BIOLOGIJA. 2010. Vol. 56. No. 1–4. P. 41–48
DOI: 10.2478/v10054-010-0012-5

Phylogeny of structural domains of plant serine β-lactamase family proteins

Danas Baniulis[1]*,

Julius Liobikas[2],

Subhash Mohan Agarwal[3],

Birutė Frercks[1],

Gintautas Vaitiekaitis[4],

Vidmantas Stanys[1]

[1] *Institute of Horticulture, Lithuanian Research Center for Agriculture and Forestry, Kauno Str. 30, LT-54333 Babtai, Kaunas reg., Lithuania*

[2] *Institute for Biomedical Research, Kaunas University of Medicine, Kaunas, Lithuania*

[3] *Institute of Cytology & Preventive Oncology, NOIDA 201 301, India*

[4] *Department of Physics, Mathematics and Biophysics, Kaunas University of Medicine, Kaunas, Lithuania*

Bacterial serine β-lactamases (βL) constitute a class of penicillin-binding proteins that hydrolyze β-lactams and prevent the inhibition of enzymes involved in peptidoglycan synthesis. Genes homologous to βL have been discovered in *Protozoa*, *Metazoa* and *Plantae*; however, the function of the non-bacterial genes remains elusive. Structural and phylogenetic analysis has revealed important hints related to the function. The non-bacterial protein structure includes a conservative three-motif amino acid sequence signature characteristic of the active center of βLs. In addition, plant βL homologue represents a multimodular enzyme containing an ABC1 domain at the N-terminal part of the polypeptide chain. Characterization of the diversity of βL homologues provides guidelines for experimental investigations on the protein function. Four distinct alloparalogous lineages of the protein have been identified in metazoan βL homologues. In this study, the diversity of plant βL homologue was characterized by assessing the phylogeny of functional domains. Sequence collection from genomic and assembled EST databases resulted in a set of 57 sequences with at least partial coverage of the structural domains. A phylogenetic diversity was estimated at the ABC1 domain and at three regions corresponding to the active-site motifs SXXK, [SY]X[NT], [K / H] [T / S]G within the βL domain. Protein sequence parsimony analysis revealed the presence of at least four distinct phylogenetic clusters in each region of the catalytic motifs of the βL domain, suggesting a structural and functional diversity of the protein family in plants. Sequences of the ABC1 domain are closely related among the multimodular proteins homologous to βL and constitute a distinct phylogenetic group within the ABC1 protein family in plants.

Key words: serine β-lactamase, LACTB, ABC1, sequence alignment, structural domain, phylogeny

Abbreviations: βL – β-lactamases; HMM – PBPs motif; – penicillin-binding proteins; PUTs – plant GDB-assembled unique transcripts.

INTRODUCTION

Penicillin-binding proteins (PBPs) represent specialized acyl serine transferases involved in the assembly and metabolism of the bacterial wall peptidoglycan that are inhibited by penicillins and cephalosporins [1, 2]. Serine β-lactamases constitute a class of PBPs that hydrolyze the β-lactams via formation of hydrolytically labile serine ester-linked β-lactamoyl enzyme intermediate. The production of β-lactamases is a defensive mechanism developed by bacteria to protect their wall peptidoglycan-synthesizing machinery against the toxic effect of penicillin. Genes homologous to βL have been discovered in *Protozoa*, *Metazoa* [3, 4] and *Plantae* [5]. However, in these organisms, neither the role of βL has been defined nor have the genes involved in the peptidoglycan synthesis pathway been detected.

* Corresponding author. E-mail: d.baniulis@lsdi.lt

Therefore, the presence of this particular gene is intriguing and raises the fundamental question about its biological role in the organism.

The phylogenetic analysis of metazoan βL homologs has revealed that the family is composed of four distinct alloparalogous protein lineages deriving from four separate bacterial genes of low molecular weight PBPs of βL subclass [6]. An extensive diversity of proteins has been observed in nematodes. In vertebrates, only homologoues of two types, LACTB and LACTB-like 2, have been found. All of these genes retain the same fold, three-motif amino acid sequence signature characteristic of the active site. Mitochondrial filaments of the metazoan homologoue LACTB have been studied, and a function involved in intramitochondrial membrane organization [7], as well as a hypothetical function related to bacterial peptidoglycan sensory mechanisms in intestine [6] have been proposed.

