Hum Genet (2002) 111 : 350–359
DOI 10.1007/s00439-002-0793-4

ORIGINAL INVESTIGATION

**Michael Linnebank · Eva Tschiedel · Johannes Häberle
Anja Linnebank · Holger Willenbring · Wim J. Kleijer
Hans G. Koch**

Argininosuccinate lyase (ASL) deficiency: mutation analysis in 27 patients and a completed structure of the human ASL gene

Received: 21 March 2002 / Accepted: 18 June 2002 / Published online: 14 August 2002

Abstract Argininosuccinic aciduria is an urea cycle disorder caused by argininosuccinate lyase (ASL) deficiency and is inherited as an autosomal-recessive trait. To date, mutation analysis has been limited because of incomplete sequence information about the human ASL gene. As a consequence, only 12 different mutations in 12 patients have been reported, so far. This study aimed at the completion of the structure and the sequence of the human ASL gene, the development of a genomic DNA-based system for mutation analysis and, finally, the characterisation of the molecular genetic background of ASL deficiency in 27 unrelated patients. This report provides transcript variants, the complete sequence of the human ASL gene and a complete ASL homologue on chromosome 22. The homologue was formerly thought to be a pseudogene but was found, in this study, to be correlated with an immunoglobulin-lambda-like mRNA. On the basis of the novel sequence data, a polymerase reaction chain system for mutation-screening in all 16 coding exons of the ASL gene was established and applied to the analysis of the ASL-deficient patients. We found mutations in all of the 54 investigated alleles and identified 23 (19 novel) different mutations. Some mutational hot-spots were identified (mainly in exons 4, 5, and 7) as were several predominant mutations: IVS5+1G→A (15 alleles), c.532G→A (7), c.346C→T (6), c.1153C→T (4). This study introduces a system for mutation analysis in the ASL gene, thereby elucidating the genetic background of ASL deficiency, which was found to be associated with considerable allelic heterogeneity.

Electronic database information: accession numbers and URLs for the data in this article are as follows:

Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for MIM 207900)

GenBank, http://www.ncbi.nlm.nih.gov/irx/cgi-bin/ (for AC068533; AF376770; NM_000048; NT_001454.20; gi 38322; gi 292400; gi 5103010)

Motif, http://www.motif.genome.ad.jp/ (for sequence analyses)

genscan, http://genes.mit.edu/GENSCAN.html (for sequence analyses)

BLAST, http://www.ncbi.nlm.nih.gov/BLAST (for sequence analyses)

Nucleotide sequence data reported are available in the GenBank database under accession no. AF376770

M. Linnebank · E. Tschiedel · J. Häberle · A. Linnebank
H. Willenbring · H.G. Koch (✉)
Universitätsklinikum Münster,
Klinik und Poliklinik fuer Kinderheilkunde,
48149 Münster, Germany
e-mail: kochha@uni-muenster.de,
Fax: +49-251-8356085

W.J. Kleijer
Clinical Genetics, Erasmus Medical Centre,
Rotterdam, The Netherlands

Present address:
M. Linnebank · A. Linnebank
Universitätsklinikum Bonn, Klinik und Poliklinik für Neurologie,
Bonn, Germany

Introduction

The urea cycle-disorder argininosuccinic aciduria is caused by deficiency of argininosuccinate lyase (ASL: EC 4.3.2.1; ASL deficiency: ASLD, MIM 207900). ASLD, which was first recorded in the late 1950s (Allan et al. 1958), is transmitted as an autosomal-recessive trait and occurs with an incidence of approximately 1:70,000 live births (Levy et al. 1980). It leads to hyperammonaemia, accumulation of argininosuccinic acid in body fluids, and a depletion of arginine. The clinical course of ASLD is of marked heterogeneity. The onset ranges from neonatal (in the majority of cases with severe symptoms, such as life-threatening hyperammonaemic crises, seizures, hypothermia, hyperventilation, vomiting and lethargy) to late (mainly with psychomotor retardation and mental disorders). Trichorrhexis nodosa may occur, but primarily in severe cases. Therapy consists of protein restriction, supplementation of arginine and stimulation of alternative routes of nitrogen detoxification (Brusilow and Horwich 2001; Brusilow and Maestri 1996; Burton 1998; Wu 1991). The clinical outcome is unsatisfactory in many cases, although asymptomatic individuals, diagnosed in routine urine tests, have

been reported (Applegarth et al. 1975; Brusilow and Horwich 2001).

The cleavage of argininosuccinate to fumarate and arginine is an essential step in the process of detoxification of ammonia via the urea cycle. This reaction is catalysed by ASL, a cytosolic homotetramer, which is predominantly expressed in the liver. In other tissues, ASL is involved in the conversion of citrulline to arginine. The ASL subunits consist of 463 amino acids with a predicted molecular weight of about 52 kDa (O'Brien et al. 1986). ASL belongs to a superfamily of enzymes that have homologous parts and catalyse similar cleavages (Yu and Howell 2000).

An ASL-cDNA clone containing an open reading frame of 464 amino acids has been isolated from a human liver cDNA library (O'Brien et al. 1986). The human ASL gene has been mapped to chromosome 7cen→q11.2 (Naylor et al. 1978; Todd et al. 1989) and is reported to consist of 16 exons. Of the genomic 5'-untranslated region, 252 bp have been published, as have up to 22 bp of exon-flanking intronic sequences (Abramson et al. 1991). One region on human chromosome 22 cross-hybridising to 5'-ASL-cDNA is thought to be an ASL pseudogene (O'Brien et al. 1986; Todd et al. 1989).

Gene sharing of ASL and delta-crystallin has been demonstrated for birds and reptiles. Delta2-crystallin of duck and chicken have significant ASL enzyme activity attributable to an evolutionary recruitment of the enzyme ASL to the lens in birds (and, similarly, in reptiles), where it functions as a structural protein via gene duplication and subsequent specialisation (Barbosa et al. 1991b; Piatigorsky et al. 1988). The human ASL enzyme shares up to 71% homology with delta-crystallins of various species (Matsubasa et al. 1989; Piatigorsky et al. 1988).

