

CLONING AND BIOINFORMATICS ANALYSIS OF ASPARTATE AMINOTRANSFERASE GENE IN *ARABIDOPSIS THALIANA*

HONGYING DUAN, XIAOSHENG DING, YONGANG YU, ZHIQIANG DUAN, YUNLONG HE, ZHIKUN DUAN AND YANQING ZHOU

College of Life Science Henan Normal University, Xinxiang 453007, China

*Corresponding author. Tel.: +86-373-332-3880; Fax: +86-373-332-6341; E-mail address: dxdhy@163.com

Abstract

In this study, gene encoding Aspartate aminotransferase was cloned from *Arabidopsis thaliana* by RT-PCR, named F23N and its bioinformatics analysis was performed. It was found that there is one pyridoxal-phosphate attachment site at 250-270 (SYAKnmGLyGERIG) and other domains, furthermore, the transmembrane helices was not found in F23N, suggesting that F23N encoding product might not belong to membrane-protein. Combined with other Aspartate aminotransferase in *Arabidopsis thaliana*, they have obvious similarities, F23N and *ASP4* take on 96% positives to each other, especially the similarity between F23N and Aspartate aminotransferase (BAE99790) reached 100%. In addition, the conserved genes encoding Aspartate aminotransferase from various species were analyzed by DNAMAN and found that F23N shows obvious comparability with those from some species in Dicotyledons and Monocotyledons. Furthermore, in the phylogenetic tree, these species analyzed were divided into two branches, one branch was composed of Protist, Eukaryotes, the other consist of *Pan troglodytes* and *Tribolium castaneum*, and Eukaryotes was further divided into two branches, which is mostly consistent with biological system.

Introduction

Aminotransferase was first found in 1930, is an effective catalyst for the amino-transformation reaction, and plays an important role in metabolism process of carbon and nitrogen in cells. Aminotransferases could transfer amido from α -amino acid to α -keto acid and catalyze synthesis and hydrolysis (Francisco *et al.*, 1991; Puppo *et al.*, 1992; Mavrides & Orr 1975; Gall *et al.*, 1983). Based on the different donor of amido, Aminotransferases are divided into Aminotransferase A, Aminotransferase B and Aminotransferase C (Rudman & Meister, 1953), Aminotransferase A was shown to catalyze transamination between glutamate and aspartate, tryptophan, tyrosine, and phenylalanine, as well as to a lesser extent, methionine and leucine. Furthermore Aspartate aminotransferase belongs to Aminotransferase A, currently is found in animal, plant and microbe, when pyridoxal phosphate is its coenzyme, Aspartate aminotransferase catalyzes the reverse reaction between do-carboxyl amino acid and α -keto acid (Christen & Metzler 1985), therefore functions of Aspartate aminotransferase are very broad, mostly participates in convergency, transfer and stockpile of nitrogen, and synthesis and hydrolysis of amino acid (Tordjman *et al.*, 2007; Funakoshi *et al.*, 2008). In addition, the main activity of Aspartate aminotransferase is relational with cytoplasm, especially the cytoplasm in leaf fresh (Gelfand & Steinberg, 1977; Powell & Morrison, 1978; Rej, 1981; Yagi & Kagamiyama, 1982; Tordjman *et al.*, 2007).

At present, theoretical research of Aspartate aminotransferase is performed at the molecular level (Jager & Pauptit, 1994). Alongwith development of technology, the catalyse mechanism of Aspartate aminotransferase is studied by molecule simulation and then instruct its reconstruction and reasonable design. In this article the gene encoding Aspartate aminotransferase was cloned from *Arabidopsis thaliana*, the similarity of its putative protein to other Aspartate aminotransferases was detected by BLAST in database and the bioinformatics analysis was performed in order to provide reference for further theory and application studies of Aspartate aminotransferases.

Materials and Methods

Plant materials and growth conditions: Seeds of *Arabidopsis thaliana* were incubated in sterile water for 30min, surface-sterilized with 75% ethanol for 30s, and then sterilized with 5% Sodium hypochlorite for 10min, and washed several times. Subsequently, seeds of *Arabidopsis thaliana* were sown on MS medium and then cultured at 22°C/18°C with a 16h light and 8h dark photoperiod. Two weeks later, *Arabidopsis thaliana* seedlings were harvested, quickly immersed in liquid nitrogen and then stored at -80°C.

RNA isolation: Total RNA was extracted from *Arabidopsis thaliana* seedlings with total RNA isolation system (Promega Corporation, US) according to the instructions of product Z5110 and Z5651 at www.promega.com. The yield of total RNA was determined by spectrophotometry at 260nm, and the purity of total RNA was estimated also by spectrophotometry. The integrity of total RNA was determined by denaturing agarose gel electrophoresis.

RT-PCR amplification: One microgram of total RNA of seedlings was used for RT-PCR with One Step RNA PCR Kit (AMV) (TaKaRa biotechnology Co., Ltd. Japan). The RT-PCR primers were designed according to DNA sequences encoding Aspartate aminotransferases in *Arabidopsis thaliana*, were AF: 5'-AACAAGGTACTATCATTTGAAAAG-3', AR: 5'-ACAAACTTTTAAAT GAATTTTGGG G-3', respectively, and synthesized commercially at GeneCore BioTechnologies Co. Ltd. (Sangon, Shang-hai, China). RT-PCR products were separated by 1.0% agarose gel electrophoresis.

DNA sequencing and bioinformatics analysis: Sequencing of cDNA fragments was performed by GeneCore BioTechnologies Co. Ltd., (Sangon, Shang-hai, China) using the ABI373A automatic sequencer. DNA sequence data were analyzed with DNAMAN, the nucleotide and amino acid sequences were aligned in

NCBI database using BLAST analysis server. Protein identification and characterization, Post-translational modification prediction and so on were performed in <http://www.expasy.org/tools/#proteome>. Alignments of amino acid sequences and construction of gene tree were done with DNAMAN.

