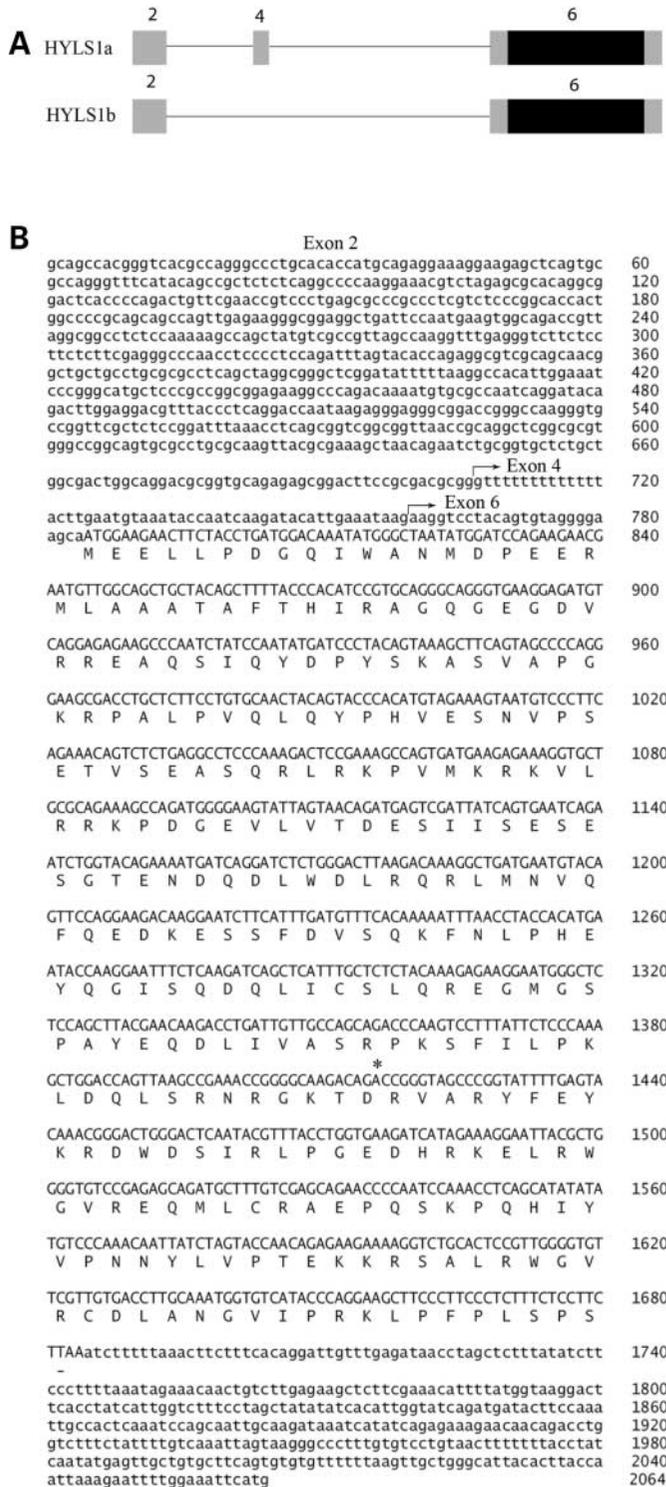


Figure 2. Genomic structure and alternative splicing of the *HYLS1* gene. *HYLS1* is composed of six putative exons and spans a region of ~17 kb. (A) Ubiquitously expressed transcripts *HYLS1a* and *HYLS1b* are shown. *HYLS1a* (AK127394) is composed of exons 2, 4 and 6, whereas *HYLS1b* is composed of exons 2 and 6. Black shading represents the translated region conserved across all *HYLS1* variants and gray shading represents untranslated regions. (B) The cDNA sequence of *HYLS1a* (AK127394) with the predicted amino acid translation of the open reading frame following the first methionine. A to G mutation is located at nucleotide 1416 (asterisk). Arrows represent sites of exon boundaries.

limbs. Figure 4G shows coronal sections of 3 month adult brains. *Hyls1* is predominately expressed in both CA1 and CA3 fields of the hippocampus and the cortex. In E12.5 mouse embryos, there were detectable expression levels in many regions including the developing brain, neural tube, heart, lung epithelium, olfactory and oral epithelium, cloaca, urethra and cartilage tissue (data not shown).

Expression and stability of wild-type (WT) and D211G mutant *HYLS1* polypeptides

HYLS1 is predicted to encode a 299 amino acid polypeptide with a deduced molecular weight of 34.4 kDa. To determine the relative molecular weight of the *HYLS1* protein, *in vitro* transcription and translation were performed using an *in vitro* rabbit reticulocyte lysate system. The *in vitro* translation product of both WT and D211G mutant constructs produced polypeptide chains of ~39 and 40 kDa (data not shown). The identity of the 39 kDa polypeptide is most likely the result from the utilization of an alternative initiation as a downstream internal Met residue can be identified at amino acid 14, the 40 kDa polypeptide representing the complete *HYLS1* polypeptide. To monitor for the relative stability of the WT and the mutant *HYLS1* polypeptide, COS-1 cells were transfected with WT and D211G mutant *HYLS1* constructs. The expression of WT and mutant *HYLS1* resulted in immunopositive bands of ~40 kDa, which is in good agreement with the molecular weight of the *in vitro* translated polypeptides. Western blot analyses of WT and mutant proteins showed no significant difference between the corresponding protein levels (Fig. 5). To address the potential effect of the D211G mutation on the stability of the polypeptide, the quantitation of both the WT and the mutant proteins was performed by pulse-chase analyses. In agreement with the western blot analyses, pulse-chase analyses did not reveal any significant difference in the amount of immunoprecipitated polypeptides between cells transfected with WT and mutant constructs. The half-lives of transiently expressed WT and mutant *HYLS1* polypeptides were ~3 h (data not shown).

Subcellular localization of WT and D211G *HYLS1* polypeptides *in vitro*

To determine the cellular destination of the WT *HYLS1* polypeptide and to determine whether D211G would result in alterations of intracellular trafficking of this protein, COS-1 cells were transiently transfected with either WT or mutant *HYLS1* constructs, and synthesized polypeptides were monitored using rabbit anti-*HYLS1* antibody. Three different staining patterns were revealed using DAPI staining as a marker (data not shown) (Table 3). WT transfected cells showed characteristically intense staining of cytoplasm (Fig. 6A), whereas most mutant polypeptides localized to nuclear inclusions (Fig. 6B). Although WT and mutant cells had a similar percentage of both cytoplasmic and nuclear inclusions, the proportion of cytoplasmic to nuclear inclusions observed in WT cells was significantly higher than in mutant cells. The co-localization experiments using antibodies specific to