Table 1 Overview of all ASLD alleles

Exon	DNA	Kind of mutation[a]	Peptide[b]	CR[c]	Known alleles (*n*=22)	Alleles in this paper (*n*=54)	All alleles (*n*=76)	Ref.[d]
2	c.35G→A	ms	R12Q	+	1	–	1	6
2	c.175G→A	ms	E59 K	+	–	2	2	–
3	c.257A→C	ms	E86A	+	1	–	1	4
3	c.260A→G	ms	D87G	+	3	1	4	3
3	c.283C→T	ms	R95C	+	2	–	2	1
4	c.G292del[e]	del/ss	fs/del19aa	+	–	1	1	–
4	c.299T→C	ms	I100T	+	–	1	1	–
4	c.331C→T	ms	R111 W	+	2	–	2	2
4	c.337C→T	ms	R113 W	+	1	–	1	4
4	c.346C→T[f]	ns/ss	Q116X	+	–	6	6	–
5	c.392C→T	ms	T131 M	+	–	1	1	–
IVS5	IVS5+1G→A[f]	ss	decay/del21aa	+	–	15	15	–
6	c.461T→C	ms	L154P	+	–	1	–	
6	c.476C→A	ms	T159 N	+	–	1	1	–
7	c.532G→A	ms	V178 M	+	–	7	7	–
7	c.545G→A	ms	R182Q	+	–	1	1	–
7	c.578G→A	ms	R193Q	+	1	1	2	2
7	c.580ins6bp[g]	ins	193insK,R	+	–	1	1	–
IVS9	IVS9+5G→A[f]	ss	decay/del21aa	+	–	2	2	–
10	c.762C→A	ms	S254R	+	–	1	1	–
11	c.857A→G	ms	Q286R	+	5	1	6	2; 3
11	c.889C→T	ms	R297 W	+	–	1	1	–
12	c.925G→A	ms	G309R	(+)	–	2	2	–
13	c.1044del13[h]	del / ss[d]	del28aa	+	2	–	2	2; 3
14	c.1121insC	ins/ fs	fs	+	2	–	2	5
14	c.1135C→T[i]	ms	R379C	+	–	2[i]	2[i]	–
15	c.1153C→T	ms	R385C	+	–	4	4	–
15	c.1193C→A	ms	A398D	+	1	1	2	3
16	c.1275C→T[i]	silent	I425I	–	–	2[i]	2[i]	–
16	c.1331C→T	ms	A444 V	+	–	1	1	–
16	c.1360C→T	ns	Q454X	+	1	–	1	2

[a]*del* deletion, *fs* frameshift, *ins* insertion, *ms* missense, *ns* nonsense, *ss* splice site
[b]*aa* Amino acid(s), *del* deletion (of)
[c]*CR* Conserved region according to Barbosa et al. (1991a)
[d]*Ref*. References: *1* Walker et al. 1990; *2* Barbosa et al. 1991a; *3* Walker et al. 1997; *4* Linnebank et al. 2000; *5* Stadler et al. 2001; *6* Sampaleanu et al. 2001
[e]Splice acceptor site affected; see Fig. 2B
[f]Splice donor site affected; see Fig. 2A
[g]Tandem repeat of 6 bp (AGCGGA) on transcripts and on genomic DNA
[h]Leads to deletion of exon 13 from transcripts
[i]Double mutated allele (1135C→T and 1275C→T)

Until now, 12 ASLD patients have been characterized at the molecular genetic level; 22 alleles have been identified carrying 12 different mutations (Table 1). The feature of interallelic complementation is known for some oligomeric enzymes (e.g. Gravel et al. 1977, 1994; Qureshi et al. 1994). In cell culture experiments, interallelic complementation extensively occurs between distinct ASLD subunits but any clinical relevance is hypothetical (Howell et al. 1998; McInnes et al. 1984; Sampaleanu et al. 2001; Yu et al. 2001). Genetic analysis of ASL transcripts has revealed excessive skipping of various exons, whichs appears in all investigated cells and tissues of ASLD patients and of controls: this mainly involves exons 2 and 7 but other "in-frame" exons are also frequently skipped (Abramson et al. 1991; Linnebank et al. 2000; Walker et al. 1990). The biological impact of these shortened transcripts or possible effects of corresponding shortened peptides on the tetramer remain unknown. If shortened transcripts are translated, the corresponding polypeptides might be misfolded.

Subjects and methods

Nomenclature

Gene mutation nomenclature used in this article follows the recommendations of den Dunnen and Antonarakis (2001). Gene symbols used in this article follow the recommendations of the HUGO Gene Nomenclature Committee (Povey et al. 2001).

Patients

Cultured fibroblasts of ASLD patients from various European countries were sent to a specialized laboratory (W. J. Kleijer) for confirmation of clinically or biochemically suspected ASLD. In this study, 27 patients (10 female, 17 male) were analysed at the molecular genetic level. ASL activity was measured in cultured skin fibroblasts of all patients by ^{14}C-citrulline incorporation into trichloroacetic-acid-precipitable proteins as described (Kleijer et al. 1984, 2002). All patients, except for five variants described below, showed an almost complete absence of ^{14}C-citrulline incorporation, i.e. less than 2% of the level in intra-assay normal controls, suggesting almost complete ASLD. The five patients with a biochemical variant type of ASLD (Table 2) exhibited considerable residual ^{14}C-citrulline incorporation (6%–21% of controls; Kleijer et al. 2002).

Gene analysis

Sequence analyses were performed electronically (GenBank, motif and genscan). BLAST-homology searches were based on the human ASL cDNA (NM_000048) and subsequently on the obtained sequences.

Transcript analysis

The characterization of the various liver ASL transcripts was performed with the 5' rapid amplification of cDNA ends (RACE) technique. The method was as previously described (Homberger et al. 2000) and was essentially according to the manufacturer's instructions (5'-RACE kit, Roche Diagnostics, Mannheim; marathon cDNA, Clontech, Heidelberg). The specific RACE primer was 5'-CAGCTTCCCTGCCGTTGCACCAATGAGCTC.