Results

Cloning of gene encoding aspartate aminotransferase:

The total RNA of *Arabidopsis thaliana* seedlings was extracted by total RNA isolation system (Promega Corporation, US), primers used to clone DNA sequence were respectively AF and AR (as in materials and methods). One microgram total RNA was amplified, RT-PCR products were determined by 1.0% agarose gel electrophoresis, and there was one special DNA fragment which located between 1000bp and 2000bp. RT-PCR products were reclaimed and ligated into pGEM-T easy vector, subsequently was sequenced by GeneCore BioTechnologies Co. Ltd.

The total cDNA sequence and its predicted amino acid sequence are shown in Fig. 1. This cDNA was composed of 1318bp, contains a single open reading frame and encodes Aspartate aminotransferase which is same with the amino acid sequence of F23N19.17 (GenBank accession number: AAF19543.1), their nucleotide acid sequences only took on various in section. Therefore, the cDNA sequence was named F23N. By analysis of F23N amino acid sequence predicted with PROSITE software, we found that there is one pyridoxal-phosphate attachment site at 250-270 (SYAKnmGLyGERIG) listed in Fig. 1, which exhibits Aminotransferases share certain mechanistic features with other pyridoxal-phosphate dependent enzymes. In addition, some other features and domains in F23N amino acid sequence are found too with SMART (<http://smart.embl-heidelberg.de/>), and listed in Table 1, but their scores are less significant than the required threshold, or they overlap with some other source of annotation.

1	GTACTATCATTGAAAAAGTAACTAGATATTTCCCTTACCCATCAGTCATGAAGTTCGATT	
		M N S I 4
61	CTATCTTCCGTTCTACCAGCTCCAGAAGACCCAGTACTCTCAGTGTACTCAAAGGGATTT	
	L S S V L P A P E D P V L S V Y S K G F	24
121	TCCTCAATTTCACTTATCTCATTTCTAGACGACCCATCACCAGTAAACTAAATCTATCG	
	S S I S L I S F L D D P S P V K L N L S	44
181	GCTGGTACATATCGCACCGAAGAAGGTAAGCCACTAGTACTAGATGTGGTACGACGTGCG	
	A G T Y R T E E G C P L V L D V V R R A	64
241	GAACAGCAACTAGCGAACGACCTATCTCGAGATAAAGAGTACCTACCATCAATGACTC	
	E Q Q L A N D L S R D K E Y L P L N G L	84
301	CCTGAGTTCAATAAGCTTTCAACAAAACCTTATACTAGGGGACGATTCCCCGCTCTTAAA	
	P E F N K L S T K L I L G D D S P A L K	104
361	GAAAACCGAGTAGTTACTACGAGTGTCTCTCGGGAACAGGATCCCTTCGCGTTGGTGCC	
	E N R V V T T Q C L S G T G S L R V G A	124
421	GAGTTTCTTGCCACACATAACAAAGAATCAGTTATTTTCGTCCCGAACCCGACATGGGGA	
	E F L A T H N K E S V I F V P N P T W G	144
481	AATCATCCACGTATCTTTACTCTCGCCGGTCTATCAGTGCAGTACTTTTCGTTACTATGAC	
	N H P R I F T L A G L S V Q Y F R Y Y D	164
541	CCCAAATCACGCGGACTCGATTTTAAGGGAATGCTGGAAGACCTAGGTGCCGCGCCCTT	
	P K S R G L D F K G M L E D L G A A P P	184
601	GGAGCCATCGTAGTTCTACAGGCTTGTGCTCATAACCCCTACCGGAGTAGATCCAACCTTT	
	G A I V V L Q A C A H N P T G V D P T F	204
661	GAGCAATGGGAAAAGATTGCGCGTCTTGTGCGGTCAAAGTCACTTCTACCTTTTTTCGAC	
	E Q W E K I R R L V R S K S L L P F F D	224
721	TCAGCTTATCAAGGATTTGCTTCTGGGTCCCTCGACGCAGATGCACAGGCAGTGCATATG	
	S A Y Q G F A S G S L D A D A Q A V R M	244
781	TTCGTTGCTGATGGTGGTGAATGCCTAATAGCACAGTCCTATGCCAAAAATATGGGCCTC	
	F V A D G G E C L I A Q S Y A K N M G L	264
841	TACGGAGAACGCATTGGATCTCTAACGATTGTATGCACGTCAGAGGATGTTGCTAAAAAA	
	Y G E R I G S L T I V C T S E D V A K K	284
901	GTTGAGAATCAGGTTCTTCTAGTAGTTTCGTCCCATGTACCTGACACCACCCATCCATGGT	
	V E N Q V L L V V R P M Y L T P P I H G	304
961	GCCTCGATAGTAGCAACAATTCTCAAAAACCTCTGATATGTATAACGATTGGACCATCGAA	
	A S I V A T I L K N S D M Y N D W T I E	324
1021	CTAAAGGGAATGGCGGACCGCATTATTTCTATGCGTCAGCAGCTTTATGCAGCACTAGAG	
	L K G M A D R I I S M R Q Q L Y A A L E	344
1081	GCGCGGGTACTCCTGGAGACTGGTCACATATAATCAAACACATAGGCATGTTTACCTTC	
	A R G T P G D W S H I I K H I G M F T F	364
1141	ACCGGCCTATCAGAGGAACAGGTCCGGCTTATGGCGAAAGAGTACCATATCTACATGACA	
	T G L S E E Q V R L M A K E Y H I Y M T	384
1201	TACGATGGTTAATTGAATCAGTGATGAAATGAATTTGTGAGAATTGATTGAATCAGTTGA	
	Y D G *	387
1261	AATGAATTTGACCCCAAAATTCATTAAACCCCAAAATTCATTAAAAA	

Fig.1. Full-length cDNA and putative amino acid sequences of F23N.

Table 1. Other features and domains in F23N protein.

