Bioinformation
www.bioinformation.net
open access
Hypothesis

Characterization of allergenic epitopes of Ory s1 protein from *Oryza sativa* and its homologs

Ruchi Sharma[1,*], Ashok Kumar Singh[1], Vetrivel Umashankar[2]

[1]Department of Botany, Udaya Pratap College, Varanasi, Uttar Pradesh, India; [2]Department of Bioinformatics, School of Biosciences, SRM University, Ramapuram, Chennai, Tamil Nadu, India; Ruchi Sharma - Email: ruchivns@rediffmail.com; * Corresponding author

Received March 16, 2009; Revised May 04, 2009; Accepted June 17, 2009; Published August 18, 2009

Abstract:
Vaccination is the most effective technique suggested now days for allergy treatment. Recombinant-based approaches are mostly focused on genetic modification of allergens to produce molecules with reduced allergenic activity and conserved antigenicity. The molecules developed for vaccination in allergy possess significantly reduced allergenicity in terms of IgE binding, and therefore will not lead to anaphylactic reactions upon injection. This approach is probably feasible with every peptide allergen with known amino acid sequence. In this study an *in silico* approach was used to investigate allergenic protein sequences. Motif analysis of these sequences reveals the allergenic epitopes in the amino acid sequences. Physicochemical analysis of protein sequences shows that the homolog allergens of *Ory s1* are highly correlated with the aromaticity, GRAVY and cysteine content. Moreover, phylogenetic analysis of Ory s1 with other sequences reveals that *Oryza sativa japonica* and *Zea mays* are close homologs, whilst *Lolium perenne* and *Dactylis glomerata* are found to be remote homologs. The multiple sequence alignment reveals of Ory s1 with all its homologs in this study reveals the high conservation of residues in DPBB_1 domain (amino acid residue positions 86- 164) and was found distinctly in all the sequences. These findings support the proposal that allergenic epitopes encompass conserved residues. The consensus allergenic was found to be mainly composed of hydrophobic residues. The functional sites of allergenic proteins reported in this study shall be attenuated to develop hypoallergenic vaccine. The sequence comparison strategy adopted in this study would pave way effective evolutionary analysis of these allergens.

Keywords: Ory s1, sequence analysis, physicochemical analysis, allergenic epitopes, phylogenetic analysis.

Background:
The Ory s1 protein from *Oryza sativa* has been studied extensively to gain a better understanding of its remarkable allergenicity **[1]**. Pollen allergens of *Oryza sativa* is recognized by the International Union of Immunological Societies (IUIS) official list of allergens which include Ory s1, Ory s7, and Ory s12. Much information has been published focusing on the physicochemical and epitope analysis of the allergenic pollen proteins **[2, 3, 4].** Majority of the world's population depend on rice, wheat, maize for daily sustenance. These provide important models for evolutionary studies of the grasses since various aspects of their biology have been well documented **[5].** The present investigation focuses on sequence and epitope analysis of Ory s1 and its homologs. Chemical modification of allergen vaccines to reduce IgE binding improves safety while maintaining clinical efficacy. Analysis of molecular size and allergen content may be useful techniques for characterization and standardization of allergoid products **[6]**. Identification of potentially allergenic proteins is needed for the safety assessment of genetically modified foods, certain pharmaceuticals and various other products on the consumer market. Features that differentiate allergens from non-allergens are difficult to find by manual inspection of amino acid sequences. Current methods in bioinformatics allergology exploit common features among allergens for the detection of amino acid sequences of potentially allergenic proteins. Features for identification still unexplored include the motifs occurring commonly in allergens, but rarely in ordinary proteins **[7]**. In addition to laboratory experimentation and clinical testing, current procedures for allergenicity assessment involve an introductory comparison of the novel protein's amino acid sequence with those of known allergens **[8]**. Several regions of the amino acid sequence of the homologs are evolutionarily highly conserved. Highly conserved primary sequences of allergenic homologs have been used in an attempt to establish evolutionary relationships. The studies of different allergen protein sequences suggest that allergens tend to share certain sequence similarities. Thus, the potential allergenicity of query proteins can be predicted by examining their sequence similarities with known allergens **[9]**. The application of computational techniques in biological discovery was possible due to the availability of extensive sequence data. The most widely-used and conceptually easiest to understand of these techniques is database homology searching, where sequence similarity can be used to assign target for hypoallergenic vaccine production. A recent paper showed the efficacy of recombinant birch pollen vaccine for the treatment of birch-allergic rhinoconjunctivitis **[10]**. The similarities found between the homologous sequences argue that there should be similarities in their three-dimensional structures, strengthening the hypothesis that proteins with similar sequences perform a similar function. The main objective of this study is to analyze the comparative abundance and distribution of allergenic epitopes in the sequences and to help in identifying target amino acid positions in the course of vaccine development.