Table 2 ASLD patients: genetics

Patient	Gender	Origin	Enzyme[a]	Allele 1	Allele 2
831	F	Dutch	<2	c.545G→A	IVS5+1G→A
832	F	Dutch	<2	c.532G→A	IVS5+1G→A
833	M	Italian	<2	c.260A→G	c.1193C→A
834	M	Dutch	<2	c.580ins6[c]	IVS5+1G→A
835	F	Dutch	<2	c.G292del	IVS5+1G→A
836	F	Bulgarian	<2	c.532G→A	IVS5+1G→A
837	M	Swiss	<2	IVS5+1G→A	IVS5+1G→A
839	M	Czech	<2	c.532G→A	IVS5+1G→A
846	M	Dutch	<2	c.476C→A	IVS5+1G→A
847	F	Dutch	<2	IVS5+1G→A	IVS5+1G→A
848	F	Belgian	<2	c.925G→A	c.925G→A
850	M	Turkish	<2	IVS9+5G→A	IVS9+5G→A
851[b]	M	Lebanese	7	c.1153C→T	c.1153C→T
852[b]	M	Italian	6	c.857A→G	c.578G→A
853[b]	M	Finnish	12	c.1153C→T	c.1153C→T
858	M	Dutch	<2	c.762C→A	IVS5+1G→A
859	F	Arab	<2	c.346C→T	c.346C→T
860	F	Turkish	<2	c.889C→T	c.392C→T
861	M	Moroccan	<2	c.175G→A	c.175G→A
862	M	Dutch	<2	c.532G→A	IVS5+1G→A
863[b]	M	Gypsy/Spanish	8	c.532G→A	c.532G→A
883	M	Dutch	<2	c.532G→A	IVS5+1G→A
892[b]	M	Arab	21	c.1135C→T; c.1275C→T	c.1135C→T; c.1275C→T
893	F	Arab	<2	c.346C→T	c.346C→T
894	F	Arab	<2	c.346C→T	c.346C→T
W	M	Swiss	n.a.	c.461T→C	c.1331C→T
R	M	German/Egypt	n.a.	c.299T→C	IVS5+1G→A

[a]Deficient ASL enzyme activity was measured in cultured fibroblasts indirectly by using an ^{14}C-citrulline incorporation assay. Results are given as % relative to intra-assay control (Kleijer et al. 2002)
[b]Variant patients with residual enzyme activity. Cultured fibroblasts were not available for patients W and R
[c]Tandem repeat of 6 bp (AGCGGA) in transcripts and in genomic DNA(*n.a.* not attempted)

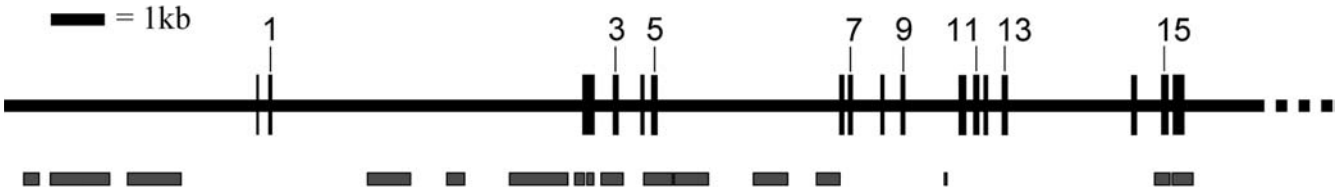

Fig. 1 Structure of the human ASL gene. In this scaled figure, the structure of the entire human ASL gene is schematised. Exons are depicted as *boxes* (*odd numbers* are given; see Table 4). The regions homologous to the "ASL pseudogene" (see Table 5) are shown *below* as *grey boxes*

Polymerase chain reaction system for mutation analysis in genomic DNA

A polymerase chain reaction (PCR) system for mutational screening based on genomic DNA was established in this study. All genomic primers included 5'-attached M13 sequences (TGTAAAACGACGGCCAGT, M13U; and CAGGAAACAGCTATGACC, M13R) as templates for IRD800-labelled cycle sequencing primers for direct sequencing as previously described (Linnebank et al. 2000). Sequencing of the ten PCR-amplified fragments allowed the analysis of the entire coding region and of all flanking intronic regions simultaneously on both alleles. PCR conditions and primer sequences are given below.

Mutation analysis

Cultured fibroblasts were used for extraction of transcripts and of genomic DNA. Alternatively, peripheral lymphocytes were prepared from EDTA-treated blood samples. Aliquots of the EDTA-treated blood were used for isolation of genomic DNA. The analysis of transcripts was possible from cultured fibroblasts and from peripheral lymphocytes, even if blood samples had been shipped for several days without special handling or treatment. The methods of preparation were as described earlier as were the direct sequencing of genomic DNA and the molecular cloning by which ASL transcripts were analysed by full-length reverse transcription/PCR (RT-PCR; Linnebank et al. 2000).

In this study, mutation analysis was carried out on samples from 27 ASLD patients. In each case, RT-PCR was performed and all mutations found in transcripts were verified at the genomic level. Mutations were additionally confirmed in parental DNA, if available.

Results

ASL gene

GenBank searches revealed two overlapping continuous genomic sequences (both included in AC068533), which contained the whole human ASL gene. Therefore, the entire ASL gene sequence could be deduced. The complete and corrected sequence of the human ASL gene is now available at AF376770 (Fig. 1).

Sequence alignments showed that the formerly published gene structure (Abramson et al. 1991) was correct with respect to the exonic sequences, except for only one nucleotide that had formerly been estimated to be the 5'-terminal nucleotide of exon 10 but that turned out to be the 3'-terminal nucleotide of exon 9. Additionally, some minor corrections had to be made to the short intronic sequences published previously (not shown). The size of the ASL gene was found to be 17 kb instead of 35 kb, indicating a prior over-estimation. We corrected the overlapping continuous genomic sequences (AC068533) at one position of the coding region (not shown) prior to deposition under GenBank accession no. AF376770.