There is no experimental data to provide a hint on the function of the β-lactamase homologue in plants. The enzymatic βL activity has been identified in plants, but it has been shown to be characteristic of enzymes of the glyoxilase family [8]. There have been no reports on a biological function involving β-lactams in plants; therefore, it is not clear why there should be βL activity. However, multiple sequence alignments and a secondary structure analysis of βL homologue revealed conserved regions and catalytic site residues characteristic of the active center of serine acyl transferases [5], and a hypothetical function of serine peptidase activity involved in the programmed cell death pathway has been proposed for the plant protein. The complexity to the problem of the identification of the function of the plant βL homologue is enhanced by the fact that the protein comprises a unique multimodular protein containing an ABC1 (activity of bc1 complex) domain at the N-terminal part of the polypeptide chain. Bacterial multimodular PBPs have been known for a long time [1]. Class A high molecular mass PBPs are bifunctional enzymes that catalyze glycosyltransfer and transpeptidation reactions during peptidoglycan biosynthesis. However, none of the bacterial βLs or their metazoan homologues identified so far include an ABC1 domain. This domain displays a similarity to protein kinases, and it has been suggested that members of the ABC1 family are novel chaperonins [9]. However, the exact molecular function of ABC1 proteins is not clear.

Sequence analysis revealed a diversity of homologues of βLs in plants [5]. Previous studies demonstrated a variation of the sequence composition of plant homologue sharing a 38% to 63% amino acid identity within the βL homology domain among the proteins identified in plants from the genera *Arabidopsis*, *Medicago*, *Oryza*, *Phaseolus*, *Solanum* and *Pinus*. The motif [K / H][T / S]G characteristic of the βL active center was not detected in *Medicago*, *Phaseolus*, *Solanum* and *Pinus* sequences, suggesting a structural and, likely, a functional diversity of plant βLs.

A more complete characterization of the diversity of plant homologues of βL would guide experimental investigations on the protein function. Therefore, in this study, the diversity of plant βL homologue was characterized by a phylogenetic analysis of its functional domains. We searched for plant genetic information databases and collected sequences containing a full or a partial βL domain. Separate phylogenetic analyses were carried out for the segments including three active-site motifs SXXK, [SY] X[NT] and [K / H][T / S]G of the βL and ABC1 domains. The study revealed the presence of different groups of βLs and sequence diversity within species. At least four distinct phylogenetic clades were identified in segments including active-site motifs of the βL domain. Sequences of the ABC1 domain of the multimodular proteins containing a βL domain constituted a distinct group within the ABC1 protein family in plants.

MATERIALS AND METHODS

Sequence collection. The pfam beta-lactamase family (Pfam id PF00144) HMM model v.15 was employed to collect sequences of plant βL homologues [10]. Sequences published in the Pfam-A families database v. 22 based on UniProt release v. 9.7 were used. Additional sequences were collected from the Plant GDB PUT database [11] using the hidden Markov motif (HMM) model v. 15 of the Pfam beta-lactamase family domain. A six-frame translation of the PUT nucleotide sequence databases was made using the *transeq* program of the EMBOSS package [12]. The translated databases were searched using the *hmmsearch* of the HMMER package v. 2.3.2 [13].

Sequences of plant ABC1 family proteins were obtained from the Pfam-A families database v. 22. To characterize the ABC1 domain of the collected βL homologoues, a HMM model v. 9 of the Pfam ABC1 family was employed.

Multiple sequence alignment. Multiple sequence alignment was built using *hmmalign* v. 2.3.2 from the HMMER package [13], using a corresponding HMM model. The alignments were reviewed manually, and short or redundant sequences were removed.

Phylogenetic analysis. Segments for phylogenetic analysis were identified using Gblocks v. 0.91b [14] and were adjusted manually. A phylogenetic protein sequence parsimony analysis of amino acid alignments was performed using *protpars* from the PHYLIP package v. 3.68 [15]. A maximum likelihood analysis of amino acid alignments was performed employing *proml* from the PHYLIP package and using the Jones, Taylor and Thornton probability model [16]. The robustness of inferred trees was assessed by bootstrapping using *seqboot*, *protpars* programs (in case

of parsimony analysis) or *proml* (maximum likelihood analysis) and *consense* from the PHYLIP package; 100 replicates was used in the bootstrapping analysis. The mouse βL and ABC1 domain homologue sequences (SwissProt id Q9EP89 and Q9D0L4, respectively) were used to root the inferred tree.

3D modeling of protein structure. A secondary structure diagram of the βL domain model was built using the I-TASSER server for protein 3D structure prediction [17]. Three-dimensional structures were visualized using the *ccp4mg* program v. 1.1.1 [18].