Methodology:
Sequence retrieval:
All databases and software used in these studies are publicly available on the world-wide web. The primary sequence of Ory s1 from *Oryza sativa* was acquired from the NCBI's GenPept, a publicly available database **[11]**. BLAST (psi blast) search, using the non redundant database, was performed that resulted homologous sequences. Twenty homologs from distant organisms were selected and the sequences were acquired from GenPept. This provided data required to predict the primary structure (sequence) and to perform the allergenicity assessment study.

Physiochemical analysis:
Physiochemical analysis of the sequences namely molecular weight, theoretical pI, amino acid composition, instability index, aliphatic index and grand average of hydropathicity (GRAVY) were done using PROTPARAM tool **[12]**.

Phylogenetic analysis:
Clustal W analysis software available online from Moscow State University's A. N. Belozersky Institute of Physio-Chemical Biology was used to compare sequence alignment of allergenic homologs with the default settings **[13]**. Here the dendrogram is calculated in 2 stages: first all pairs of sequence are compared using of Wilbur and Lipman method **[14]** and then the similarity scores resulted are used to

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 4(1): 12-18 (2009)
© 2009 Biomedical Informatics

construct the dendrogram using the UPGMA cluster analysis method of Sneath and Sokal **[15]**. Tree construction was done by using Phylodraw **[16]**.

Allergenic domain detection:
Allergenic domains of the sequences were extracted using ProScan **[17]**. Motif search in the multiple sequence alignment was carried out using Multiple Em for Motif Elicitation (MEME tool) technique **[18]**. Antigenic sites on proteins was discovered, using EMBOSS antigenic program **[19]**, a semi-empirical method which makes use of physicochemical properties of amino acid residues and their frequencies of occurrence in experimentally known segmental epitopes **[20].**