Table 3 Exons and intervening sequences (*IVS*) of the human ASL gene. The genomic positions refer to AF376770, whereby the A of the start-codon (position/bp 5001 in AF376770) was set as nucleotide/bp no. 1 (*Cds* coding sequence). The 5' border of exon 1 refers to O'Brien and co-workers (1986), whereas supplementary data are given in this report

Exon	5' border	3' border	Size (bp)	IVS	5' border	3' border	Size (bp)
1	–244	12	256	1	13	5,721	5,709
Cds(1)	1	12	12				
2	5,722	5,916	195	2	5,917	6,286	370
3	6,287	6,370	84	3	6,371	6,798	428
4	6,799	6,855	57	4	6,856	6,995	140
5	6,996	7,093	98	5	7,094	10,502	3,409
6	10,503	10,581	79	6	10,582	10,661	80
7	10,662	10,739	78	7	10,740	11,251	512
8	11,252	11,304	53	8	11,305	11,646	342
9	11,647	11,709	63	9	11,710	12,724	1,015
10	12,725	12,839	115	10	12,840	13,008	169
11	13,009	13,093	85	11	13,094	13,193	100
12	13,194	13,253	60	12	13,254	13,529	276
13	13,530	13,613	84	13	13,614	15,923	2,310
14	15,924	16,004	81	14	16,005	16,474	470
15	16,475	16,581	107	15	16,582	16,685	104
16	16,686	16,868	183				
Cds(16)	16,686	16,830	145				
Poly-A	16,852	16,857	6				

In this report, the A of the ATG-start codon of the ASL gene has been set as nucleotide no. 1 in genomic DNA (g.1) and cDNA (c.1) and the respective exon is depicted as exon 1, in concordance with the established ASL nomenclature.

Regarding the complete ASL gene (AF376770), the sizes of the ASL exons ranged from 53 bp (exon 8) to 195 bp (exon 2) and the sizes of introns from 80 bp (IVS6) to 5.7 kb (IVS1). The positions and sizes of exons and introns are given in Table 3. Transcript analyses (see below) revealed the existence of a non-coding upstream mRNA sequence referred to as exon 0. Sequence analysis with genscan and motif suggested the ASL gene to be under control of a TATA-box (g.–3727 to g.–3713). Moreover, several other putative regulatory domains were identified: CCAAT/enhancer binding proteins (g.–3205 to g.–3192 and g.–674 to g.–662), SP1 (g.–2954 to g.–2945 and g.–319 to g.–310), upstream stimulating factor (g.–2472 to g.–2459), core-promoter (g.–2411 to g.–2372), activator-protein 1 (g.–1022 to g.–1014), activator-protein 4 (g.–413 to g.–404), activating transcription factor (g.–296 and g.–283) and poly-A signal (g.16,852 to g.16,857).

ASL-transcripts

The results of the RACE-experiments revealed one possible major transcript, which started at c.–45 (exon 1). This transcript is shorter than that previously published (O'Brien et al. 1986), which was not detected in this study. Further transcripts carried an additional upstream sequence (exon 0) followed by exon 1 from c.–45 on. In addition to these two, inserts were found with a transcription initiation at c.–30 (exon 1) and within exon 2 (c.58). Concerning the latter, no in-frame ATG in agreement with the requirements for start codons existed in the adjacent downstream sequence, suggesting that this transcript might be an artefact. The other three mentioned transcripts, beginning with exon 0 or exon 1 (c.–45 or c.–30), most probably make use of the same start codon, namely the first in-frame ATG. In accordance with these findings, the mentioned ATG had previously been published as the assumed start codon, and the A was numbered as c.1/g.1, which could be kept in this study (O'Brien et al. 1986).

ASL pseudogene

A BLAST search at GenBank with the entire ASL gene revealed the first complete sequence of the ASL pseudogene, which was found to map to chromosome 22q11.2 (NT_001454.20; subclone KB113H7). Spread over 88.4 kb, the pseudogene included 16 continuous sequences with a length between 53 bp and 1175 bp and contained homologies to the ASL gene between 83% and 92% (Fig. 1; Table 4). Homologous parts of the ASL gene were found in exonic (exons 2–5, 15, 16) and intronic sequences (IVS 1–5, 9, 14, 15) and in the 5'- and 3'-untranslated regions including the putative promoter domain and the putative poly-A signal. Altogether, 8.9 kb genomic ASL sequence were shown to exhibit an average of homology of 88% with the ASL pseudogene.

Genscan analyses suggested the so-called pseudogene to be a regular gene with a promoter region, a poly-A signal and 11 exons containing a typical initial exon and a terminal exon. The predicted coding sequence of the pseudogene contained 1239 bp. More than 0.4 kb shared high homology with the human ASL cDNA (data not shown). A GenBank search with the predicted cDNA pointed out that this so-called pseudogene might be a regular gene, coding for immunoglobulin-lambda-like mRNA (gi 38322, 292400, 38323 and 5103010).

Table 4 Alignment of the ASL gene (AF376770) and its so-called pseudogene (NT_001454; see also Fig. 1; *UTR* untranslated region, *prom* promoter, *IVS* intervening sequence, *ex* exon). Each line describes a homologous part of the ASL gene and its pseudogene; the position of this part is given for both genes. The lengths of homologous sequences refer to the ASL gene, not to the pseudogene

5' position ASL	3' position ASL	Domain/region in the ASL gene	5' position pseudo-gene	3' position pseudo-gene	Homology: length in bp (total: 8.9 kb)	Homology in % (88% in average)
–4,733	–4,432	5'-UTR	3,466,894	3,466.603	302	87
–4,271	–3,097	5'-UTR	3,466,613	3,465,440	1175	90
–2,798	–1,736	5'-UTR including prom	3,456,449	3,464,377	1063	89
1,673	2,530	IVS1	3,386,191	3,385,333	858	89
3,135	3,493	IVS1	3,385,077	3,384,725	359	88
4,320	5,476	IVS1	3,384,236	3,383,071	1157	87
5,526	5,727	IVS1, ex2	3,383,064	3,382,867	202	90
5,749	5,891	ex2	3,382,871	3,382,731	143	92
6,013	6,449	IVS2, ex3, IVS3	3,382,730	3,382,302	437	89
6,803	7,373	ex4, IVS4, ex5, IVS5	3,382,291	3,381,720	571	87
7,390	8,063	IVS5	3,381,713	3,381,070	674	83
8,849	9,528	IVS5	3,380,584	3,379,881	680	86
10,031	10,505	IVS5	3,379,697	3,379,224	475	90
12,427	12,479	IVS9	3,380,371	3,380,311	53	90
16,298	16,613	IVS14, ex15, IVS15	3,379,227	3,378,914	316	87
16,645	17,059	IVS15, ex16, 3'-UTR	3,378,912	3,378,498	415	91