RESULTS AND DISCUSSION

βL belongs to a large family of serine proteases, which is distinguished by three conserved amino acid motifs that contribute to the formation of the catalytic site. The SXXK motif (I) contains the catalytic serine residue which undergoes reversible acylation through substrate binding, whereas the [SY]X[NT] (II) and the [KH][ST]G (III) motifs contribute to substrate docking [19]. A 3D structure diagram of a model of the plant βL domain based on *A. thaliana* sequence is shown in Fig. 1A. However, it should be noted that two regions corresponding to the N-terminal ABC1 domain and 86 a long C-terminal region (Val817 through Arg903 of *A. thaliana* βL) were not considered during homology modeling as they have no known equivalent three-dimensional configuration. Thus, these regions are highly specific to plant βL, and we hypothesise that the C-terminal region is involved in the interaction with the ABC1 domain. The resulting model demonstrates that the putative structure of the plant βL homologue closely resembles bacterial a βL and includes the same structural features involving two helix bundles separated by a β-sheet. The three catalytic motifs are located on the conserved βL domain of the plant homologue.

To collect sequences of plant βL homologues, the HMM model of the Pfam beta-lactamase family domain (Pfam id PF00144) was employed. Twenty-six sequences of *Embryophyta* (land plants) were obtained from the Pfam database [10].

The pfam database incorporates sequences from the UniProtKB database. The database represents only a small part of plant genome data contained in EST libraries, which remains unexploited. Since ESTs generally correspond to only partial cDNA sequences and are typically highly redundant, data collection and analysis is complicated. However, the complexity is resolved by using the assemblies of putative unique transcript contigs of ESTs; threfore, the PlantGDB-assembled Unique Transcripts (PUTs) database [11], including assembled DNA sequences from 144 species, has been employed in this study. The search using a statistical model of the Pfam beta-lactamase family domain identified 176 homologous sequences with a significant E-value score. A multiple alignment of the collected sequences was built using the statistical model of the Pfam beta-lactamase family domain. Analysis of the alignment revealed that only 13 of the collected sequences covered the complete beta-lactamase family domain (corresponding to the 329 residues of the statistical model), and only seven sequences contained both domains (ABC1 and βL). The partial domain sequence coverage might be a result of in-

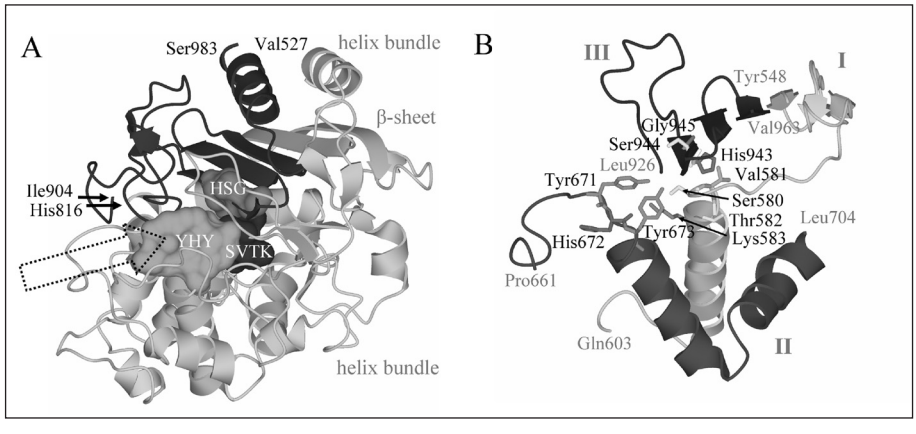

Fig. 1. A 3D structure model of plant βL domain: A – a 3D structure ribbon diagram of the model of βL domain was built using the I-TASSER server for protein 3D structure prediction [Zhang, 2008]. The model includes *A. thaliana* βL (SwissProt id Q93YN5) sequence regions Val527-His816 and Ile904-Ser983. The predicted model C-score = 0.61; the expected TM-score = 0.80 ± 0.09, the expected RMSD = 5.3 ± 3.4Å. The 86 a. a. non-homologous region (starting and ending residues are indicated by small arrows) is omitted. Residues involved in the putative acyl serine transferase active center motif [Liobikas, 2006] are shown as a molecular surface, and the corresponding residue triads and a tetrade are indicated in a single-letter code. An active site cavity characteristic of bacterial βL is indicated with a dotted arrow; B – a Ribbon diagram of sequence segments (labeled with Roman numerals) used for phylogenetic analysis. The starting and ending residues of the segment are labelled in gray. Putative acyl serine transferase active center motif residues are shown as cylinders and labelled in black