Discussion:
Sequence from NCBI server was retrieved for *Oryza sativa* allergenic protein Ory S1 A86533.1. BLAST search (Psi BLAST) obtained homologous Sequences, and 20 significant homologous sequences were short listed from diverse species varying from grasses to higher plants for comparative study: AAA86533.1, NP_001048686.1, CAA81613.1, CAA10520.1, CAA10140.1, CAB63699.1, AAP96760.1, AAS48882.1, ABF81662.1, CAC40805.1, *ABB83474.1*, AAZ08315.1, NP_190182.2, ABK93417.1, AAV85475.1, ACB45301.1, ABJ90221.1, AAT11859.2, BAC67192.1, BAC66787.1. Amino acid residues Ala (7.6-11.9), Cys (2.70-3.80), His (1.0-2.2), Leu (4.6-6.7), Trp (2.2-3.5), Tyr (2.3-4.5), Val (6.2-9.8) are found quite equally distributed among all the sequence than other amino acids (**Table 1),** while Cys and Val residues are predominant found in allergenic epitopes **(Table 4). Table 4** shows most of the hydrophobic residues in allergenic site. **Table 2** provides details of the physiochemical analysis which shows all the sequences as stable (instability index ranging from 17.35 to 44.02) with a theoretical pI ranging from 5.32 to 9.58. Though GRAVY was found negative (-.003 to -0.643) for complete sequences the distribution of sequences in alleregnic motif was predominantly hydrophobic **(Table 1).** The information content diagram provides an idea of the positions in the motif that are most highly conserved. It is very interesting to note the simplified position specific matrix result where the four conserved cystiene residues was found in all the sequences at the same position **(Table 3).** Interestingly the phylogenetic tree of Ory s1 homologs included several apparently eukaryotic orthologs. Therefore it seems most likely that the progenitors of each of these orthologous sets might also cause allergy. The phylogenetic analysis has resulted that pollen allergen sequence from *Oryza sativa japonica* (NP_001048686.1) was most closely related sequence and EXPB10 sequence from *Zea mays* (ABF81662.1) as next closely related sequence with the query sequence Ory s1 from *Oryza sativa indica* (AAA86533.1). Beta expansin B2 from *Festuca pratensis* (CAC40805.1) and Beta-expansin EXPB4 from Hordeum vulgare (ACB45301.1) were next most closely related sequences found. Expansin 1 from *Mangifera indica* (AAT11859.2) and sequence from *Eucalyptus globulus* (AAZ08315.1) were found less related. Pollen allergen from *Lolium perenne* (CAB63699.1) and Group 1 allergen Dac g 1.02 precursors from *Dactylis glomerata* (AAP96760.1) were found as distantely related. It shows that the relation ship of the sequences in phylogenetic analysis **(Figure 1).**

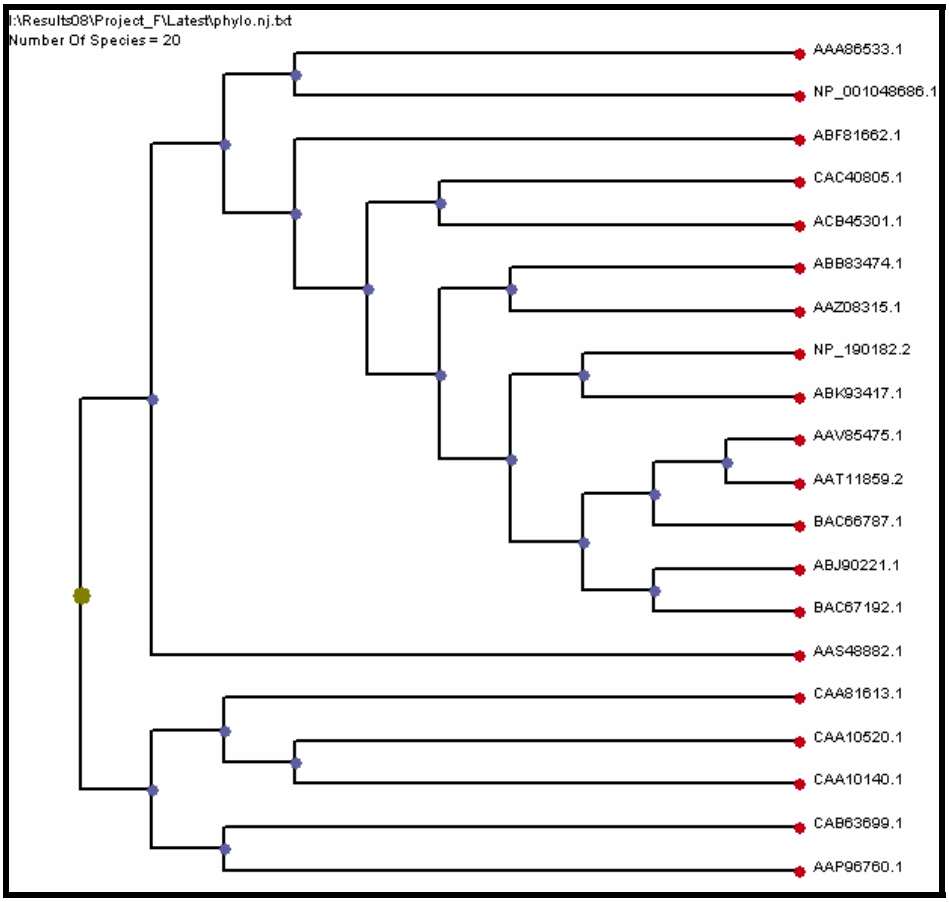

Figure 1: Figure showing phylogenetic relationship of Ory s1 protein sequence with other homologs.