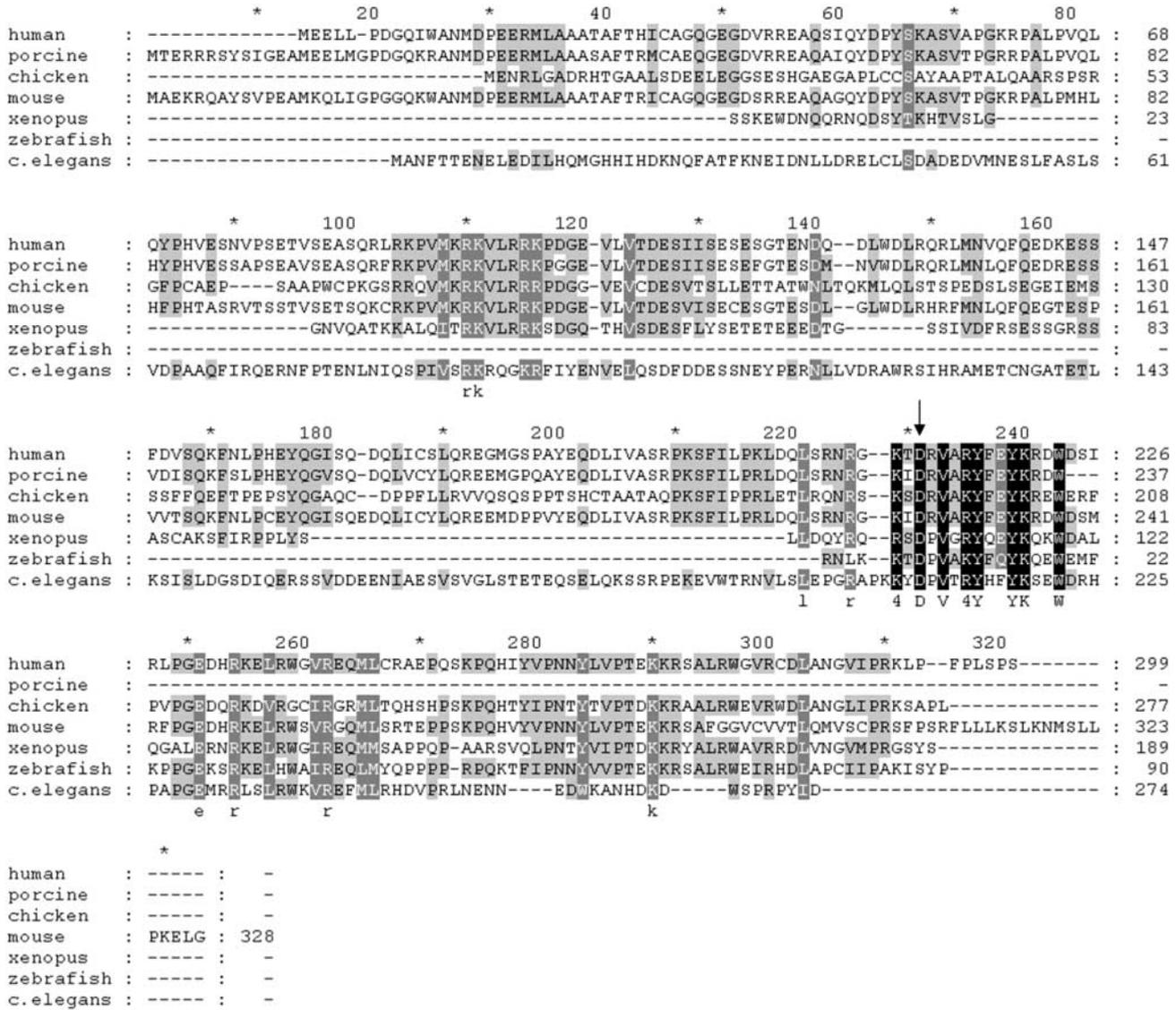


Figure 3. Comparison of human (NP_659451.1), porcine (BP435891), mouse (BAB29044.1), chicken (CR354360), xenopus (BJ623896), zebrafish (BM034949) and *C. elegans* (AAK95869.1) HYLS1 orthologs using ClustalW followed by GeneDoc. Arrow indicates the mutated amino acid conserved between species. This conserved aspartic acid corresponds to amino acid 211 in the human sequence. Owing to the additional 5' protein sequence in other species and gaps in the sequence alignment, this aspartic acid appears as amino acid 231 in the figure.

Golgi, ER or lysosomes failed to show any significant colocalization of the immunostaining (data not shown).

The specificity of the anti-HYLS1 antibody was confirmed by tagging both WT and mutant HYLS1 constructs with separate FLAG and Myc tags. COS-1 cells were transiently transfected individually with either FLAG-tagged, Myc-tagged or untagged WT or mutant HYLS1 constructs which were monitored using mouse anti-FLAG, mouse anti-Myc and rabbit anti-HYLS1 antibodies, respectively. DAPI staining was used as a nuclear marker. As observed in Figure 7A–C, FLAG-tagged, Myc-tagged and untagged WT HYLS1 polypeptides all localized to cytoplasmic vesicles, whereas FLAG-tagged, Myc-tagged and untagged mutant HYLS1 proteins showed identical staining patterns in the nucleus (Fig. 7D–F).

Expression of HYLS1 in *D. melanogaster*

To monitor whether antibodies against human HYLS1 peptides represent useful tools in the analyses of transgenic *D. melanogaster* tissues, western blot analysis of proteins extracted from adult fly heads expressing WT HYLS1 and from *yw* controls was performed. An immunopositive band corresponding to the size of the HYLS1 protein (~40 kDa) was observed in head lysates expressing WT HYLS1, whereas no band was observed in control head lysates (data not shown).

To determine the subcellular localization of HYLS1 polypeptides *in vivo*, WT and mutant HYLS1 were overexpressed in *D. melanogaster* third instar eye discs using a GMR-GAL4 driver. The pGMR-transgene is expressed posterior to the