Table 5 Oligonucleotide primers for PCR and direct sequencing (*s* sense, *as* antisense)

Exon	s/as	Position g.	Sequence[a]	Product (bp)[b]
1	s	(–249) – (–230)	ACACTATCCGTGCGGCCAGG	349
	as	63 – 44	ACTCTCTCCTTTGGAGGCTA	
2	s	5,621 – 5,640	CCTGCTACCATCAGACTTGA	420
	as	6,004 – 5,985	CAACACTGCACTGTTTGCTC	
3	s	6,231 – 6,250	CTCTTGGCTGCTGATGCCTG	247
	as	6,441 – 6,422	AATGCTCAAAGTGGCCCTGG	
4+5	s	6,732 – 6,751	ACTGCCCTGCCTGGGTTGAC	448
	as	7,143 – 7,124	GGCTTCCATCACACCTCTGT	
6+7	s	10,445 – 10,464	GTCACAGGCAGGCCTTGCAT	390
	as	10,798 – 10,779	CCTAAGGCTATGCAGCAGCC	
8+9	s	11,201 – 11,219	TCAGGGCTGCCTGCCAGGA	595
	as	11,759 – 11,740	GAAGGACTAAAGTGAGGCAC	
10–12	s	12,655 – 12,674	TCCTGCCCCCTGTATGGTCA	680
	as	13,298 – 13,278	CCTGAGAACTGGCACGTACTA	
13	s	13,495 – 13,514	TGCTAGGCCCTCACCTCCTG	246
	as	13,704 – 13,685	CCACCTCACCTCCTCCCAAG	
14	s	15,869 – 15,888	CAGCTTCAGATCCCAGGGTC	309
	as	16,141 – 16,122	CACCTCGGCTTCCCAAAGTG	
15+16	s	16,425 – 16,444	CAGGGCATGGAGAAACCTGC	572
	as	16,960 – 16,941	CCTCTCCAGTCCCTGACTGT	

[a]All primers contained 5'-attached 18 bp of M13 tails (sense or antisense) for direct sequencing as mentioned in the text

[b]Size in bp including 36 bp of the M13 tails

PCR system of genomic screening for ASL mutations

The novel sequence information was used to generate genomic DNA-based oligonucleotide primers for the amplification of single or adjacent exons of the ALS gene including the respective splice sites. Cross-hybridisation with the homologous sequence on chromosome 22 was avoided, which was established by the sequencing of all PCR products. The 16 coding exons of the ASL gene were amplified by 10 PCRs. The PCR design allowed the simultaneous performance of all PCRs in a two-step procedure with an annealing/synthesis temperature of 68°C. Primer sequences are outlined in Table 5.

Mutations

Defects on both ASL-alleles were identified for each patient (Table 2). Different typically shortened transcripts attributable to exon-skipping were detected in the RT-PCR products of all patients. This was in accordance with previously published observations in patient and control cell-lines (Linnebank et al. 2000). Missense mutations (Table 2) did not have any obvious influence on the distribution or quantity of shortened transcripts. However, in some patients, RT-PCR products were undetectable or showed a changed pattern of shortened transcripts (not shown). These patients were found to be affected by splice-site or frameshift mutations in subsequent genomic analysis.

In addition to the previously known 21 ASL deficient alleles harbouring 12 different mutations, 54 mutant alleles were found in this study (Table 1). They comprised 24 different mutations, including 19 novel ones: 17 missense mutations (in total 29 alleles), one nonsense mutation (6 alleles), two splice-site mutations (17 alleles), and one silent mutation (2 alleles) plus one genomic deletion (1 allele) and one genomic insertion (1 allele). All missense mutations affected conserved regions of the ASL peptide according to BLAST analyses and to Barbosa and co-workers (1991a). Of the 27 patients, 12 were found to carry mutations in a homozygous state. For all patient lines, sequence alterations other than the diagnosed mutations in the coding region were excluded (a doubly mutated allele was found only in one case, see below).

Mutations of splice sites

The novel, but obviously frequent, splice donor site mutation IVS5+1G→A was associated with low product amounts in RT-PCR. The transcripts of this allele were found to be affected by the deletion of exon 5 instead of an insertion of IVS5 (which we rather had expected, since the mutation IVS5+1G→A affected the splice donor site). Similarly, the other two splice donor site mutations (c.346C→T, IVS9+5G→A) of this study were also found to cause deletions of the respective upstream exon (Fig. 2A).

Mutation c.G292del in the splice acceptor site of exon 4 (patient 835; Fig. 2B) was correlated with two groups of transcripts, since this splice site mutation was found not to lead to altered splicing in each transcript: Missplicing led to the in-frame deletion of exon 4, whereas obviously, the presence of correct splicing conserved exon 4, which however was affected by the deletion of c.G292. There-