complete gene model assemblies published in the PUT database. However, the possibility of variation in the domain and the catalytic motif composition among the plant βL homologues was demonstrated in a previous study which identified the [K / H][T / S]G motif only in sequences of *Arabidopsis*, *Oryza* plants, but not in *Medicago*, *Phaseolus*, *Solanum* and *Pinus*, suggesting a structural diversity of plant βLs [Liobikas, 2006]. The requirement of the βL active site motifs and βL and ABC1 domains for the functioning of plant βL homologues remains inexplicable. Therefore, for a comprehensive assessment of the diversity of βL homologues, a set of 132 non-redundant sequences covering at least one complete catalytic motif of the βL domain were selected for a phylogenetic analysis. The phylogenetic analysis was performed at four separate segments that include the ABC1 domain and regions surrounding SXXK (segment I), [SY]X[NT] (segment II), [K / H][T / S]G (segment III) of catalytic motifs of the βL domain (Fig. 1B). Based on the results of the analysis, a set of 57 sequences representing the highest diversity was defined (Table).

Phylogenetic analysis using the protein sequence parsimony method revealed several clusters for three segments of the βL domain, and four clusters for each segment were identified as reliable by bootstrapping analysis (Fig. 2). All 11 sequences containing a complete βL domain were represented within one cluster (cluster 1), except for segment III where a significant diversity (clusters 1–3) was identified among these sequences. Additional diversity was represented by sequences with partial βL domain coverage, which clustered into several significant clades.

Table. **Accession numbers and domain structure of representative protein sequences used in phylogenetic analysis**

Species	Accession[a]	Number of sequences[b]	E-value score[c]	Domains[d]
Bryophyta – Moss Superclass V – Bryopsida				
Physcomitrella patens	A9SIA0			■ 1 1 2
	A9U6U8			■ 4
	A9RTH6			■ ■
	A9SM18			■ 1 1 2
Tracheophyta – Spermatophyta – Gnetophyta				
Gnetum gnemon	PUT-163a-Gnetum_gnemon-3922	6193	1.2×10^{-7}	4
Tracheophyta – Spermatophyta – Coniferopsida				
Picea engelmanni × *Picea glauca*	PUT-157a-Picea_engelmannii_×_Picea_glauca-4983	13880	3.9×10^{-8}	■
Picea engelmannii × *Picea sitchensis*	PUT-155a-Picea_engelmannii_×_Picea_sitchensis-7665	27495	3.6×10^{-19}	4
	PUT-155a-Picea_engelmannii_×_Picea_sitchensis-6990		5.7×10^{-5}	■
Picea glauca	PUT-163a-Picea_glauca-17795	53255	2.0×10^{-4}	1 1 3
	PUT-163a-Picea_glauca-42544		1.4×10^{-35}	4 ■
	PUT-163a-Picea_glauca-15720		4.2×10^{-23}	4
	PUT-163a-Picea_glauca-21372		1.9×10^{-21}	3 4
	PUT-163a-Picea_glauca-10461		2.3×10^{-3}	4
Pinus taeda	PUT-157a-Pinus_taeda-5918	77540	4.2×10^{-37}	4 3
	PUT-157a-Pinus_taeda-14936		1.7×10^{-19}	■
	PUT-157a-Pinus_taeda-14658		1.7×10^{-14}	1 3
Tracheophyta – Magnoliophyta – Liliopsida				
Hordeum vulgare	Q6KB67			■
	PUT-161a-Hordeum_vulgare-35109815	102435	6.2×10^{-38}	■ 1 1 1
	PUT-161a-Hordeum_vulgare-79997		1.7×10^{-6}	■
	PUT-161a-Hordeum_vulgare-92236		1.7×10^{-5}	1
Oryza sativa	Q5Z8Y3			■ 1
	B8B250			■ 1 1
	PUT-157a-Oryza_sativa-043714	235277	9.0×10^{-36}	1 1
	PUT-157a-Oryza_sativa-0158191		2.4×10^{-11}	■
	PUT-155a-Oryza_sativa__japonica_cultivar_group_-120711	149282	2.7×10^{-4}	1
Saccharum officinarum	PUT-157a-Saccharum_officinarum-82059	131381	7.4×10^{-29}	■ 1
	PUT-157a-Saccharum_officinarum-17017		6.0×10^{-13}	■
	PUT-157a-Saccharum_officinarum-7818		9.9×10^{-6}	1