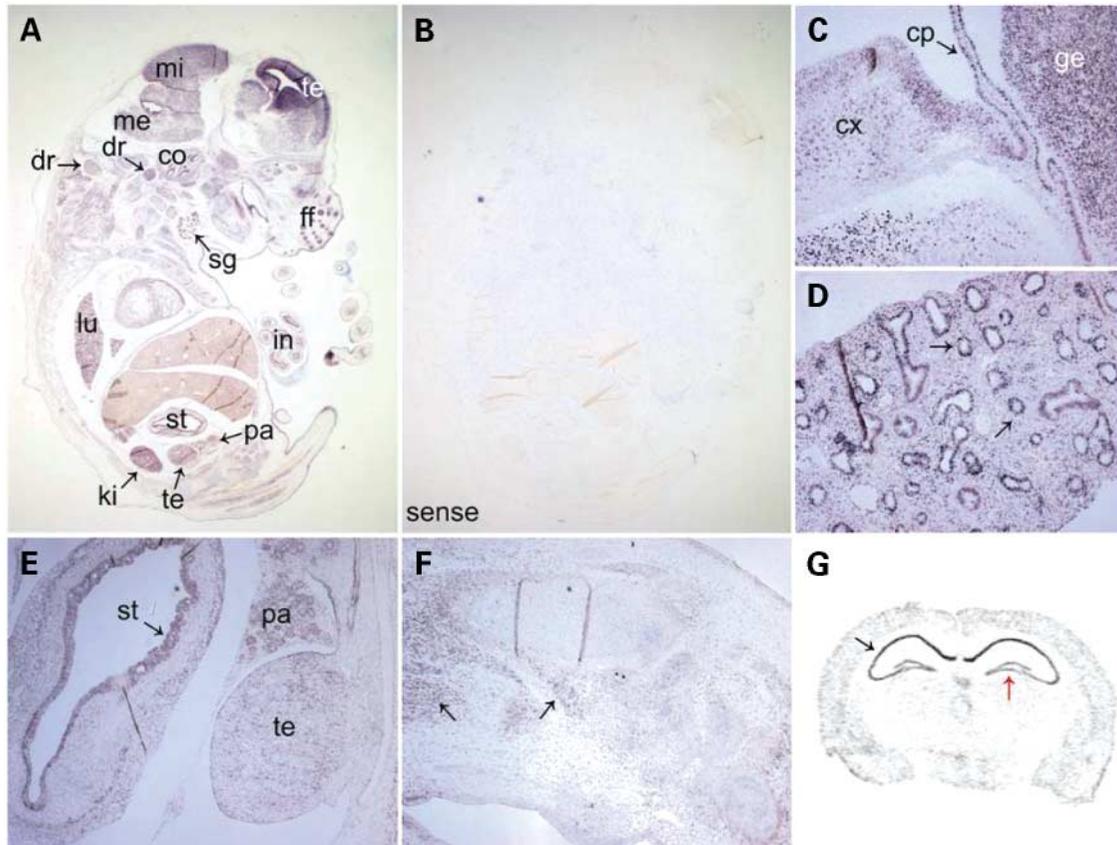


Figure 4. Expression profile of the *Hyls1* gene during mouse development. Sagittal sections of the whole embryo at embryonic day (E) 15.5 with anti-sense (A) and sense (B) RNA probes show ubiquitous expression pattern throughout the embryo. In the head, expression is seen in the brain including the telencephalon (te), midbrain (mi) and medulla (me). Expression is prominent also in the dorsal root ganglia (dr), cochlea (co), submandibular gland (sg) and follicles of fibrissae (ff). Several developing organs show expression of *Hyls1* including lungs (lu), kidney (ki), testis (te), pancreas (pa), stomach (st) and intestine (in). Higher magnification figures show expression in the brain (C), including the developing cortex (cx), choroid plexus (cp) and ganglionic eminence (ge), in the epithelial cells of the lung (arrows) (D) and in several organs including testis (te), pancreas (pa) and the epithelial layer of stomach (st) (E). Expression is also seen in the cartilage tissue of the developing limbs (arrows) (F). Coronar section of 3-month adult brains (G) shows strong expression in hippocampus (black arrow) and dentate gyrus (red arrow). Original magnifications: A and B 1.6 \times , C and D 20 \times and E and F 10 \times .

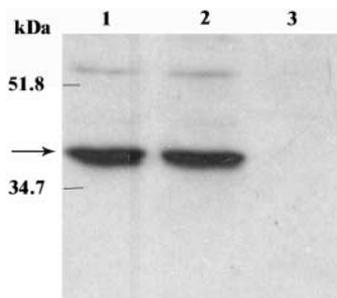


Figure 5. Western blot analysis of HYLS1 polypeptides. COS-1 cells transiently transfected with WT and mutant HYLS1 constructs were subjected to western blot analysis using rabbit anti-HYLS1 antibody. Each lane was loaded with 20 μ g protein. Lane 1: WT HYLS1. Lane 2: mutant HYLS1. Lane 3: untransfected COS-1 cell lysate. Arrow indicates detection of HYLS1 polypeptides at \sim 40 kDa.

morphogenetic furrow in the eye imaginal disc and was, accordingly, the site of HYLS1 expression. Immunostaining of larval eye discs with rabbit anti-HYLS1 revealed the same cellular localization pattern observed *in vitro* for both

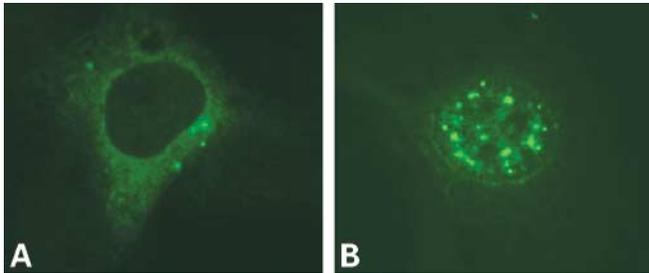
WT and mutant HYLS1 expressing cells. Discs were double stained with anti-Elav, a nuclear marker. Similar to findings in COS-1 cells, cells expressing WT HYLS1 in the eye disc predominantly showed cytoplasmic staining with some perinuclear inclusions (Fig. 8A–C), whereas cells expressing mutant HYLS1 in the eye disc revealed distinct staining of nuclear aggregates (Fig. 8D–F). HYLS1 was not observed in controls (Fig. 8G–I).

Bioinformatic analysis of the *HYLS1* gene and polypeptide

Web Promoter Scan Service (PROSCAN) predicted two most likely promoter regions for *HYLS1* to a region 3.4 kb upstream from the first exon (nucleotides –3382–3133) and inside the gene in intron 3 (nucleotides 4626–4876). When comparing the pI values calculated with ExPASy's tool, the WT protein pI value was 6.98 versus 7.67 for mutant protein. PeptideCutter estimated an Asp-N endopeptidase cleavage site at position 211 for normal HYLS1 polypeptide, whereas this cleavage site was absent from the mutant form of HYLS1. According to Pfam database, there was a low complexity region between

Table 3. Distribution of HYLS1 polypeptide in transiently transfected COS-1 cells

HYLS1 construct	Cytoplasm (%)	Cytoplasm and nucleus (%)	Nucleus (%)
WT	49.4	48.8	1.8
D211G mutant	6.2	53	40.7

**Figure 6.** The subcellular localization of WT (A) and Asp211Gly mutant (B) HYLS1 proteins in transiently transfected COS-1 cells. The HYLS1 proteins were detected by anti-HYLS1 antibody and FITC-conjugated secondary antibody (green). Magnification = 63 \times .

amino acids 112 and 120. Such regions are typical for structurally less complex proteins (12).

DISCUSSION

Genetic evidence presented here strongly suggests that HLS results from a D211G amino acid change in the novel *HYLS1* gene. Initially, this gene was positionally cloned to a critical 904 kb interval on chromosome 11q24.2 using the power of LD and haplotype analyses (13). Sequencing of the coding regions of the 10 candidate genes in the region and detailed analyses of identified nucleotide variants revealed that the A to G change in the novel *HYLS1* gene was the only variant that segregated completely with the disease in all HLS families. The observed carrier frequency of this variant (2.5%) was slightly higher than the expected carrier frequency (2%) in the late settlement region of Finland, but as for most early malformation syndromes, it is likely that some cases remain non-diagnosed as HLS. There is no evidence of the phenotypic effects in heterozygote carriers as the clinical manifestations in the parents have not been systematically studied. However, one of the mothers has duplicated uterus, which is very common in females with HLS (1,4). One cousin of two affected individuals had thoracic meningocele with anomalies in the vertebral bodies, anal atresia with urethral fistula and left hand polydactyly. The significance of these findings and the impact of a single mutated copy of *HYLS1* gene to clinical phenotype should be addressed in future studies.