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 4(1): 12-18 (2009)
© 2009 Biomedical Informatics

Conclusion:
Consensus epitope identification using the accessible allergenic region has geared up the pace. If bioinformatics approaches are standardized and optimized, it can be used for the swift identification of potential antigenic regions to target allergenic proteins in course of development of hypoallergenic vaccines. Allergenic epitopes in this study show abundance of hydrophobic residues in the motif. The role of highly conserved cysteines residues at four positions shall also play a major role in determining the allergenicity. Cystiene residues were found highly conserved for the motif 1 with a width of 29 and 20 sites. It was found in all the sequences at 9, 21, 24, 29 positions as shown in the information content diagram **[21].** This is in order with the results documented in table 3. The function of disulfide bonds formed between cysteines in IgE binding has been investigated in studies with several other allergens this falls in line with the previously documented studies **[22, 23]**. Hence, the procured consensus region shall be utilized for effective vaccine design against food allergens.

References:

[1] D Scott *et al., Molecular Plant* **1:**751 (2008) [PMID: 2660330]
[2] F Takaiwa *et al, Immunol. Allergy Clin. North Am*. **27:**129 (2007) [PMID: 17276883]
[3] JS Ye *et al, DNA Research* **12:**167 (2005) [PMID: 16303748]
[4] MM Sen *et al., J. Environ. Monit.* **5:**959 (2003) [PMID: 14710939]
[5] P Rajendrakumar *et al., In Silico Biology* **8:** 9 (2008) [PMID: 18928198]
[6] J Carnes *et al., Clin. Exp. Allergy* **39:**426 (2009) [PMID:19134021]
[7] K Bjorklund *et al., Bioinformatics* **21:**39 (2005) [PMID: 15319257]
[8] W Kong *et al., In Silico Biology* **7:**77 (2006) [PMID: 17688432]
[9] GS Ladics, MK Selgrade, *Regul Toxicol Pharmacol.* **54:**S2. (2009) [PMID: 19028539]
[10] G Pauli *et al, J. Allergy Clin. Immunol.* **122:**951 (2008) [PMID: 19000581]
[11] http://blast.ncbi.nlm.nih.gov
[12] http://www.expasy.ch/tools/protparam.html
[13] http://www.genebee.msu.su/clustal/
[14] WJ Wilbur, DJ Lipman, *Acad. Sci. Vold*., **80:**726 (1983)
[15] PHA Sneath, RR Sokal, *Numerical Taxonomy: Freeman*, San Francisco (1973)
[16] http://pearl.cs.pusan.ac.kr/phylodraw
[17] http://npsa-pbil.ibcp.fr/
[18] http://meme.sdsc.edu/meme/meme-intro.html
[19] http://inn.weizmann.ac.il/cgibin/EMBOSS/
[20] AS Kolaskar, PC Tongaonkar, *FEBS Lett.* **276:**172 (1990) [PMID: 1702393]
[21] R Sharma, AK Singh, *Am. J. Infectious Diseases* **5:**149 (2009)
[22] M Lombardero *et al., J. Immunol.* **144:**1353 (1990) [PMID: 1689351]
[23] S Olsson *et al., Mol. Immunol.* **35:**1017 (1998) [PMID: 10068036]

Edited by P. Kangueane
Citation: Sharma *et al,* Bioinformation 4(1): 12-18 (2009)
License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 4(1): 12-18 (2009)
© 2009 Biomedical Informatics

Supplementary material

Table 1: Table showing amino acid composition table (%) for the residues found most conserved in the homologs