Name	Begin	End	E-value
Calx_beta	28	112	2.28e+02
CASc	162	300	2.96e+03
B3_4	192	316	3.79e+04
BID_2	213	288	6.79e+02
DAGKa	234	349	2.36e+03
GED	261	359	1.69e+03
HTH_XRE	331	386	2.56e+03
SANT	336	382	3.45e+02
PBD	347	378	8.18e+01

The numbers on the left and right represent the positions of the nucleotide sequence or amino acid sequence, respectively. The positions of Aminotransferases class-I pyridoxal-phosphate attachment site is underlined. The asterisk denotes the termination codon.

Bioinformatics analysis of F23N encoding product: In this article, the bioinformatics analysis of F23N amino acid sequence was performed, the Hydrophobicity profile is shown in Fig. 2, the secondary structure was also predicted with DNAMAN (Fig. 3), and the transmembrane helices was not found in F23N with

HMMTOP 2.0, suggesting F23N protein might be secretory protein. It is well known that glycosylation site, phosphorylation site and other sites belong to important post-translational modification, and have main effects on structure and function of protein (Zhao & Liu, 2003). In this research, F23N protein also has several interesting sites predicted with NetPhos 2.0 Server, which were listed in Table 2, there are 18 Phosphorylation sites in all, 11 Ser, 4 Thr and 3 Tyr, respectively, which are consanguineous with activity of protein, possibly activity and function of F23N protein are regulated by these sites.

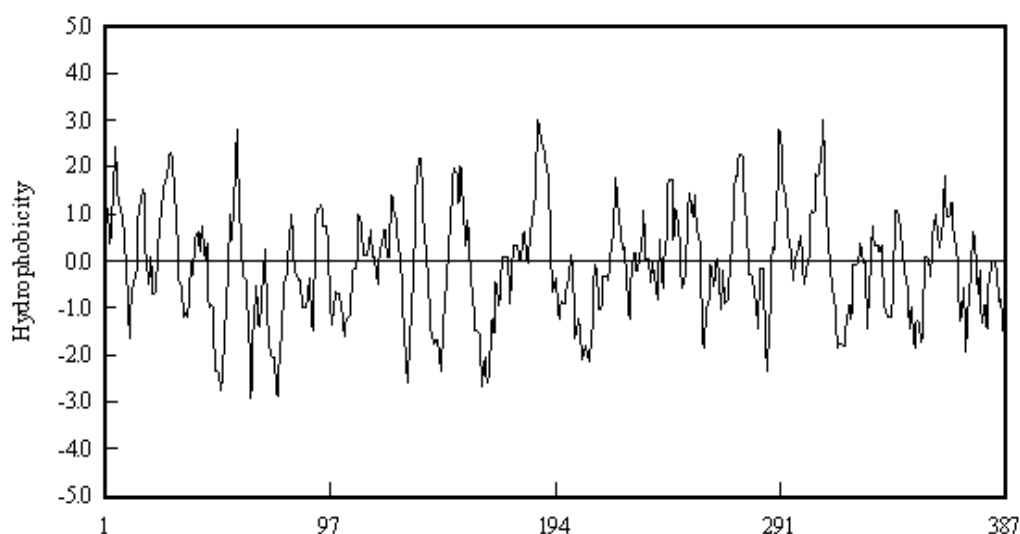


Fig. 2. The Hydrophobicity profile predicted of F23N amino acid sequence by DNAMAN.

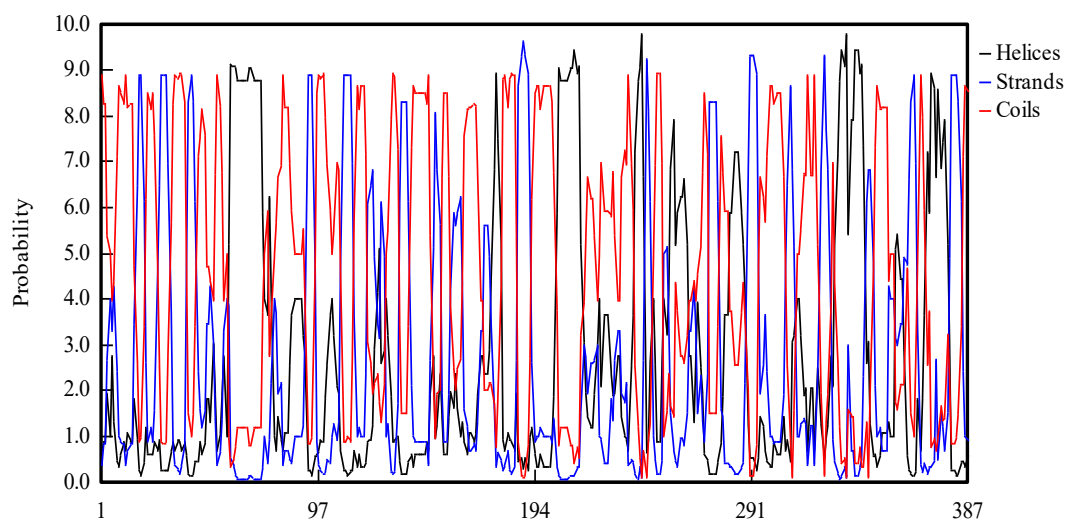


Fig. 3. The secondary structure predicted of F23N protein.

Table 2. The phosphorylation sites predicted in F23N amino acid sequence.