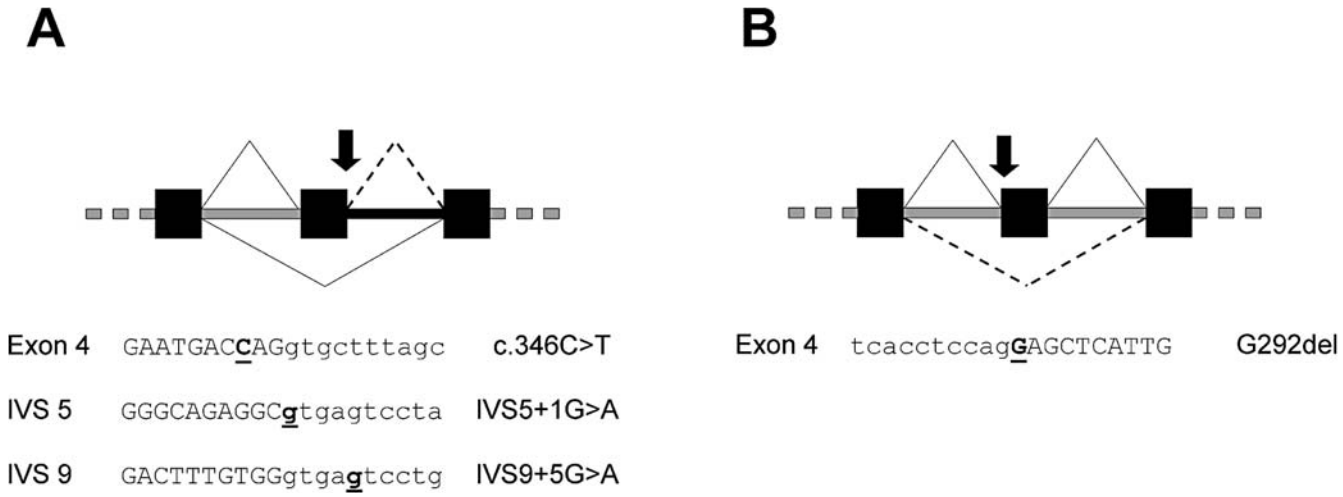

Fig. 2A, B Splice site mutations of this study. Scheme for the effects of the various splice donor (**A**) and acceptor (**B**) sites detected in this study. **A** In general, splice donor site mutations are expected to disturb normal splicing (*fine dotted lines*). On respective transcripts, an insertion of the downstream IVS (*black rectangle*) may occur (*boxes* exons). The splice donor site mutations (*arrow*) found in this study (c.346C→T; IVS5+1G→A; IVS9+5G→A) led to low amounts of RT-PCR products lacking any insertion. Instead, the 5' exon was deleted (*middle box*). **B** Delineation of the splice acceptor site mutation G292del. This deletion of a single basepair led to two different transcripts: in the one (which seemed to be better amplifiable by RT-PCR, data not shown), altered splicing (*lower fine dotted lines lines*) led to the skipping of exon 4, whereas the other was normally spliced (*upper lines*) and showed the deletion of c.G292. Therefore, transcripts resulting from c.G292del alternatively are affected by a frameshift or by the deletion of exon 4

fore, the mutation c.G292del either resulted in the deletion of exon 4 or in a frameshift.

Silent mutations

Patient 892 carried the silent mutation c.1275C→T (I425I), together with the missense transition c.1135C→T (R379C) in *cis* in a homozygous state. This alteration was neither detected in other patients nor in controls. No further silent mutations were found in this study.

Distribution of mutations

Several ASLD-mutations were found to be more common (Table 1): IVS51+G→A (15 of 54 alleles), c.532G→A (7 alleles), c.346C→T (6 alleles) and c.1153C→T (4 alleles). Exons 3–5, 7 and 11 contained the most mutations (Table 1) Homozygosity was diagnosed for 12 of the 27 unrelated patients: c.175G→A (1 patient), c.346C→T (3 patients), IVS5+1G→A (2 patients), c.532G→A (1 patient), IVS9+5G→A (1 patient), 925G→A (1 patient), c.1153C→T and c.1275C→T (double mutated allele; 1 patient) and 1153C→T (2 patients).

Discussion

Genomic and transcript data

This study made use of database sequence information, which allowed the computer-based deduction of the complete human ASL gene and a homologous gene and the analysis of putative functional domains. A possible promoter region including a TATA-box homologue was identified. This feature is also present in other genes encoding for urea-cycle or related enzymes: liver-type arginase (Takiguchi et al. 1988), ornithine transcarbamylase (Veres et al. 1986), ornithine delta-aminotransferase (Mitchell et al. 1988). RACE experiments revealed diverse ASL transcripts with various initial exons, which most probably encoded the same ASL peptide.

The phenomenon of gene sharing had been reported for the avian ASL gene. In this study, a homologue of the human ASL gene, the so-called ASL pseudogene, could be characterized and was shown to be most probably a functional gene, encoding an immunoglobulin-lambda-like peptide. However, there are no experimental data, so far, providing evidence that the immunoglobuline-lambda-like gene is biologically active or that it is related to the ASL gene in terms of gene sharing. The detailed information established for the "ASL pseudogene" sequence was used in this study to avoid cross-hybridisation of ASL primers.

PCR screening system

Our protocol for genomic mutation screening has allowed the confirmation of mutations found at the cDNA level and might also be used as a stand-alone tool for mutation screening at the genomic level. This is important, since cells might not always be available for mutation analysis and since some alleles give small amounts of RT-PCR product in this study. Some of the 16 coding exons were amplified together (because some primer pairs spanned two or more physically close-neighbouring exons) and all

PCRs could be performed simultaneously under the same conditions. PCR products contained sense and antisense M13 sequences for convenient direct sequencing. Since the oligonucleotide primers were located intronically and at a distance to the exon-intron-boundaries, the investigation of splice sites was possible; this is especially important in ASLD. The ASL feature of common exon skipping complicates the recognition of splice-site mutations at the transcriptional level but, as indicated by this study, splice-site mutations seem to be a frequent cause of ASLD (Table 1). Differentiation between exon skipping and splice-site mutations as a reason for missing exons at the genomic level appears indispensable.

Some of the ASL mutations seem to be more common than others but the heterogeneity of the ASL mutations has demonstrated that screening of mutational hot-spots will often not be sufficient. The method presented here has proven feasible for genomic screening of the entire coding region of the human ASL gene and might be useful in future studies.

ASLD patients and mutations

In this study, we have investigated the largest ASLD collective reported so far at the molecular genetic level. ASLD is supposed to occur with an equal prevalence in female and male patients and has been observed in several different ethnic populations (Brusilow and Horwich 2001). Here, we report 10 female and 17 male patients of different ancestry (Table 2). Cultured skin fibroblasts of 22 out of the 27 patients exhibit almost undetectable citrulline incorporation, whereas the remaining five patients are biochemical variants with high residual enzyme activity. Details of these variants have been published elsewhere (Kleijer et al. 2002).