S. No	Organism Name	Ala (A)	Cys (C)	His(H)	Leu (L)	Val(V)	Trp(W)	Tyr (Y)
1	AAA86533.1	9.90	3.00	1.50	5.70	7.60	1.50	2.30
2	NP_001048686.1	10.60	3.40	1.10	5.30	6.80	2.30	3.00
3	CAA81613.1	8.70	2.70	1.50	4.60	8.00	2.30	3.00
4	CAA10520.1	8.40	2.70	1.50	5.70	7.20	2.30	3.40
5	CAA10140.1	8.00	2.70	1.50	6.10	8.00	2.30	3.40
6	CAB63699.1	8.00	2.70	1.50	4.90	7.20	2.30	3.40
7	AAP96760.1	7.60	2.70	1.90	5.30	7.60	2.30	3.40
8	AAS48882.1	7.80	3.00	2.20	4.40	7.80	2.20	3.30
9	ABF81662.1	7.80	3.30	1.10	5.20	8.60	2.60	4.50
10	CAC40805.1	10.40	6.30	2.20	5.90	5.90	2.20	4.10
11	ABB83474.1	8.00	3.60	1.10	6.50	8.00	2.20	3.30
12	AAZ08315.1	11.90	3.80	1.00	6.70	8.60	1.90	2.90
13	NP_190182.2	8.80	3.30	1.40	5.60	9.80	2.30	4.20
14	ABK93417.1	9.30	4.20	1.20	5.80	7.30	2.30	3.50
15	AAV85475.1	8.50	3.10	1.60	6.60	7.00	3.10	3.50
16	ACB45301.1	10.30	3.30	2.20	6.20	6.20	2.60	4.80
17	ABJ90221.1	10.30	3.20	2.40	5.50	6.70	2.80	4.70
18	AAT11859.2	8.80	3.80	1.20	5.40	6.90	3.50	3.50
19	BAC67192.1	9.90	3.20	2.40	5.90	7.10	2.80	4.70
20	BAC66787.1	9.20	3.50	1.20	5.40	8.10	3.50	3.10

Table 2: Table showing physiochemical properties of Ory s1 and its homologous sequences

SN	Accession No.	MW	NCR	PCR	TP	II	AI	G
1	AAA86533.1	28497.70	34	38	8.53	44.02	73.46	-0.41
2	NP_001048686.1	28351.40	36	35	6.34	31.01	73.22	-0.33
3	CAA81613.1	28203.10	33	34	7.53	25.33	68.97	-0.32
4	CAA10520.1	28246.00	35	34	6.46	25.62	66.39	-0.393
5	CAA10140.1	28347.20	36	35	6.46	22.07	69.70	-0.388
6	CAB63699.1	28360.90	37	33	5.69	25.47	64.52	-0.454
7	AAP96760.1	28468.10	36	33	6.01	25.04	65.00	-0.446
8	AAS48882.1	29026.90	32	34	8.02	28.96	65.00	-0.423
9	ABF81662.1	29109.40	30	33	8.25	17.35	73.20	-0.249
10	CAC40805.1	29307.00	22	23	7.60	36.29	69.70	-0.219
11	ABB83474.1	28983.50	19	19	6.79	42.20	68.07	-0.106
12	AAZ08315.1	21755.40	17	14	5.32	34.43	72.05	-0.042
13	NP_190182.2	23391.70	16	24	9.14	31.65	80.70	-0.148
14	ABK93417.1	28470.30	25	28	8.25	38.92	74.21	-0.147
15	AAV85475.1	27917.60	11	20	9.34	35.85	71.09	-0.056
16	ACB45301.1	30067.30	17	31	9.49	34.23	64.03	-0.358

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 4(1): 12-18 (2009)
© 2009 Biomedical Informatics