Name	Pos	Context	Score
Serine	25	SKGFSSISL	0.544
	31	ISLISFLDD	0.924
	100	LGDDSPALK	0.638
	115	TQCLSGTGS	0.585
	134	HNKESVIFV	0.550
	167	YDPKSRGLD	0.730
	216	RLVRSKSL	0.600
	278	IVCTSEDVA	0.638
	315	ILKNSDMYN	0.658
	334	DRIISMRQQ	0.964
	368	FTGLSEEQV	0.935
Thr	47	LSAGTYRTE	0.624
	111	RVVTTQCLS	0.762
	151	PRIFTLAGL	0.680
	348	EARGTPGDW	0.984
Tyr	78	RDKEYLPLN	0.530
	227	FDSAYQGFA	0.732
	318	NSDMYNDWT	0.847

The putative amino acid sequence of F23N was compared in database by BLAST and found that F23N has high sequence identity with other Aspartate aminotransferases in *Arabidopsis thaliana* (Fig. 4). For example, F23N and ASP4 (NP-849838) take on 96% positives to each other, F23N shows 75% positives to ASP2 (NP-197456), 48%-69% positive to ASP3 (NP-196713), *Aspartate*

aminotransferase (CAA56932) and ASP5 (NP-194927). Furthermore, the comparability between F23N and *Aspartate aminotransferase* (BAE99790) is extraordinary high (100%), but compared with 387 amino acid residues of F23N, *Aspartate aminotransferase* (BAE99790) is only composed of 240 amino acid residues. In order to further study *Aspartate aminotransferase* F23N, ASP2, ASP3, ASP4, ASP5 and *Aspartate aminotransferase* (CAA56932), their relation was analyzed according to comparability of amino acid sequences. As shown in Fig. 5, *Aspartate aminotransferase* in *Arabidopsis thaliana* were divided into two branches, *Aspartate aminotransferase* (CAA56932) and ASP5 were clustered one branch, the other *Aspartate aminotransferase* were clustered the other branch, in which the relation between F23N and ASP4 is close and ASP2 shows close relation with ASP3.

The amino acid sequences of F23N, ASP4 (NP-849838), ASP2 (NP-197456), ASP3 (NP-196713), *Aspartate aminotransferase* (CAA56932), ASP5 (NP-194927) and *Aspartate aminotransferase* (BAE99790) were compared by DNAMAN. The identity and similarity of amino acid residues were showed dark or gray, respectively. The consensus indicates the conserved amino acids in these *Aspartate aminotransferase*. Furthermore, the similarity proportion of these sequences was annotated below.

The similar relation between F23N, ASP4 (NP-849838), ASP2 (NP-197456), ASP3 (NP-196713), *Aspartate aminotransferase* (CAA56932), ASP5 (NP-194927) and *Aspartate aminotransferase* (BAE99790) were analyzed with DNAMAN. The scale bar represents length of branch.

Molecule phylogenetic analysis of Aspartate aminotransferase: Genes encoding *Aspartate aminotransferase* were screened by BLAST in <http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>, and obtain conserved gene encoding *Aspartate aminotransferase* from various species, and the phylogenetic tree was constructed by DNAMAN to study evolution relationship among gene encoding *Aspartate aminotransferase* in different species.

As shown in Fig. 6, the phylogenetic tree was divided into two branches, one branch was composed of Protist, Eukaryotes such as plant, animal and fungal, the other consist of *Pan troglodytes* and *Tribolium castaneum*. Moreover, Eukaryotes was further divided into two branches, one was composed of green alga, Dicotyledon and Monocotyledon belonging to Angiosperm and according to the evolution relationship, the other branch consist of in turn some species belonging to Protozoa such as *Dictyostelium discoideum AX4*, *Yarrowia lipolytica CLIB1229*, *Ustilago maydis 521*, *Malassezia globosa CBS 7966*, *Cryptococcus neoformans var. neoformans JEC21*

and *Laccaria bicolor S238N-H82*, *Drosophila pseudoobscura* and *Apis mellifera* in Phylum Arthropods and animals in Craniota.

The genes encoding *Aspartate aminotransferase* from *Arabidopsis thaliana* (AAF19543), *Populus trichocarpa* (ABK95824), *Vitis vinifera* (CAO65506), *Daucus carota* (P28734), *Lupinus angustifolius* (AAA50160), *Oryza sativa Japonica Group* (NP_001044317), *Glycine max* (AAC50015), *Panicum miliaceum* (CAA45023), *Medicago sativa* (CAA43779), *Lotus japonicus* (CAA63894), *Triticum aestivum* (ABY58643), *Apis mellifera* (XP_396131), *Securigera parviflora* (AAL09704), *Chlamydomonas reinhardtii* (XP_001695040), *Dictyostelium discoideum AX4* (XP_646849), *Solanum tuberosum* (ABB55364), *Xenopus laevis* (NP_001080255), *Phaseolus vulgaris* (AAN76499), *Xenopus tropicalis* (NP_001016933), *Ornithorhynchus anatinus* (NP_001016933), *Lotus corniculatus* (AAC12674), *Ustilago maydis 521* (XP_756742), *Oryza sativa* (AAO23563), *Pan troglodytes* (XP_523381), *Tribolium castaneum* (XP_969549), *Drosophila pseudoobscura* (XP_001356945), *Mus musculus* (AAB91426), *Monodelphis domestica* (XP_001376001), *Yarrowia lipolytica CLIB122* (XP_500415), *Rattus norvegicus* (NP_037309), *Bos taurus* (NP_777231), *Macaca mulatta* (XP_001103601), *Cryptococcus neoformans var. neoformans JEC21* (XP_568414), *Macaca fascicularis* (Q4R559), *Gallus gallus* (7AATA), *Pongo pygmaeus* (Q5REB0), *Sus scrofa* (NP_999093), *Danio rerio* (NP_998544), *Canis lupus familiaris* (XP_535278), *Homo sapiens* (AAH00525), *Laccaria bicolor S238N-H82* (XP_001873821), *Sus scrofa domestica* (0308236A), *Malassezia globosa CBS 7966* (XP_001731601), *Tetrahymena thermophila SB210* (XP_001017054) and *Pan troglodytes* (NP_001092011) respectively were analyzed. The scale bar represents the branch length.