Table **(continued)**

Species	Accession[a]	Number of sequences[b]	E-value score[c]	Domains[d]
Tracheophyta – Magnoliophyta – eudicotyledons				
Arabidopsis thaliana	Q93YN5			
	PUT-157a-Arabidopsis_thaliana-81461	150533	3.0×10^{-20}	
Artemisia annua	PUT-163a-Artemisia_annua-23935	24963	7.7×10^{-44}	
	PUT-163a-Artemisia_annua-24765		2.7×10^{-36}	
	PUT-163a-Artemisia_annua-15193		4.0×10^{-38}	
	PUT-163a-Artemisia_annua-23957		2.7×10^{-26}	
	PUT-163a-Artemisia_annua-62		3.5×10^{-14}	
	PUT-163a-Artemisia_annua-13455		3.8×10^{-8}	
Citrus clementina	PUT-157a-Citrus_clementina-33359	37350	7.6×10^{-34}	
	PUT-157a-Citrus_clementina-21997		1.7×10^{-33}	
Citrus sinensis	PUT-157a-Citrus_sinensis-14434	27239	2.7×10^{-4}	
Lactuca virosa	PUT-157a-Lactuca_virosa-4209	12839	6.6×10^{-37}	
	PUT-157a-Lactuca_virosa-6956		3.4×10^{-40}	
	PUT-157a-Lactuca_virosa-6690		6.6×10^{-37}	
	PUT-157a-Lactuca_virosa-11139		3.6×10^{-32}	
Lotus corniculatus	PUT-151a-Lotus_corniculatus-6993	39940	9.5×10^{-15}	
Poncirus trifoliata	PUT-163a-Poncirus_trifoliata-2895	11204	6.0×10^{-40}	
	PUT-163a-Poncirus_trifoliata-4961		1.6×10^{-25}	
Populus trichocarpa	B9N6B3			
	B9NA19			
Ricinus communis	B9TKN7			
	B9TEE0			
	B9SPR5			
	B9TP52			
	B9TP64			
	B9TB72			
	B9T9A0			
Vitis vinifera	PUT-157a-Vitis_vinifera-2851297	47846	2.5×10^{-38}	
	PUT-157a-Vitis_vinifera-316		4.6×10^{-23}	

[a] SwissProt or PlantGDB PUT id; [b] the number of DNA sequences in the PlantGDB PUT database for indicated species; [c] E-value score for search using the statistical model of the beta-lactamase family domain; the score based on the number of sequences in the PlantGDB PUT database for indicated species; [d] presence, absence and partial sequences of the ABC1 domain and I–III segments of the βL domain (consecutively) are represented by black, white and grey rectangles, respectively; numbers indicate a phylogenetic group.

It is remarkable that several sequences originating from one species were assigned to several different clusters. This diversity could point to different classes of βLs with a different structure and function. For three segments, such species include *P. trifoliata* and *P. glauca* (sequences in cluster 1 and 3). For segments I and II, *R. communis* sequences fall into clusters 1 and 3, and clusters 1, 2 and 3, respectively. *Physcomitrella patens* sequences are assigned to clusters 1 and 4 of segment II. On the other hand, several unique sequences originating from one species (such as *R. communis*, *A. annua*, *L. virosa*) were assigned to a single cluster indicating the cases of intraspecific gene polymorphism.

The application of the parametric method of maximum likelihood analysis to assess the phylogenetic distance among the identified clusters came out unsuccessful. Although a similar overall clustering of the sequences was obtained as compared to the sequence parsimony method, the analysis resulted in a poor robustness of the inferred trees (data not shown). The reason could be the irrelevance of the parametric model for the regions used in the analysis, and an accurate application of the model would require using parameters inferred from functional and structural data.

Phylogeny of the ABC1 family domain (PF03109) was assessed using 197 *Embryophyta* sequences available from Pfam database, including eight sequences representing the multimodular proteins containing βL domain (Table). Two additional sequences containing ABC1 domain identified in the PUT database search using beta-lactamase family domain model were included in the analysis (Table). The