The Finnish *HYLS1* mutation changing Asp to Gly is expected to have a significant effect on the structure of the polypeptide. This Asp is strictly conserved across various species from *C. elegans* to human, suggesting that it is a critical amino acid for the normal function of the protein. In

addition, predicted consequences to the synthesized polypeptide are significant: Asp is bulky and negatively charged, whereas Gly is the smallest amino acid with a neutral charge facilitating turns in the polypeptide backbone. This transition in the mutant polypeptide changes the isoelectric point, most probably affecting the solubility of the protein. Other potentially relevant changes caused by the mutation include disappearance of a protease cleavage site in the mutant polypeptide by ExPASy's PeptideCutter tool.

Database searches (NCBI, Celera, UCSC) suggest that the *HYLS1* transcript has five isoforms. Although RT-PCR analyses of fetal cDNA confirmed the expression of only two of these five variants, *HYLS1a* and *HYLS1b*, it is possible that these variants are expressed throughout the fetal life, whereas the other *HYLS1* isoforms become expressed only during restricted periods of fetal development. Sequencing of the fibroblast cDNA identified *HYLS1a* as the major isoform, most likely correlating with the 1.7 kb bands identified in the northern blots. Further analyses will be required to determine the expression pattern of the various isoforms in different tissues and cell types. As exon 6 is the only untranslated region of the gene, alternative splicing most likely does not affect the structure of the synthesized *HYLS1* polypeptide. However, alternative splicing may affect the transcript structure and expression level of *HYLS1* in different cell types and developmental stages. Other genes with only one protein coding exon are often reported to contain a 5' untranslated region acting as a critical regulatory region for the transcription or translation (14,15). This scheme might hold for *HYLS1* as well.

Neither the *HYLS1* gene nor the corresponding polypeptide possesses any previously characterized domains (except the low complexity region) which would provide some clue of its function. However, the expression pattern of *Hyls1* based on *in situ* hybridization studies on mouse embryos correlate well with the disease phenotype, the expression pattern nicely reflecting the most affected tissues in HLS patients: CNS, cardiovascular system, lung, genitalia and cartilage of developing limbs. Expression of *Hyls1* in numerous embryonic tissues suggests its essential role in correct embryological development, particularly the development of the CNS.

Computer assisted prediction programs would indicate that both WT and mutant forms of the *HYLS1* polypeptide are soluble. Although SignalP analysis does not predict a nuclear localization signal (NLS), PSORTII analysis predicts a preferential nuclear localization of the *HYLS1* polypeptide (52%). Further, the sequence is deficient for signal peptide or any secretory signal. Our *in vitro* expression studies suggest the predominant location in the cytoplasm, whereas immunostaining of the mutant polypeptide reveals nuclear dots. This would indicate an alteration of the intracellular trafficking of the mutant *HYLS1* polypeptide in patients' cells. Further, our data on *D. melanogaster* transgenics (16) expressing the human *HYLS1* protein also show that the mutant form of the protein, opposite to the WT counterpart, forms nuclear aggregates in the eye disc of flies, providing additional evidence for the biological significance of the identified mutation. Our preliminary data from additional immunostaining studies suggest that mutant *HYLS1* aggregates colocalize with lamin and caspase-3 in the eye discs, implying a potential role for

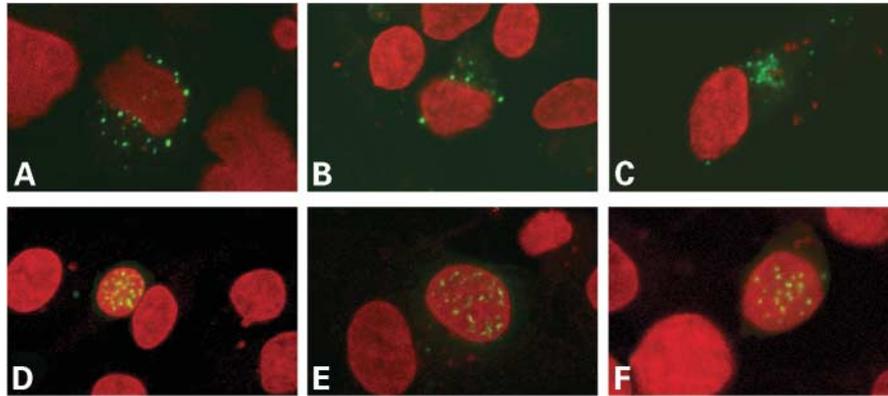


Figure 7. Confirmation of HYLS1 antibody specificity in transiently transfected COS-1 cells. (A–C) show cytosolic localization of WT HYLS1 polypeptides with FLAG-tagged (A), Myc-tagged (B) and untagged WT HYLS1 constructs (C). (D–F) show nuclear localization of mutant HYLS1 polypeptides with FLAG-tagged (D), Myc-tagged (E) and untagged mutant HYLS1 constructs (F). (A and D) green, anti-FLAG; (B and E) green, anti-Myc; (C and F) green, anti-HYLS1; (A–F) red, DAPI-stained nuclei; (A–F) yellow, merge.

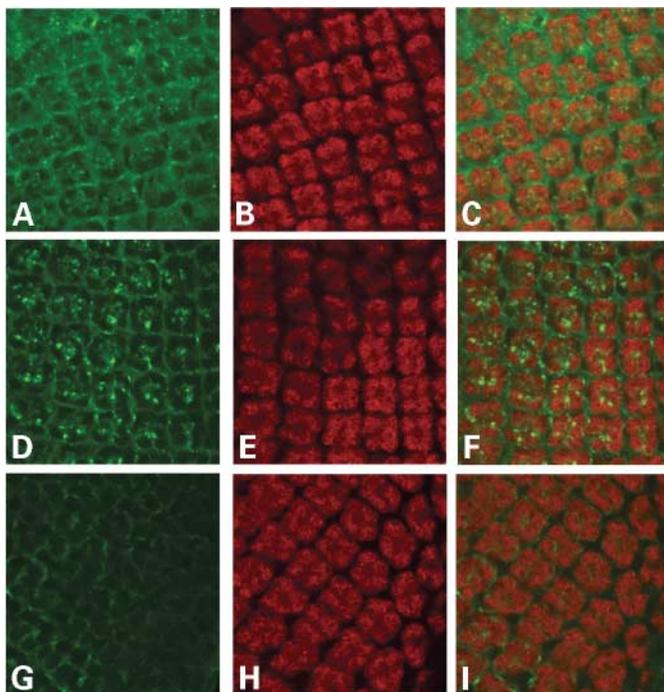


Figure 8. Subcellular localization of HYLS1 in third instar larval eye discs. (A–C) WT HYLS1 overexpressed in third instar larval eye discs stained with anti-HYLS1 and anti-Elav. WT HYLS1 shows localization to cytoplasm and perinuclear aggregates. (D–F) Overexpression of D211G mutant HYLS1 in third instar larval eye discs stained with anti-HYLS1 and anti-Elav. Mutant HYLS1 shows localization to nuclear inclusions and colocalizes with Elav. (G–I) Control w^{1118} third instar larval eye discs stained with anti-HYLS1 and anti-Elav. (green, rabbit anti-HYLS1; red, rat anti-Elav; yellow, merge). Magnification = 100 \times .

HYLS1 in apoptosis (data not shown). Further studies are required to understand the potential role of HYLS1 in this or any other biological pathway.