F23N	MNSILSSVLPAP	12
ASP2	MDSVFSNVARAP	12
ASP3	MKTTHFSSSSSSDRRIGALLRHLNSGSDSDNLSSLYASPTSGGTGGSVFSLVQAP	56
ASP4	MNSILSSVLPAP	12
ASP5	MASLMLSIGSTSLPREINKDKLKLGTASNPFLKAKSFSRVMTMTAVAKPSRFEGITMAP	60
BAE99790		0
CAA56932	MASLMLSIGSTSLPREINKDNVKGLTASNPFLKAKSFSRVMTMTAVAKPSRFEGITMAP	60
Consensus		
F23N	EDPVLSVYSKGFSSISLISFLDDPSPVKLNLSAGTYRTEEGKPLVLDVVRRAEQQLANDL	72
ASP2	EDPILGVT.....VAYNNDPSPVKINLGVGAYRTEEGKPLVLDVVRKAEQQLVNDP	63
ASP3	EDPILGVT.....VAYNKDPSPVKLNLGVGAYRTEEGKPLVLDVVRKAEQQLINDR	107
ASP4	EDPVLSV.....IFACRDDPSPVKLNLSAGTYRTEEGKPLVLDVVRRAEQQLANDL	63
ASP5	PDPIILGVS.....EAFKADTNGMKLNLGVGAYRTEELQPYVLNVVKKAE.N.LMLER	110
BAE99790		0
CAA56932	PDPIILGVS.....EAFKADTNGMKLNLGVGAYRTEELQPYVLNVVKKAE.N.LMLER	110
Consensus		
F23N	SRDKEYLPLNGLPEFNKLSTKLILGDDSPALKENRVVTTQCLSGTGSLRVGAEFLLATHNK	132
ASP2	SRVKEYIPIVGISDFNKLSAKLILGADSPAITESRVTTVQCLSGTGSLRVGAEFLLKTHYH	123
ASP3	TRIKEYLPVGLVEFNKLSAKLILGADSPAIRENRIITVECLSGTGSLRVGAEFLAKHYH	167
ASP4	SRDKEYLPLNGLPEFNKLSTKLILGDDSPALKENRVVTTQCLSGTGSLRVGAEFLLATHNK	123
ASP5	GDNKEYLPIEGLAAFNKATAELLLFGAGHPVIKEQRVATIQQLSGTGSLRLAAALIERYP	170
BAE99790		0
CAA56932	GDNKEYLPIEGLAAFNKATAELLLFGAGHPVIKEQRVATIQQLSGTGSLRLAAALIERYP	170
Consensus		
F23N	ESVIFVPNPTWGNHPRIFTLAGLSVQYFRYYDPKSRGLDFKGMLEDLGAAPPGAIVVLQA	192
ASP2	QSVIYIPKPTWGNHPKVFENLAGLSVEYFRYYDPATRGLEDLGAAPSGAIVVLHA	183
ASP3	QKTIYITQPTWGNHPKIFTLAGLTVKTYRYYDPATRGLEDFKGMLEDLGAAPPGAIVVLHA	227
ASP4	ESVIFVPNPTWGNHPRIFTLAGLSVQYFRYYDPKSRGLDFKGMLEDLGAAPPGAIVVLQA	183
ASP5	GAKVVISSPTWGNHKNIFENDAKVPWSEYRYYDPKTIIGLDFEGMIADIKEAPEGSFILHGH	230
BAE99790	MLEDLGAAPPGAIVVLQA	18
CAA56932	GAKVVISSPTWGNHKNIFENDAKVPWSEYRYYDPKTIIGLDFEGMIADIKEAPEGSFILHGH	230
Consensus	d a g l	
F23N	CAHNPTGVDPDTFEQWEKIRRLVRSKSLLPFFDSAYQGFASGSLDADAQAVRMFVADGGEC	252
ASP2	CAHNPTGVDPDTSEQWEQIRQLMRSKSLLPFFDSAYQGFASGSLDADAQAVRMFVADGGEC	243
ASP3	CAHNPTGVDPDTIQWEQIRKLMSKSLLPFFDSAYQGFASGSLDADAQAVRMFVADGGEC	287
ASP4	CAHNPTGVDPDTFEQWEKIRRLVRSKSLLPFFDSAYQGFASGSLDADAQAVRMFVADGGEC	243
ASP5	CAHNPTGVDPDTPEQWVKIADVIQEKNIHPPFDVAYQGFASGSLEDAASVRLFAERGMF	290
BAE99790	CAHNPTGVDPDTFEQWEKIRRLVRSKSLLPFFDSAYQGFASGSLDADAQAVRMFVADGGEC	78
CAA56932	CAHNPTGVDPDTPEQWVKIADVIQEKNIHPPFDVAYQGFASGSLEDAASVRLFAERGMF	290
Consensus	cahnptg dpt qw i k pffd ayqgfasgsld da r f g e	
F23N	LIAQSYAKNMGLYGERIGSLTIVCTSEDVAKKVENQVLLVVRPMYLTPTTHGASIVATIL	312
ASP2	LIAQSYAKNMGLYGERVGLSIVCKSADVASKVESQVKLVVRPMYSSPPTTHGASIVATIL	303
ASP3	LVAQSYAKNMGLYGERVGLSIVCKSADVAGRVESQVKLVVRPMYSSPPTTHGASIVATIL	347
ASP4	LIAQSYAKNMGLYGERIGSLTIVCTSEDVAKKVENQVLLVVRPMYLTPTTHGASIVATIL	303
ASP5	FVAQSYSKNLGLYAERIGAINVVCSSADAATRVKSQKRIARPMYSNPPVHGARTIVANV	350
BAE99790	LIAQSYAKNMGLYGERIGSLTIVCTSEDVAKKVENQVLLVVRPMYLTPTTHGASIVATIL	138
CAA56932	FVAQSYSKNLGLYAERIGAINVVCSSADAATRVKSQKRIARPMYSNPPVHGARTIVANV	350
Consensus	aqsy kn gly er g vc s d a v q rpmy pp hga iva	
F23N	KNSDMYNDWTIELKGMADRIISMRRQLYAAL EARGTPG. DWSHI IKHIGMFTFTGLSEBQ	371
ASP2	KSSDMYNNTWIELKEMADRIKSMRQQLFEAIQARGTPG. DWSHI IKQIGMFTFTGLNKEQ	362
ASP3	RDKNLFNEWTIELKAMADRIISMRRQLFEALRTRGTPG. DWSHI IKQIGMFTFTGLNPAQ	406
ASP4	KNSDMYNDWTIELKGMADRIISMRRQLYAAL EARGTPG. DWSHI IKHIGMFTFTGLSEBQ	362
ASP5	GDVTMFSEWKAEMEMMAGRIKTVRQELYSLSVSKDKSGKDWFSILKQIGMFSFTGLNKAQ	410
BAE99790	KNSDMYNDWTIELKGMADRIISMRRQLYAAL EARGTPG. DWSHI IKHIGMFTFTGLSEBQ	197
CAA56932	GDVTMFSEWKAEMEMMAGRIKTVRQELYSLSVSKDKSGKDWFSILKQIGMFSFTGLNKAQ	410
Consensus	w e ma ri r l g dws i k igmf ftgl q	
F23N	VRLMAKEYHIYMTYDG.....	387
ASP2	VEFMTKEFHIIYMTSDGRISMAGLSSKTVPHLADAMHAAVTRL	404
ASP3	VSPMTKEYHIYMTSDGRISMAGLSSKTVPHLADAIHAVVTKA	448
ASP4	VRLMAKEYHIYMTYDGRISMAGLSSKTVPLADAIHAVVTRI	404
ASP5	SDNMTDKWHVYMTKDGRISLAGLSLAKCEYLADAIIDSYHN	452
BAE99790	VRLMAKEYHIYMTYDGRISMAGLSSKTVPLADAIHAVVTRI	239
CAA56932	SDNMTDKWHVYMTKDGRISLAGLSLAKCEYLADAIIDSYHN	452
Consensus	m h ymt dg	