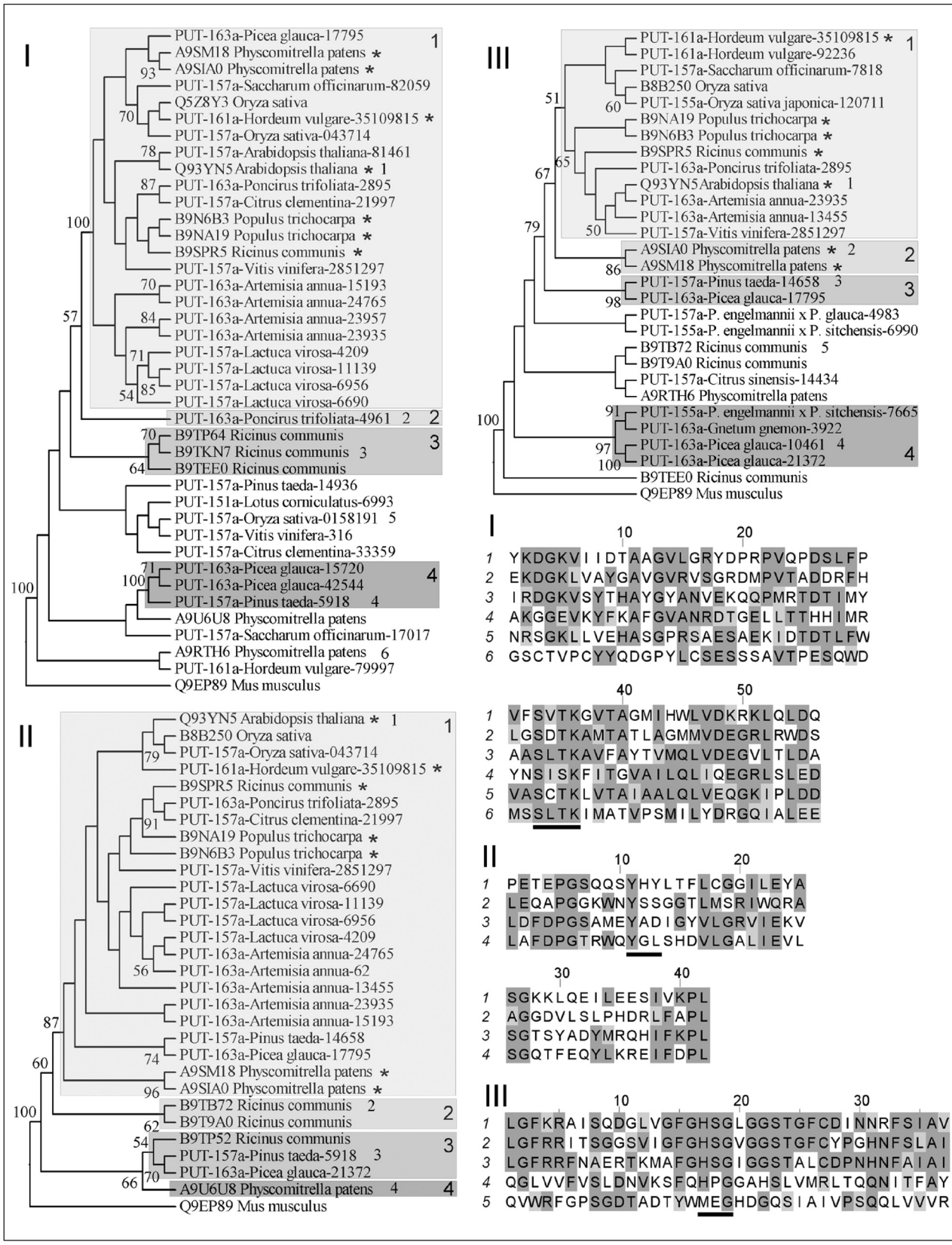

Fig. 2. Phylogenetic tree and alignment of sequences of segments I, II and III of the βL domain characteristic of plant βL homologues. Sequence alignment was built using the HMM model of the Pfam beta-lactamase family domain. The mouse LACTB sequence (Q9EP89) was used to root the phylogenetic tree. Phylogenetic groups are indicated as grey rectangles. Sequences including the ABC1 domain are labelled with a star. Typical sequences of the group shown in the alignment are labelled by numbers. A sequence alignment is coloured depending on the BLOSUM62 score. Lines indicate active centre motifs