Our data would imply that normally cytoplasmic HYLS1 protein gets trapped in the nuclear dots when mutated. On the basis of this, two potential scenarios of the molecular

mechanism of HLS can be speculated: The WT HYLS1 protein could shuffle between the cytoplasm and the nucleus, being transported into the nucleus by a carrier protein and when mutated, binds more efficiently to the carrier resulting in more transportation of the protein into the nucleus. Alternatively, the WT HYLS1 protein could also diffuse into the nucleus due to its small size and when mutated binds to another protein or protein complex inside the nucleus causing it to essentially become stuck in the observed dot-like nuclear structures.

Many findings in this study, including that HYLS1 does not possess any secretory signal, the broad expression of HYLS1 in embryonic tissues and the immunostaining of the WT and mutant protein, would lend support to the hypothesis that HYLS1 could be a transcription factor. Altered localization between the cytoplasm and the nucleus is a common regulatory mechanism for numerous transcription factors involved in the developmental processes (17). Identification of protein partners interacting with HYLS1 will be critical for understanding the biological role of HYLS1, its role in the pathogenesis of HLS and the mechanisms underlying CNS development.

MATERIALS AND METHODS

Subjects

DNA samples isolated from 24 affected individuals and 40 unaffected family members from 16 unrelated Finnish HLS families were used for genotype and sequence analyses. In addition, DNA was obtained from four unrelated affected Finnish patients with no family members' samples. The affected children of two of the families had fewer anomalies than usual, but all of them were typical of HLS. Neither of these cases had polydactyly, but they both had hydrocephaly, defective lobation of the lungs, anomalies of the respiratory tract, small chin and anomalous nose. Four-hundred and fifty-four control DNA samples collected from western, central and eastern Finland, as well as 49 control DNA samples of mixed European descent, were used for the

carrier screening. Fibroblast cell lines from one control and three affected individuals from families 20, 34 and 40 (labeled A, B and C, respectively) were cultured for mRNA extraction. This study has been approved by the IRB committee of UCLA and the Joint Authority for the Hospital District of Helsinki and Uusimaa.

Genotyping

Eight dinucleotide microsatellite markers spanning the restricted 1 cM HLS locus were genotyped on the critical chromosomal region. Six of these markers (D11S1752, D11S933, D11S4158, D11S1896, D11S934 and D11S4110) were previously used in the initial HLS genome scan (4), whereas D11S990 and D11S975 were new dinucleotide markers identified from the NCBI (www.ncbi.nlm.nih.gov) and the Genome Database (www.gdb.org). The marker order was confirmed by blasting marker sequences to public (NCBI, UCSC, genome.cse.ucsc.edu) and private (Celera, public.celera.com/cds/login.cfm) databases. PCR amplification of each marker was performed using fluorescently labeled primers. Amplified products were separated in a 6.5% denaturing acrylamide gel using LI-COR 4200 DNA Analysis System and alleles were analyzed using SAGA Automated Genetic Analysis Software (LI-COR).

Sequence analysis

Coding regions of the positional candidate genes (*PKNOX2*, *FEZ1*, *EI24*, *ITM1*, *CHEK1*, *ACRV1*, *PATE*, *PUS3*, *FLJ32915* and *DDX25*) were amplified with primers in flanking intronic sequences using genomic DNA from two controls and two affected fetuses. Primers were designed based upon sequence information obtained from either NCBI or Celera database. Sequencing was subsequently performed on both strands using ABI 3700 and ABI Big Dye v.2 chemistry. Each informative SNP identified was subsequently genotyped using genomic DNA from all family material by PCR amplification and sequence analyses.

Haplotype and LD analyses

To identify the ancestral disease haplotype, chromosomes of affected individuals were compared with non-transmitted chromosomes. Haplotypes were constructed manually by assuming a minimum number of recombinations per family. Multiple informative SNPs identified in any single candidate gene (*PKNOX2*, *EI24* and *PUS3*) were arranged into haplotype blocks which were assigned a single number. The two-point LD analysis was performed using the DISLAMB computer program. A disease-allele frequency of 0.01 was assumed for the analyses.

Mutation analysis of the *HYLS1* gene

All exons of the *HYLS1* gene were initially PCR amplified and sequenced using genomic DNA from two controls and two affected individuals. PCR primers were designed using sequence information from NCBI and Celera. The PCR amplification was carried out using a 66–58°C touchdown program.

The purified PCR products were sequenced as described earlier.

The region harboring the mutation in exon 6 of the *HYLS1* gene was PCR amplified and sequenced from genomic DNA from all available family members using primers 5'-GG AATTTCTCAAGATCAGCTCA and 5'-CTAGGTTATCTCA ACAATCCTG.

mRNA was extracted from one patient (A) fibroblast cell line using Qiagen's Oligotex Direct mRNA midi/maxi kit (no. 72041) and cDNA was generated with an Advantage RT-for-PCR kit (BD Biosciences; no. 639505). From the control and two patient (B and C) fibroblast cell lines, mRNA was extracted using Qiagen's Oligotex Direct mRNA Mini Kit (no. 72022). Template cDNA was obtained using Invitrogen's SuperScript First-Strand Synthesis System for RT-PCR (no. 11904-018). Primers 5'-CGCGAAAGCTAAC AGAATCTG and 5'-TCTTAAGTCCCAGAGATCCTG, corresponding to exons 2 and 6, respectively, for placenta and control fibroblasts as well as primers in exon 2, 5'-GAAAG CTAACAGAATCTGCGG and in exon 6, 5'-CATCACTGGC TTTCGGAGTC, for cell line A were initially used for RT-PCR amplification and sequencing of cDNA in order to confirm the presence of this transcript. Primers flanking the entire available cDNA sequence of *HYLS1a*, the major isoform of the *HYLS1* gene, in exon 2 (5'-CGCGAAAGCT AACAGAATCTG) and exon 6 (5'-GTAAGTGTAATGCCCA GCAAC) were used to PCR amplify placenta, control fibroblast and patient fibroblast (B) cDNA. A 66–60°C touchdown program was used for PCR. Nested primers were used for sequencing.

Pyrosequencing

To determine the frequency of the *HYLS1* mutation, control samples were genotyped by pyrosequencing technique using the PSQ96 instrument and the SNP reagent kit (Pyrosequencing AB) (18). Genomic DNA of all control samples was initially PCR amplified using 10 μM forward primer (5'-CA AGACCTGATTGTTGCCAGC) and 10 μM reverse biotinylated primer (5'-CAAAGCATCTGCTCTCGGAC) in a 50 μl PCR reaction to produce a product length of 190 bp. PCR amplification was performed using a 66–58°C touchdown program of 38 cycles. An aliquot of 20 μM detection primer (5'-ACCGGGCAAGACAG) was used and designed using the SNP Primer Design Software Version 1.01.

Expression analysis of the *HYLS1* transcript

A human fetal cDNA panel (Clontech; K1425-1) was RT-PCR amplified using primers located in exons 1 (5'-GCAGA TCACCTGAGGTTAGG) and 6 (5'-TCTTAAGTCCCAGAG ATCCTG), exons 1 (5'-GCAGATCACCTGAGGTTAGG) and 2 (5'-CAGATTCTGTTAGCTTTCGCG), exons 1 (5'-GC AGATCACCTGAGGTTAGG) and 3 (5'-CCAGGCAGGACC TGGAATC), exons 2 (5'-GAAAGCTAACAGAATCTGCGG) and 3 (5'-CTTGACCCAGGTCGCTAG), exons 3 (5'-CAAC GTCTGCGTTCCTGAG) and 6 (5'-TCTTAAGTCCCAGAG ATCCTG) and exons 2 (5'-CGCGAAAGCTAACAGAATC TG) and 6 (5'-TCTTAAGTCCCAGAGATCCTG). A 68–60°C touchdown program was used for PCR amplification of

all PCR reactions except for the exons 3 and 6 reaction for which a 66–58° touchdown program was used.