Homology matrix of 7 sequences

F23N	100%					
ASP2	77.5%	100%				
ASP3	71.2%	81.0%	100%			
ASP4	98.7%	78.0%	72.3%	100%		
ASP5	49.1%	52.5%	49.1%	48.6%	100%	
BAE99790	100.0%	82.1%	75.8%	100.0%	50.4%	100%
CAA56932	49.1%	52.5%	48.9%	48.6%	99.6%	50.4%

Fig. 4. Alignment between amino acid sequences of F23N and other Aspartate aminotransferase in *Arabidopsis thaliana*.

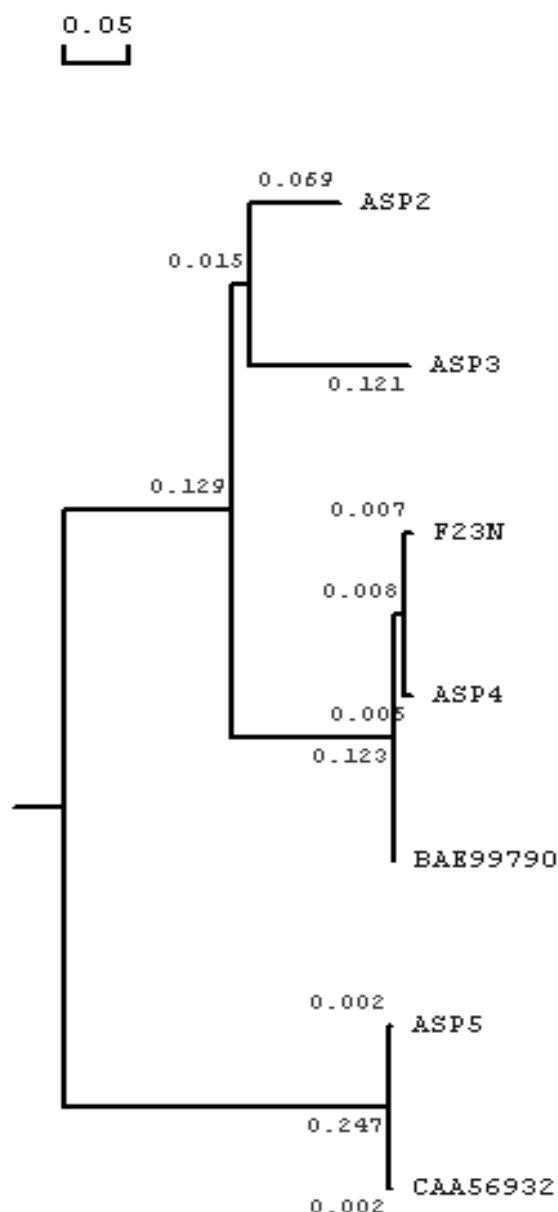


Fig. 5. The relation of Aspartate aminotransferase in *Arabidopsis thaliana*.

In addition, genes encoding Aspartate aminotransferase from various species were screened by BLAST and analyzed with DNAMAN and discovered that Aspartate aminotransferase F23N from *Arabidopsis thaliana* shows obvious comparability with some species belonging to Dicotyledon and Monocotyledon, 71%~75% or so. Furthermore, in the phylogenetic tree these species analyzed were divided into two branches, and mostly consistent with biology system among species. However there are not consistent with the classic classification system in phylogenetic tree, for example, relationship between *Pan troglodytes* (NP_001092011) and *Homo sapiens* is far, and the comparability is 48.9%, but *Pan troglodytes* (XP_523381) exhibit higher similarity with *Homo sapiens*, 99.5%. The amino acid sequence of NP_001092011 and XP_523381 was further compared, their similarity is merely 48.6%, which probably shows

Discussion

In this article, one gene encoding Aspartate aminotransferase was cloned from *Arabidopsis thaliana* by RT-PCR, composed of 1318bp, and its encoding product is same with the amino acid sequence of F23N19.17 (AAF19543.1), but their nucleotide acid sequences are partly various, inferring RT-PCR amplification or sequencing might result in the different nucleotide acid sequences. Therefore, the cDNA sequence was named F23N, belongs to one member of Aspartate aminotransferases family which catalyzes the reverse reaction between do-carboxyl amino acid and α -keto acid, mostly participates in convergency, transfer and stockpile of nitrogen, and synthesis and hydrolysis of amino acid (Christen & Metzler 1985; Vipin & Robert, 1991; Mark *et al.*, 2003; Zhou *et al.*, 2009).