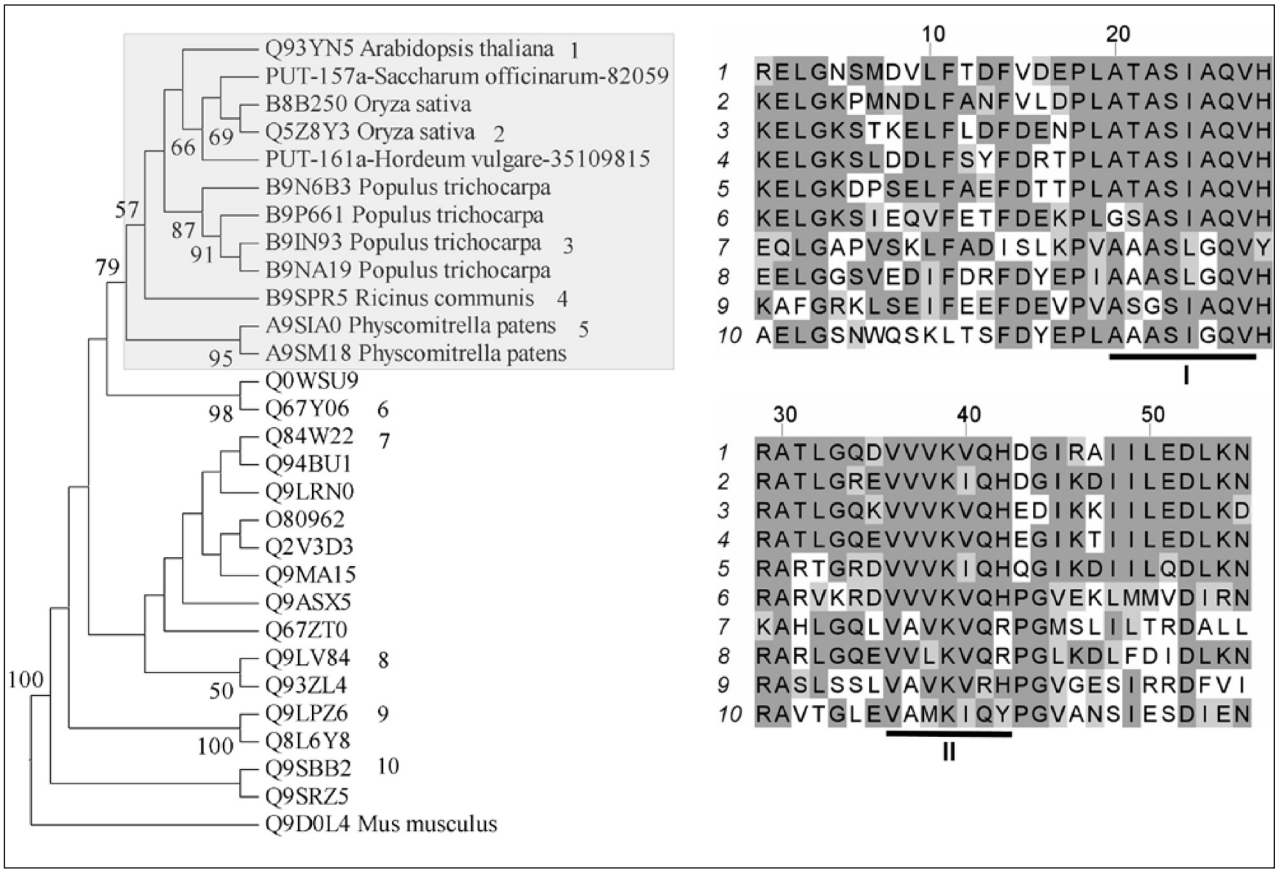

Fig. 3. Phylogenetic tree and alignment of typical sequences of ABC1 family domain. Sequence alignment was built using the HMM model of the Pfam ABC1 family domain. The mouse ABC1 sequence (Q9D0L4) was used to root the phylogenetic tree. The phylogenetic group including multimodular βL homologues is indicated as a grey rectangle. Only 16 representative *A. thaliana* sequences are shown outside the multimodular βL homologue group. Typical sequences shown in the alignment are labelled by numbers. Sequence alignment is colored depending on the BLOSUM62 score. Active centre motifs involved in ATP binding (I and II) are indicated

results of protein sequence parsimony analysis presented in Fig. 3 demonstrated that the sequences of the multimodular proteins containing a βL domain clustered within one characteristic group differing from other ABC1 family proteins. There were only two sequences of non-multimodular proteins included in the group. These were the predicted protein sequences of *P. trichocarpa* (B9P661 and B9IN93) that contained no βL domain. The sequences might represent paralogues that contain no βL domain, or theoretical models of the gene. Therefore it is likely that the ABC1 domain of the multimodular βL homologues forms a unique group differing from the other proteins of the ABC1 family.

In conclusion, the analysis of plant βL homologue protein sequences revealed a phylogenetic diversity which implies the structural and functional uniqueness of the protein family in plants. At least four distinct phylogenetic groups were identified for each region, corresponding to catalytic segments of the βL domain. Sequences of the ABC1 domain are closely related among the multimodular proteins homologous to βL and constitute a distinct phylogenetic group within the ABC1 protein family in plants. The further refinement of phylogenetic data requires additional sources of complete sequence information on the plant proteins and the availability of structural and functional information that would provide the means required for an efficient application of the parametric methods of phylogenetic analysis.

ACKNOWLEDGEMENT

The publication has been supported by the project VP1-3.1-ŠMM-05-K-01-022 "Strengthening the potential of the Lithuanian Biochemical Society in developing knowledge society".