Mouse brain cDNAs of varying post-embryonic developmental stages (P1, P7, P14, P30, 3 months and 6 months) were RT-PCR amplified using primers crossing boundaries of three exons (5'-GGATTCGCGTTTTCTCCCTG and 5'-CAGGTCCTATCAGTTGCTTC) of the *Hyls1* mouse ortholog obtained from NCBI (AK013896). A 66–58° touchdown program was used for RT-PCR amplification.

Control and patient (B) polyA mRNAs were extracted from human fibroblasts for northern blot analyses. A total of 2 µg of mRNA was run on a 0.8% gel and transferred to a Hybond-XL nylon membrane (Amersham Pharmacia Biotech). A northern blot probe covering 600 bp within the sixth exon was designed using primers 5'-GACTCCGAAAGCCAGTGATG and 5'-TGACACCATTTGCAAGGTCAC. PCR amplification products were run on a 1.5% agarose gel and bands were gel purified using Qiagen's QIAquick Gel Extraction Kit (no. 28706). For each probe, 20 ng of DNA was labeled with ³²P-dCTP/dGTP using Invitrogen's Random Primers DNA Labeling System (18187-013). The same probe was hybridized to a human fetal multiple tissue northern blot (Clontech; no. 7756-1). A mouse probe of 546 bp was generated by PCR amplification of *Hyls1* exon 4 from mouse brain cDNA using primers 5'-GCTAGTGACAGATGAGTCTG and 5'-CATGACACCATTGCAAGGTC. These primer sequences were derived from NCBI's sequence information (AK013896) and hybridized to a mouse embryo northern blot (Clontech; no. 7763-1). Labeling and purification were the same as described earlier. Hybridization of the fibroblast and Clontech blots was performed for 1 h using ExpressHyb (Clontech) at 68°C. Human β-actin mRNA served as a control for each blot.

***In situ* hybridization**

Radioactive *in situ* hybridization for mouse brain was performed as described previously (19). Slides were prepared in the following manner: 20 µm coronal brain sections from a 3-month-old WT post-embryonic mice were fixed in 4% paraformaldehyde (PFA), rinsed with 0.2 M phosphate buffer (PB) and stored at -70°C. Probes were designed by PCR amplifying the same region of cDNA used as a probe for the mouse northern blot as described earlier. Anti-sense and sense probes were generated by ligating either a T7 or an Sp6 promoter sequence, respectively, to 300 ng of PCR template using Ambion's Lig'n Scribe Kit (no. 1730). *In vitro* transcription was performed in the presence of ³⁵S-UTP using the Stratagene RNA transcription kit protocol (200340). Following hybridization, sections were washed, treated with RNase A (0.03 mg/ml Rnase A, 0.01 M Tris-HCl pH 8.0, 0.5 M NaCl and 0.25 µM NaCl), washed and dried. Slides were exposed to autoradiography film for 5 days, dipped in NTB2 emulsion (Kodak) and exposed at 4°C for 5 weeks. After being developed, sections were counterstained with hematoxylin and eosin and coverslipped with Permount. Images were captured with an Olympus digital camera and processed with Adobe Photoshop Elements 2.0.

Non-radioactive *in situ* hybridization was also performed to sections of mouse embryos (C57BL/National Public Health Institute, Helsinki, Finland) at embryonic days 12.5 (E12.5)

and 15.5 (E15.5). The embryos were fixed by overnight immersion in 4% PFA in 0.1 M PB, pH 7.4. Samples were embedded in paraffin and cut into 5 µm thick sagittal sections. Mounting was performed onto SuperFrost/Plus microscope slides (Merck, Germany). Animal care and handling were at all times consistent with the guidelines set out in the National Research Council's guide for laboratory animal care and use.

Probes were generated as described earlier in the mouse northern blot procedure. The probe fragment was cloned into the pGEM-T Easy Vector (Promega) and the anti-sense and sense probes with the digoxigenin label were created with a DIG RNA labeling kit (Roche), according to the manufacturer's protocol. The probes were diluted in 1:100 and in 1:200 in a hybridization buffer containing 2× SSC, 10% dextran sulfate, 0.01% sheared salmon sperm DNA, 0.02% SDS and 50% formamide. A method described by Breitschopf *et al.* (20) was used for the hybridization with slight modifications. Photographing and analysis of the samples were performed using a Leica microscope equipped with DC300F camera and IM1000 software and a Zeiss Axioplan 2 imaging microscope (Zeiss). The final figures were prepared using CorelDraw v.12 software.

Antibody production

To obtain the HYLS1-specific peptide antibody, rabbits were immunized with a synthetic peptide (SFDVVSQKFNLPHEYQ) corresponding to amino acid residues 147–161 of the HYLS1 polypeptide. The keyhole limpet hemocyanin (KLH)-coupled peptide was purchased from ResGen. Rabbits were immunized by intradermal injection with 500 µg of peptide-KLH conjugate emulsified in complete Freund's adjuvant. Injections were performed at 3 weeks intervals for a total of 9 weeks. The blood was collected 1 week after the last booster and the serum was isolated by centrifugation.

The lysosome/late endosome-specific lamp1 antibody H4A3, developed by Thomas August (Johns Hopkins University, Baltimore, MD, USA) under auspices of the NICHD, was obtained from Developmental Studies Hybridoma Bank and maintained by the Department of Biological Sciences at the University of Iowa. The microtubule-binding peripheral Golgi membrane antibody 58K was obtained from Sigma and the protein-disulfide isomerase ER marker (PDI) was obtained from StressGen Biotechnologies Corp. (SPA-891). Fluorescein isothiocyanate (FITC)-conjugated anti-rabbit IgG and tetramethylrhodamine isothiocyanate (TRITC)-conjugated anti-mouse IgG were obtained from Sigma.

Construction of expression plasmids

As the entire translated region of the *HYLS1* gene is covered by exon 6, this region was PCR amplified from genomic DNA from one control and one affected individual using primers 5'-GAATTCGAAGCAATGGAAGAAGACTTCTAC and 5'-TCTAGATTAAGAAGGAGAAAGAGGGAAG (TD66-60). FLAG and Myc tags were added to both WT and mutant HYLS1 constructs by using linker primers introducing either the FLAG or the Myc sequence immediately before the original start codon. Each insert was initially cloned into a pGEM-T Easy Vector (Promega; A1360) and subcloned as *EcoRI*-*XbaI*

fragments into either a pCMV5 vector (Anderson 89) or a pGEM-3Z vector (Promega; P2151). Constructs were confirmed by sequencing.