Validity of data

The mutations presented in this paper were considered to be disease-causing, since they affected the splicing of RNA, caused frameshifts or at least changed conserved amino acid residues. Mutations were mostly established in parental DNA and at both the transcriptional and the genomic level.

Splice donor site mutations

Three novel splice donor site mutations (IVS5+1G→A, IVS9+5G→A, c.346C→T) were found to lead to the deletion of upstream exons, whereas insertions of respective downstream intervening sequences were not detected in RT-PCR analysis. We suppose that enlarged transcripts with an insertion might occur but might be unstable or at least difficult to amplify by the chosen RT-PCR. Furthermore, the ASL feature of exon skipping might be of enhanced relevance for respective transcripts. If the exon is skipped (splicing as shown by the lower lines in Fig. 2), the splice defect (shown by the upper lines), which would otherwise lead to a large insertion (black rectangle), is of minor relevance. Therefore, transcripts with the skipped exon seem to be advantageous in terms of transcript stability and PCR amplification.

Interallelic complementation

In this study, 12 patients showed homozygous allelic states, whereas the other 15 exhibited compound heterozygous allelic states. Previous studies indicated interallelic complementation of ASLD polypeptides as an important feature for modulating enzyme activity (Howell et al. 1998; Sampaleanu et al. 2001; Walker et al. 1997; Yu et al. 2001). However, in our study, we could not observe any major impact of this phenomenon on the biochemical phenotype of a large series of patients; except for five cases, all patients revealed absent enzyme activity in cultured fibroblasts. Remarkably, of the five patients with residual activity, four were homozygous for individual mutations excluding interallelic complementation (Table 2). Concerning the mentioned reports, interallelic complementation at the ASL locus seems to be restricted to distinct mutations; two missense mutations were found to play a major role, viz. D87G and Q286R. In the present study, we report two patients as having a compound heterozygous state for each of these two mutations: one of them (852) carried the mutations Q286R and R193Q and was indeed a biochemical variant but the other (833) was a typical case without marked residual activity, in spite of previous reports that demonstrated a complementing effect for the same mutations D87G and A398D (Yu et al. 2001). This might be attributable to the application of different assays for the characterization of residual ASL enzyme activity.

The transition c.532G→A was found in five patients in a compound heterozygous allelic status, always together with IVS5+1G→A on the other allele but not with other mutations. We do not know whether the concurrence of those two mutations is of any special relevance. One patient (863) was homozygous for c.532G→A. Remarkably, he was one of the patients who revealed residual enzyme activity, which was not detected in the five compound heterozygous patients affected by c.532G→A together with IVS5+1G→A. Therefore, it seems that the frequently found combination c.532G→A/IVS5+1G→A does not lead to interallelic complementation in relevant amounts. We cannot explain why the patient homozygous for c.532G→A (or the other variant patients with homozygous allelic status) exhibits high residual ASL activity. Interallelic complementation does not seem to be a possible explanation for the biochemical variants, since four of the five variant patients carry mutations in homozygous states.

Obviously, interallelic complementation does not play a major role for modifying the biochemical phenotypes in the population analysed in this study. However, we can neither provide clinical data nor exclude some clinical impact of interallelic complementation, since one could

speculate that "complementers" might be mildly affected without manifestation of clinical symptoms. In conclusion, our large series of ASLD patients who have been investigated by means of enzyme activity and genetics provides no evidence that naturally occurring interallelic complementation plays a major role in modifying the biochemical phenotype.

Ethnic distribution of mutations

Three of the four Arab patients were homozygous for c.346C→T, which was not found in patients of other origins. A founder effect of c.346C→T in the Arab population seems likely or c.346C→T might be more common in ASLD in Arab patients and could therefore be the worthwhile subject of first-line molecular analysis in such patients.

This study provides data concerning the molecular genetic background of ASL deficiency, introduces a system for genetic diagnosis and presents information about the ASL gene and its so-called pseudogene.

Acknowledgements The technical assistance of I. Du Chesne, M. Herting and V. Sablitzky is gratefully acknowledged. We thank Drs. B. Steinmann and J. Herwig for the contribution of samples of two patients. This study was supported by a grant from Innovative Medizinische Forschung (IMF) Münster to H.W. and H.G.K.

References

Abramson RD, Barbosa P, Kalumuck K, O'Brien WE (1991) Characterization of the human argininosuccinate lyase gene and analysis of exon skipping. Genomics 10:126–132

Allan JD, Cusworth DC, Dent CE, Wilson VK (1958) A disease, probably hereditary, characterised by severe mental deficiency and a constant gross abnormality of amino acid metabolism. Lancet I:182–187

Appelgarth D, Davidson A, Perry L, Podon S, Chrichton J, Flandwick D (1975) Argininosuccinic acidemia in a healthy individual detected by a urine screening program. Clin Chem 21:950–951

Barbosa P, Cialkowski M, O'Brien WE (1991a) Analysis of naturally occurring and site-directed mutations in the argininosuccinate lyase gene. J Biol Chem 266:5286–5290

Barbosa P, Wistow GJ, Cialkowski M, Piatigorsky J, O'Brien WE (1991b) Expression of duck lens delta-crystallin cDNAs in yeast and bacterial hosts. Delta 2-crystallin is an active argininosuccinate lyase. J Biol Chem 266: 22319–22322

Brusilow SW, Horwich AL (2001) Urea cycle enzymes. In: Scriver CR, Beaudet AL, Valle D, Sly WS (eds) The metabolic and molecular bases of inherited disease, 8th edn. McGraw-Hill, New York, pp 1909–1963

Brusilow SW, Maestri NE (1996) Urea cycle disorders: diagnosis, pathophysiology, and therapy. Adv Pediatr 43:127–170

Burton BK (1998) Inborn errors of metabolism in infancy: a guide to diagnosis. Pediatrics 102:E69

Dunnen JT den, Antonarakis SE (2001) Nomenclature for the description of human sequence variations. Hum Genet 109:121–124