References

1. Goffin C, Ghuysen JM. Microbiol Mol Biol Rev 1998; 62(4): 1079–2172.
2. Knox JR, Moews PC. J Mol Biol 1991; 220(2): 435–55.
3. Liobikas J, Polianskyte Z, Speer O et al. Prot Expr Purif 2006; 45(2): 335–42.
4. Smith TS, Southan C, Ellington K et al. Genomics 2001; 78(1–2): 12–4.
5. Liobikas J, Baniulis D, Stanys V et al. Biologija 2006; 3: 126–9.
6. Peitsaro N, Polianskyte Z, Tuimala J et al. BMC Evol Biol 2008; 8: 26.
7. Polianskyte Z, Peitsaro N, Dapkunas A et al. Proc Nat Acad Sci U.S.A. 2009; 106(45): 18960–5.
8. Limphong P, Nimako G, Thomas PW et al. Biochemistry 2009; 48(36): 8491–3.
9. Bousquet I, Dujardin G, Slonimski PP. EMBO J 1991; 10(8): 2023–31.
10. Bateman A, Birney E, Cerruti L et al. Nucl Acids Res 2002; 30(1): 276–80.
11. Dong Q, Lawrence CJ, Schlueter SD et al. Plant Physiol 2005; 139(2): 610–8.
12. Rice P, Longden I, Bleasby A. Trends Genet 2000; 16(6): 276–7.
13. Eddy SR. Bioinformatics 1998; 14(9): 755–63.
14. Castresana J. Mol Biol Evol 2000; 17(4): 540–52.
15. Felsenstein J. Cladistics 1989; 5: 164–6.
16. Jones DT, Taylor WR, Thornton JM. Comp Appl Biosci 1992; 8(3): 275–82.
17. Zhang Y. BMC Bioinformatics 2008; 9: 40.
18. Potterton E, McNicholas S, Krissinel E et al. Acta Crystallogr D Biol Crystallogr 2002; 58(Pt 11): 1955–7.
19. Macheboeuf P, Contreras-Martel C, Job V et al. FEMS Microbiol Rev 2006; 30(5): 673–91.

Danas Baniulis, Julius Liobikas, Subhash Mohan Agarwal, Birutė Frercks, Gintautas Vaitiekaitis, Vidmantas Stanys

AUGALŲ β-LAKTAMAZĖS ŠEIMOS BALTYMŲ STRUKTŪROS DOMENŲ FILOGENETINĖ ANALIZĖ

Santrauka

Bakterijų serino β-laktamazės priskiriamos peniciliną sujungiančių baltymų klasei, kuriai priklausantys baltymai skaido β-laktamus ir apsaugo peptidoglikano sintezės fermentus nuo inhibicijos. β-laktamazei homologiški genai buvo rasti *Protozoa*, *Metazoa* ir *Plantae*, tačiau nebakterinės kilmės genų funkcija iki šiol nėra nustatyta. Struktūros ir filogenetiniai tyrimai atskleidė galimos funkcijos identifikavimui svarbius aspektus. Nebakterinės kilmės β-laktamazei homologiškų baltymų aminorūgščių sekoje rastas β-laktamazės aktyviam centrui būdingas trijų sekos motyvų darinys. Taip pat nustatyta, kad augalų β-laktamazės homologas yra daugianaris baltymas, kurio polipeptido grandinės amino gale yra ABC1 domenas. β-laktamazės baltymo funkcijos eksperimentinius tyrimus tikslingai nukreipia aminorūgščių sekų įvairovės charakterizavimas. Keturi aloparalogų kamienai nustatyti tarp *Metazoa* β-laktamazės homologų. Siekiant charakterizuoti augalų β-laktamazei homologiškų baltymų įvairovę, buvo panaudota funkcinių domenų filogenetinė analizė. Grupė 57-ių aminorūgščių sekų, kurios bent iš dalies atitinka struktūrinius domenus, buvo surinkta iš genomo ar sujungtų sekų raiškos žymių duomenų bazių. Filogenetinė įvairovė įvertinta panaudojant ABC1 domeno ir trijų sričių, atitinkančių β-laktamazės aktyvaus centro motyvus SXXK, [SY]X[NT], [K / H][T / S]G, sekas. Filogenetinė analizė atskleidė, kad kiekvienos β-laktamazės aktyvaus centro motyvo srities sekoms būdingos bent keturios skirtingos filogenetinės grupės. Tai rodo augalų β-laktamazės homologų struktūrinę ir funkcinę įvairovę. Taip pat nustatyta, kad daugianarių β-laktamazei homologiškų baltymų ABC1 domeno sekos yra panašios ir sudaro filogenetinę grupę, skirtingą nei kiti ABC1 šeimos baltymai.

Raktažodžiai: serino β-laktamazė, LACTB, ABC1, sekos lyginimas, struktūros domenai, filogenetinė analizė