Bioinformatics analysis of F23N encoding product with PROSITE, it was found that there is one pyridoxal-phosphate attachment site at 250-270 (SYAKnmGLyGERIG), which suggests that Aminotransferases share certain mechanistic features with other pyridoxal-phosphate dependent enzymes, such as the covalent binding of the pyridoxal-phosphate group to a lysine residue (Sung *et al.*, 1991), furthermore the sequence around the pyridoxal-phosphate attachment site of this class of enzyme is sufficiently conserved to allow the creation of a specific pattern (David & Naomi, 1977; Yoshimura *et al.*, 1993; Grishin *et al.*, 1995; Deu *et al.*, 2009). Search of the putative amino acid sequence in database revealed that F23N has high sequence identity with other Aspartate aminotransferases in *Arabidopsis thaliana* (Torre *et al.*, 2006), especially shows 100% positive to Aspartate aminotransferase (BAE99790), but compared with 387 amino acid residues of F23N, Aspartate aminotransferase (BAE99790) is only composed of 240 amino acid residues, therefore they are not the same gene, possibly Aspartate aminotransferase (BAE99790) is part of F23N. Furthermore, the relation between Aspartate aminotransferase F23N, ASP2, ASP3, ASP4, ASP5 and Aspartate aminotransferase (CAA56932) was further analyzed and found Aspartate aminotransferase (CAA56932) and ASP5, F23N and ASP4, ASP2 and ASP3 shows close, respectively.

that *Pan troglodytes* NP_001092011 and XP_523381 are paralog. Whereas, *Pan troglodytes* (NP_001092011) and *Tribolium castaneum* (XP_969549) were clustered together, and exhibit parallel relationship with other species. Therefore, it is presumed that gene (NP_001092011) in *Pan troglodytes* might be evolved from gene (XP_523381). In addition, based on the phylogenetic tree, we also found *Tetrahymena thermophila* SB210 belong to Protozoa which is the furthest primal phylum in animal kingdom, *Dictyostelium discoideum* AX4 which is one kind of amoebae locates in the bottom of Metazoa, and fungi has closer relationship with animal than plant. Taken together, the clustering result according to Aspartate aminotransferases from various species was mainly consistent with traditional classification among species.

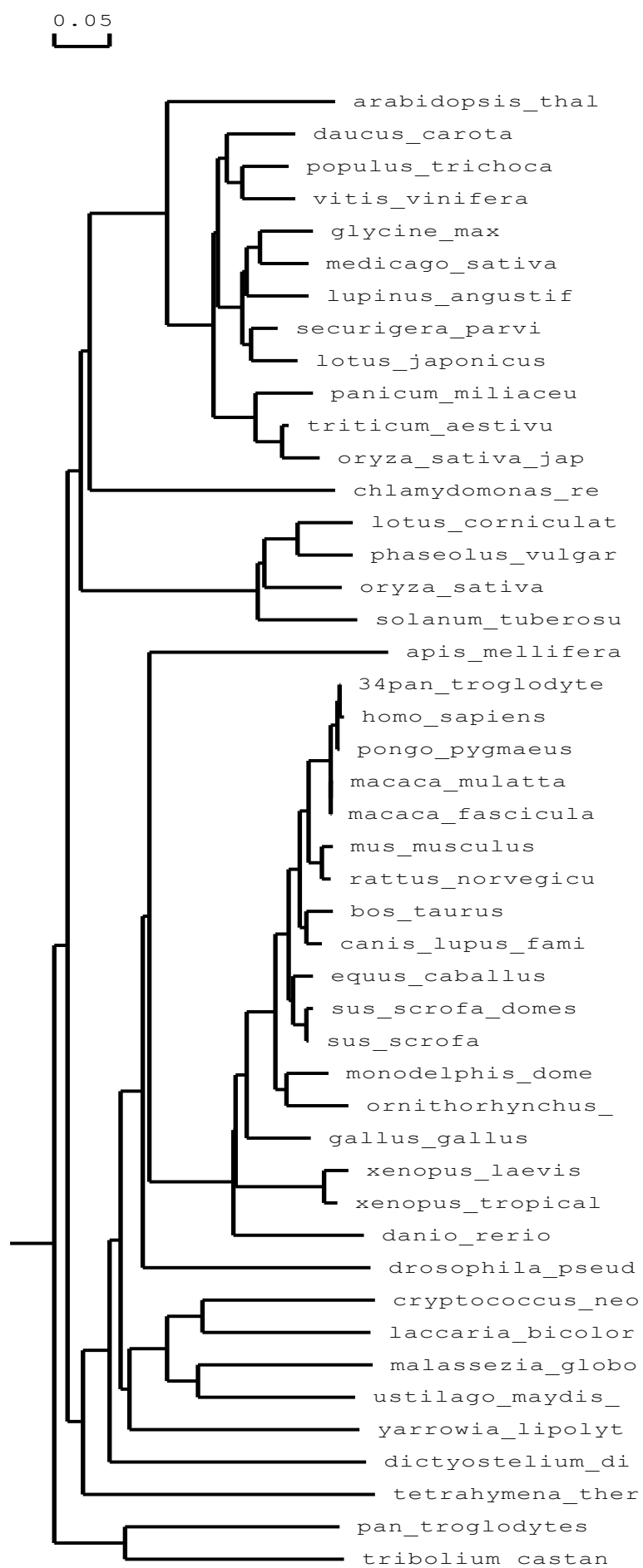


Fig. 6. The phylogenetic tree was constructed based on amino acid sequences of Aspartate aminotransferases from various species.