Cell culture, transfections and immunofluorescence microscopy

COS-1 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) (Cellgro) supplemented with 10% fetal bovine serum (FBS) (Cellgro) and antibiotics. One day before transfection, cells were seeded on 3 cm six-well plates at a density of 2.0×10^5 cells/well. Transfection was performed with 1.5 μ g of each of the constructs described earlier in pCMV5 vector using LipofectAMINE PLUS reagent (Invitrogen) following manufacturer's guidelines. Forty-eight hours after transfection, cells were incubated in DMEM for 2 h in the presence of 50 μ g/ml cycloheximide (Sigma), fixed with methanol and blocked with 0.5% bovine serum albumin (BSA) (Sigma)/0.2% saponin (Sigma) in phosphate-buffered saline (PBS). Cells were labeled with the rabbit anti-HYLS1 specific antibody (1:200), mouse anti-FLAG (1:500) (Sigma) or mouse anti-Myc (1:500) (Upstate Cell Signaling Solutions) antibodies. Double-labeling was performed with 58K (Golgi) (1:200), PDI (ER) (1:100) or H4A3 (lysosome) (1:50) antibodies. FITC-conjugated anti-rabbit IgG and TRITC-conjugated anti-mouse IgG were used as secondary antibodies. Cells were mounted on either glycerol or Vectashield mounting medium with DAPI (Vector Laboratories, Inc.; H-1200) and viewed with an inverted Leica TCS-SP confocal microscope (Heidelberg, Germany) using Leica Confocal Software v.2.

Transgenic HYLS1 *D. melanogaster* lines

cDNAs encoding WT and mutant HYLS1 were subcloned into the GAL4 responsive *D. melanogaster* transformation vector pExp-UAS (Exelixis). Germline transformation was carried out by standard methods. Transformants of four WT and five mutant lines were obtained and homozygous lines were established. GMR-GAL4 driver was used to express HYLS1 in the eye discs.

Immunohistochemistry

Third instar eye imaginal discs were dissected and fixed as described previously (21). For immunostaining, whole mount eye discs were incubated with rabbit polyclonal anti-HYLS1 (1:100) and rat monoclonal anti-elav (1:50) at 4°C overnight. FITC-conjugated goat anti-rabbit (1:100) and TRITC-conjugated goat anti-rat (1:100) secondary antibodies were incubated for 3 h at room temperature. Samples were mounted on coverslips using Fluoromount mounting medium (Vectashield). Samples were viewed on a Leica confocal laser scanning microscope.

In vitro transcription and translation

Transcription-coupled translation of both WT and D211G mutant HYLS constructs (in pGEM-3Z) was performed in

the absence of microsomes using the T_NT@T7 coupled reticulocyte lysate system (Promega), according to the manufacturer's protocol. Protein samples were separated on a 10% SDS-PAGE gel followed by autoradiography.

Western blotting

COS-1 cells were transfected with WT or D211G mutant HYLS1 constructs as described earlier, and tissue culture samples for SDS-PAGE electrophoresis were collected 48 h post-transfection. Proteins were separated on 10% SDS-PAGE and analyzed by western blotting with the rabbit anti-HYLS1 primary antibody (1:1000). For fly protein preparation, five adult fly heads of each genotype (WT HYLS1 and y^1w^{1118}) were collected into a glass homogenizer and were homogenized in 1× Laemmli buffer. The homogenates were spun for 10 min at 16060g, and the supernatants were removed. Proteins were separated on 12% SDS-PAGE and analyzed by western blotting with the rabbit anti-HYLS1 primary antibody (1:200).

Pulse-chase analysis

Metabolic labeling of COS-1 cells transfected with WT and D211G mutant constructs was performed as described previously (22). Cells were harvested at 0, 2, 4, 8, 12 and 16 h. Cell lysates were immunoprecipitated with rabbit anti-HYLS1 antibody (1:200).

Bioinformatic analysis

PSORT II (psort.ims.u-tokyo.ac.jp), a program which applies algorithms for protein sorting signals, was used to analyze the potential subcellular localization of both the WT and the D211G mutant forms of the HYLS1 polypeptide. Topology prediction servers, SOSUI (sosui.proteome.bio.tuat.ac.jp/sosui/frame0.html) and Tmpred (www.ch.emblnet.org/software/TMPRED_form.html) were used to determine the hydropathy profile for both the WT and the mutant HYLS1 polypeptide. SignalP v1.1 (www.cbs.dtu.dk/services/SignalP), a signal peptide prediction server, was used to predict NLS in the WT and the mutant HYLS1 polypeptide. PROSITE (us.expasy.org/prosite) and Pfam (www.sanger.ac.uk/Software/Pfam) were used to identify potential protein families or motifs related to the HYLS1 polypeptide. TIGR (www.tigr.org) and NCBI BLAST (www.ncbi.nlm.nih.gov/BLAST) were used to identify cDNA and protein orthologs. ClustalW (www.ebi.ac.uk/clustalw) was used to align all HYLS protein orthologs, and GeneDoc (www.psc.edu/biomed/genedoc) was used to shade the alignment. PROSCAN v. 1.7 (bimas.dcrn.nih.gov/molbio/proscan) was used to predict the most likely promoter regions for HYLS1, and ExPASy's Compute pI/Mw tool (us.expasy.org/tools/pi_tool.html) was used to calculate the pI values for both the WT and the mutant form of HYLS1. Study of the cleavage points of the polypeptide for various enzymes was done with ExPASy's PeptideCutter tool (au.expasy.org/tools/peptidecutter).

ACKNOWLEDGEMENTS

We thank the patients and families for contributing to this work and Dr Harley Kornblum and Dr Daniel Geschwind for sharing their expertise in embryonic and neurological development. Dr Kirsi Sainio and Satu Kuure are thanked for their help with *in situ* analyses and Dr Ismo Ulmanen and George Lawless for their help with the manuscript. Ritva Timonen, Tuula Manninen, Lea Puhakka, Arja Korhonen, Amir Abu-Khalil and Jenny Lee are thanked for their excellent technical assistance. This study was financially supported by the NIH program project grant, project 1P01 ES 11253-02 to L.P. and by the Academy of Finland, project 211124 to M.K. and 202887 to L.P.

Conflict of Interest statement. None declared.

ELECTRONIC DATABASE INFORMATION

Accession numbers and URLs for data presented herein are as follows:

Celera Public Database, <http://public.celera.com/cds/login.cfm> [for BAC clones (accession nos AP003069, AP000708, AP001494, AP001132, AP003087, AP000842); for *HYLS1* genomic sequence (accession no. hCG1644899); *HYLS1* cDNA (accession no. hCT1645026); *HYLS1* protein (accession no. hCP1634375); mouse *Hyls1* genomic sequence (accession no. mCG126066); mouse *Hyls1* cDNA (accession no. mCT127332); mouse *Hyls1* protein sequence (accession no. mCP56034)]
 ExPASy's pI computing tool, http://us.expasy.org/tools/pi_tool.html
 ExPASy's PeptideCutter tool, <http://au.expasy.org/tools/peptidecutter>
 GenBank, <http://www.ncbi.nlm.nih.gov/GenBank> [for *FLJ32915* human *HYLS1* cDNA (GenBank accession no. NM_145014); for mouse *Hyls1* cDNA (GenBank accession no. AK013896); for porcine, chicken, xenopus and zebrafish EST orthologs (GenBank accession nos BP435891, CR354360, BJ623896 and BM034949, respectively); for human, mouse and *C. elegans* protein orthologs (GenBank accession nos NP_659451.1, BAB29044 and AAK95869, respectively)].
 GeneDoc, <http://www.psc.edu/biomed/genedoc>
 Genome Database, <http://www.gdb.org> [for D11S990 (accession ID 195039); for D11S975 (accession ID 191712)]
 ClustalW, <http://www.ebi.ac.uk/clustalw>
 NCBI AceView, <http://www.ncbi.nlm.nih.gov/IEB/Research/AceView/index.html>
 NCBI BLAST, <http://www.ncbi.nlm.nih.gov/BLAST>
 Online Mendelian Inheritance in Man (OMIM), <http://www.ncbi.nlm.nih.gov/Omim> (for hydrolethalus [MIM 236680])
 Pfam (Protein families database of alignments and HMMs), <http://www.sanger.ac.uk/Software/Pfam>
 PROSCAN, <http://bimas.dcert.nih.gov/molbio/proscan>
 PROSITE, <http://us.expasy.org/prosite>
 PSORTII, <http://psort.ims.u-tokyo.ac.jp>

SignalP, <http://www.cbs.dtu.dk/services/SignalP>
 SOSUI, <http://sosui.proteome.bio.tuat.ac.jp/sosui/frame0.html>
 The Institute for Genomic Research, <http://www.tigr.org> [for porcine, chicken, xenopus and zebrafish EST orthologs (BP435891, CR354360, BJ623896 and BM034949, respectively); for human, mouse and *C. elegans* protein orthologs (NP_659451.1, BAB29044 and AAK95869, respectively)].
 TMPRED, http://www.ch.embnet.org/software/TMPRED_form.html
 UCSC Genome Bioinformatics, <http://genome.cse.ucsc.edu>

REFERENCES

- Salonen, R., Herva, R. and Norio, R. (1981) The hydrolethalus syndrome: delineation of a "new", lethal malformation syndrome based on 28 patients. *Clin. Genet.*, **19**, 321–330.
- Siffring, P.A., Forrest, T.S. and Frick, M.P. (1991) Sonographic detection of hydrolethalus syndrome. *J. Clin. Ultrasound*, **19**, 43–47.
- Ämmälä, P. and Salonen, R. (1995) First-trimester diagnosis of hydrolethalus syndrome. *Ultrasound Obstet. Gynecol.*, **5**, 60–62.
- Salonen, R. and Herva, R. (1990) Hydrolethalus syndrome. *J. Med. Genet.*, **27**, 756–759.
- Norio, R. (2003) The Finnish disease heritage III: the individual diseases. *Hum. Genet.*, **112**, 470–526.
- Visapää, I., Salonen, R., Varilo, T., Paavola, P. and Peltonen, L. (1999) Assignment of the locus for hydrolethalus syndrome to a highly restricted region on 11q23–25. *Am. J. Hum. Genet.*, **65**, 1086–1095.
- Krassikoff, N., Konick, L. and Gilbert, E.F. (1987) The hydrolethalus syndrome. *Birth Defects Orig. Artic. Ser.*, **23**, 411–419.
- Chan, B.C., Shek, T.W. and Lee, C.P. (2004) First-trimester diagnosis of hydrolethalus syndrome in a Chinese family. *Prenat. Diagn.*, **24**, 587–590.
- Rakheja, D., Cimo, M.L., Ramus, R.M., Rogers, B.B., Bennett, M.J., Boyer, P.J. and Galindo, R.L. (2004) Hydrolethalus syndrome, in contrast to Smith–Lemli–Opitz syndrome, is not due to a defect in post-squalene cholesterol biosynthesis: a case report. *Am. J. Med. Genet.*, **129A**, 212–213.
- Aughton, D.J. and Cassidy, S.B. (1987) Hydrolethalus syndrome: report of an apparent mild case, literature review, and differential diagnosis. *Am. J. Med. Genet.*, **27**, 935–942.
- de Ravel, T.J., van der Griendt, M.C., Evan, P. and Wright, C.A. (1999) Hydrolethalus syndrome in a non-Finnish family: confirmation of the entity and early prenatal diagnosis. *Prenat. Diagn.*, **19**, 279–281.
- Wan, H. and Wootton, J.C. (2000) A global compositional complexity measure for biological sequences: AT-rich and GC-rich genomes encode less complex proteins. *Comput. Chem.*, **24**, 71–94.
- Peltonen, L., Jalanko, A. and Varilo, T. (1999) Molecular genetics of the Finnish disease heritage. *Hum. Mol. Genet.*, **8**, 1913–1923.
- Fukada, T., Ono, M., Sakata, S., Kioka, N., Sakai, H. and Komano, T. (2003) Genomic structure and 5'-flanking sequences of rat *N*-acetylglucosaminyltransferase I gene and regulatory role of its transcriptional diversity. *Biosci. Biotechnol. Biochem.*, **67**, 1515–1521.
- Ji, H., Zhang, Y., Zheng, W., Wu, Z., Lee, S. and Sandberg, K. (2004) Translational regulation of angiotensin type 1a receptor expression and signaling by upstream AUGs in the 5' leader sequence. *J. Biol. Chem.*, **279**, 45322–45328.
- Muqit, M.M. and Feany, M.B. (2002) Modelling neurodegenerative diseases in *Drosophila*: a fruitful approach? *Nat. Rev. Neurosci.*, **3**, 237–243.
- Lei, E.P. and Silver, P.A. (2002) Protein and RNA export from the nucleus. *Dev. Cell*, **2**, 261–272.
- Pielberg, G., Olsson, C., Syvanen, A.C. and Andersson, L. (2002) Unexpectedly high allelic diversity at the KIT locus causing dominant white color in the domestic pig. *Genetics*, **160**, 305–311.
- Kornblum, H.I., Chugani, H.T., Tatsukawa, K. and Gall, C.M. (1994) Cerebral hemidecortication alters expression of transforming growth factor alpha mRNA in the neostriatum of developing rats. *Brain Res. Mol. Brain Res.*, **21**, 107–114.

20. Breitschopf, H., Suchanek, G., Gould, R.M., Colman, D.R. and Lassmann, H. (1992) *In situ* hybridization with digoxigenin-labeled probes: sensitive and reliable detection method applied to myelinating rat brain. *Acta Neuropathol. (Berl)*, **84**, 581–587.
21. Van Vactor, D.L., Jr, Cagan, R.L., Kramer, H. and Zipursky, S.L. (1991) Induction in the developing compound eye of *Drosophila*: multiple mechanisms restrict R7 induction to a single retinal precursor cell. *Cell*, **67**, 1145–1155.
22. Visapää, I., Fellman, V., Vesa, J., Dasvarma, A., Hutton, J.L., Kumar, V., Payne, G.S., Makarow, M., Van Coster, R., Taylor, R.W. *et al.* (2002) GRACILE syndrome, a lethal metabolic disorder with iron overload, is caused by a point mutation in BCS1L. *Am. J. Hum. Genet.*, **71**, 863–876.
23. Terwilliger, J.D. (1995) A powerful likelihood method for the analysis of linkage disequilibrium between trait loci and one or more polymorphic marker loci. *Am. J. Hum. Genet.*, **56**, 777–787.