Gravel RA, Lam KF, Scully KJ, Hsia Y (1977) Genetic complementation of propionyl-CoA carboxylase deficiency in cultured human fibroblasts. Am J Hum Genet 29:378–388

Gravel RA, Akerman BR, Lamhonwah AM, Loyer M, Leon-del-Rio A, Italiano I (1994) Mutations participating in interallelic complementation in propionic acidemia. Am J Hum Genet 55: 51–58

Homberger A, Linnebank M, Winter C, Willenbring H, Marquardt T, Harms E, Koch HG (2000) Genomic structure and transcript variants of the human methylenetetrahydrofolate reductase gene. Eur J Hum Genet 8:725–729

Howell PL, Turner MA, Christodoulou J, Walker DC, Craig HJ, Simard LR, Ploder L, McInnes RR (1998) Intragenic complementation at the argininosuccinate lyase locus: reconstruction of the active site. J Inherit Metab Dis 21 (Suppl 1):72–85

Kleijer WJ, Blom W, Huijmans JG, Mooyman MC, Berger R, Niermeijer MF (1984) Prenatal diagnosis of citrullinemia: elevated levels of citrulline in the amniotic fluid in the three affected pregnancies. Prenat Diagn 4:113–118

Kleijer WJ, Garritsen VH, Linnebank M, Mooyer P, Huijmans JGM, Mustonen A, Simola K, Arslan-Kirchner M, Battini R, Briones P, Cardo E, Mandel H, Tschiedel E, Wanders RJA, Koch HG (2002) Pre- and postnatal diagnosis of variants of argininosuccinic aciduria with partial argininosuccinate lyase deficiency in five unrelated families. J Inherit Metab Dis (in press)

Levy HL, Coulombe JT, Shih VE (1980) Newborn urine screening. In: Bickel H, Guthrie R, Hammersen G (eds) Neonatal screening for inborn errors of metabolism. Springer, Berlin Heidelberg New York, pp 89–103

Linnebank M, Homberger A, Rapp B, Winter C, Marquardt T, Harms E, Koch HG (2000) Two novel mutations (E86A, R113 W) in argininosuccinate lyase deficiency and evidence for highly variable splicing of the human argininosuccinate lyase gene. J Inherit Metab Dis 23:308–312

Matsubasa T, Takiguchi M, Amaya Y, Matsuda I, Mori M (1989) Structure of the rat argininosuccinate lyase gene: close similarity to chicken delta-crystallin genes. Proc Natl Acad Sci USA 86:592–596

McInnes RR, Shih V, Chilton S (1984) Interallelic complementation in an inborn error of metabolism: genetic heterogeneity in argininosuccinate lyase deficiency. Proc Acad Natl Sci USA 81:4480–4484

Mitchell GA, Looney JE, Brody LC, Steel G, Suchanek M, Engelhardt JF, Willard HF, Valle D (1988) Human ornithine-delta-aminotransferase. cDNA cloning and analysis of the structural gene. J Biol Chem 263:14288–14295

Naylor SL, Klebe RJ, Shows TB (1978) Argininosuccinic aciduria: assignment of the argininosuccinate lyase gene to the pter→q22 region of human chromosome 7 by bioautographie. Proc Natl Acad Sci USA 75:6159–6162

O'Brien WE, McInnes RR, Kalumuck K, Adcock M (1986) Cloning and sequence analysis of cDNA for human argininosuccinate lyase. Proc Natl Acad Sci USA 83:7211–7215

Piatigorsky J, O'Brien WE, Norman BL, Kalumuck K, Wistow GJ, Borras T, Nickerson JM, Wawrousek EF (1988) Gene sharing by delta-crystallin and argininosuccinate lyase. Proc Natl Acad Sci USA 85:3479–3483

Povey S, Lovering R, Bruford E, Wright M, Lush M, Wain H (2001) The HUGO Gene Nomenclature Committee (HGNC). Hum Genet 109:678-680

Qureshi AA, Crane AM; Matiaszuk NV, Rezvani I, Ledley FD, Rosenblatt DS (1994) Cloning and expression of mutations demonstrating intragenic complementation in mut0 methylmalonic aciduria. J Clin Invest 93:1812–1819

Sampaleanu LM, Vallee F, Thompson GD, Howell PL (2001) Three-dimensional structure of the argininosuccinate lyase frequently complementing allele Q286R. Biochemistry 40:15570–15580

Stadler S, Gempel K, Bieger I, Pontz BF, Gerbitz KD, Bauer MF, Hofmann S (2001) Detection of neonatal argininosuccinate lyase deficiency by serum tandem mass spectrometry. J Inherit Metab Dis 24:370–378

Takiguchi M, Haraguchi Y, Mori M (1988) Human liver-type arginase gene: structure of the gene and analysis of the promoter region. Nucleic Acids Res 16:8789–8802

Todd S, McGill JR, McCombs JL, Moore CM, Wieder I, Naylor SL (1989) cDNA sequence, interspecies comparison and gene mapping analysis of argininosuccinate lyase. Genomics 4:53–59

Veres G, Craigen WJ, Caskey CT (1986) The 5' flanking region of the ornithine transcarbamylase gene contains DNA sequences regulating tissue-specific expression. J Biol Chem 261:7588–7591

Walker DC, McCloskey DA, Simard LR, McInnes RR (1990) Molecular analysis of human argininosuccinate lyase: mutant characterization and alternative splicing of the coding region. Proc Natl Acad Sci USA 87:9625–9629

Walker DC, Christodoulou J, Craig HJ, Simard LR, Ploder L, Howell PL, McInnes RR (1997) Intragenic complementation at the human argininosuccinate lyase locus. Identification of the major complementing alleles. J Biol Chem 7:6777–6783

Wu YT (1991) Screening for inborn errors of amino acid metabolism. Ann Clin Lab Sci 21:123–142

Yu B, Howell PL (2000) Intragenic complementation and the structure and function of argininosuccinate lyase. Cell Mol Life Sci 57:1637–1651

Yu B, Thompson GD, Yip P, Howell L, Davidson AR (2001) Mechanisms for intragenic complementation at the human argininosuccinate lyase locus. Biochemistry 40:15581–15590