Acknowledgments

This study was supported by Science Fund of Henan (102300410265, 2009A180009, 2009A180008, 2008A180015, 2008B180007, 08N042, 082300430270, and 082102120019), P. R. China.

References

- Christen, P. and D.E. Metzler. 1985. Transaminases. eds. New York: John Wiley & Sons, Inc., pp. 102-202.
- David, H. G. and R. Naomi. 1977. Mapping of the Aspartate and Aromatic amino acid aminotransferase genes *tyrB* and *aspC*. *J. Bacteriology*, 130(1): 441-444.
- de la Torre, F., L.D. Santis, M.F. Suárez, R. Crespillo and F.M. Cánovas. 2006. Identification and functional analysis of a prokaryotic-type aspartate aminotransferase: implications for plant amino acid metabolism. *Plant J.*, 46: 414-425.
- Deu, E., J. Dhoot and J.F. Kirsch. 2009. The partially folded homodimeric intermediate of Escherichia coli aspartate aminotransferase contains a "molten interface" structure. *Biochemistry*, 48(2): 433-441.
- Francisco, J.G., C. Maria and M. Angel. 1991. Purification and characterization of aspartate amino transferase. *Biol. J.*, 278(15): 149-154.
- Funakoshi, M., M. Sekine, M. Katane, T. Furuchi, M. Yohda, T. Yoshikawa and H. Homma. 2008. Cloning and functional characterization of Arabidopsis thaliana D-amino acid aminotransferase--D-aspartate behavior during germination. *FEBS J.*, 275(6): 1188-1200.
- Gall, I. and K. Erzsebet. 1983. The characteristics of aspartate transferase enzyme in tobacco tissue cultures. *Acta Biol.*, 29(1): 3-10.
- Gelfand, D.H. and R.A. Steinberg. 1977. *E.Coli* mutants deficient in the aspartate and aromatic aminotransferase. *J. Bacteriology*, 130(1): 429-440.
- Grishin, N.V., M.A. Philips and E.J. Goldsmith. 1995. Modeling of the spatial structure of eukaryotic ornithine decarboxylases. *Protein Sci.*, 4: 1291-1304.
- Jager, J. and R.A. Paupit. 1994. Three dimensional structure of a mutant E. coli aspartate aminotransferase with increased enzymic activity. *Protein Engineering*, 7 (5): 605- 612.
- Mark, A.B., P.N. Linh and A.B. Christopher. 2003. Aspartate Aminotransferase Generates Proagonists of the Aryl Hydrocarbon Receptor. *Mol Pharmacol*, 64(3): 550-556.
- Mavrides, C. and W. Orr. 1975. Multispecific aspartate and aromatic amino amitransferase in *E.Coli*. *J Biol Chem.*, 250(1): 4128-4133.
- Powell, J.T. and J.F. Morrison. 1978. The purification and properties of the aspartate aminotransferase and aromatic acid aminotransferase from *E.Coli*. *Eur J Biol Chem.*, 87(9): 391-400.
- Puppo, J., J.M. Forja and J. Blasco. 1992. Biochemical characteristic of aspartate aminotransferase from gills and digestive gland of the manile clam. *Comp Biochem Physiol.*, 247(8): 2486-2492.
- Rej, R. 1981. Effects of temperature on the steady-state kinetics and measurement of aspartate aminotransferase. *Clin Chem*, 27 (2): 213-219.
- Rudman, D. and A. Meister. 1953. Transamination in *Escherichia coli*. *J. Biol. Chem.*, 200: 591-604.
- Sung M.H., K. Tanizawa, H. Tanaka, S. Kuramitsu, H. Kagamiyama, K. Hirotsu, A. A. Okamoto, T. Higuchi and K. Soda. 1991. Thermostable aspartate aminotransferase from a thermophilic *Bacillus* species, Gene cloning, sequence determination, and preliminary x-ray characterization. *J. Biol. Chem.*, 266: 2567-2572.
- Tordjman, J., S. Leroyer, G. Chauvet, J. Quette, C. Chauvet, C. Tomkiewicz, C. Chapron, R. Barouki, C. Forest, M. Aggerbeck and B. Antoine. 2007. Cytosolic aspartate aminotransferase, a new partner in adipocyte glyceroneogenesis and an atypical target of thiazolidinedione. *J. Biol. Chem.*, 282(32): 23591-23602.
- Vipin, K.R and J. Robert. 1991. Aspartate aminotransferase activity is required for aspartate catabolism and symbiotic nitrogen fixation in *Rhizobium meliloti*. *J. Bacteriology*, 173(9): 2879-2888.
- Yagi, T. and H. Kagamiyama. 1982. Aspartate aminotransferase from bakes'yeast. *J Biol Chem.*, 92(1): 35-43.
- Yoshimura, T., K. Nishimura, J. Ito, N. Esaki, H. Kagamiyama, J.M. Manning and K. Soda. 1993. Unique stereospecificity of D-amino acid aminotransferase and branched-chain L-amino acid aminotransferase for C-4' hydrogen transfer of the coenzyme. *J Am Chem Soc.*, 115: 3897-3900.
- Zhao, H.L. and Z.M. Liu. 2003. Protein Glycoengineering. *Progress in Biotechnology*, 9: 18-20.
- Zhou, Y., H. Cai, J. Xiao, X. Li, Q. Zhang and X. Lian. 2009. Over-expression of aspartate aminotransferase genes in rice resulted in altered nitrogen metabolism and increased amino acid content in seeds. *Theor Appl Genet.*, 118(7): 1381-1390.

(Received for publication 16 September 2009)