17	ABJ90221.1	26800.00	6	17	9.48	35.50	60.59	-0.163
18	AAT11859.2	28333.40	11	23	9.48	40.16	67.92	0.003
19	BAC67192.1	26755.00	6	16	9.38	35.28	62.89	-0.125
20	BAC66787.1	28019.80	10	22	9.58	33.98	68.65	-0.008

SN = S. No.; MW = molecular weight; NCR = negatively charged residues; PCR = positively charged residues; TP = Theoretical pI; II = Instability index; AI = Aliphatic index; G = Gravity

Table 3: Simplified position-specific probability matrix for motif 1

		Simplified position-specific probability matrix																														
S. No.	AA																															
1	A	4			1				5	5	1			3						1				4							
2	C									6													a		a					a	
3	D												1					5													
4	E												1															7			
5	F		1	7												8				1						7					
6	G	1	3		3	5		1	3		6	1							a		8		a								
7	H				1		1																								
8	I							1				1	1		5													8			
9	K																8			3										7	
10	L				1		1	1			3				5					3								1			
11	M				1		6													1											
12	N			1	1		2	1				6					3	3													
13	P	1	7										1	6																	
14	Q	3											1														3				
15	R																			3						1		3			
16	S	1			3	3	1		3			3	1	2				3		1	1			6							
17	T					1		7					5																		
18	V					1		1				1	1															1			
18	W																														
20	Y	1		3												1									3						
	CN	C	A	P	F	S	G	M	T	A	C	G	N	T	P	I	F	K	D	G	R	G	C	G	S	C	F	E	I	K	C

AA = amino-acids; CN = consensus

Table 4: Table showing five major allergenic epitopes in protein homologs of Ory s 1 using EMBOSS antigenic program

S. No.	ID	Epitope sequence & position	Site 1	Site 2	Site 3	Site 4	Site 5
1	AAA86533.1	Sequence	SSLLLACVVVAAMVSPSPAGHPKVPPG	PNYLALLVKYVAGDGDVVEVE	RVQVNV	CGSCFEIKCPEACSDKPALIHVT	FRRVKCKYP
		Sequence position	4-30	177-197	255-260	92-116	154-162
2	NP_001048686.1	Sequence	SSLLLACVVVAAMVSAVSCGPPKVPPG	PNYLALLVKYVAGDGDVVEVE	CGSCFEIKCSKPEACSDKPALIHVT	FRRVKCKYP	PKPLKGPFSVRVT
		Sequence position	4-30	177-197	92-116	154-162	222-234
3	CAA81613.1	Sequence	SSSVLLVVALFAVFLGSAHGIPKVPPG	CGSCFEIKCTKPEACSGEPVVVHIT	NYLALLVKFVAGDGDVVAVDI	EVEIQFRRVKCKYP	VTFHVE
		Sequence	4-30	92-116	181-201	152-165	170-175

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 4(1): 12-18 (2009)
© 2009 Biomedical Informatics

		position					
4	CAA 10520 .1	Sequence	SSSVLLVVALFA VFLGTAHGIAK VPPG	NYLALLVKYV	CGSCFEIKCTK PESCSGEPVL VHIT	DGDVVAVDI	VTFHVE
		Sequence position	4-30	181-190	92-116	193-201	170-175
5	CAA 10140 .1	Sequence	SSLVLLVVALF AVFLGTAHGIA KVPPG	NYLALLVKYV	DGDVVAVDI	CGSCFEIKCT KPESCSGEPI VVHIT	IAAYHLDLS GK
		Sequence position	4-30	181-190	193-201	92-116	123-133
6	CAB 63699 .1	Sequence	SSSVLLVVALFA VFLGSAHGIAK VPPG	PNYLAILVKYV	DGDVVAVDI	CGSCFEIKCT KPESCSGEAV TVT	ELELQFRRV KCKYP
		Sequence position	4->30	180->190	193->201	92->114	152->165
7	AAP 96760 .1	Sequence	SSSVLLVVALFA VFLGSAHGIPKV PPG	NYLALLVKYV	DGDVVAVDI	CGSCFEIKCT KPESC SGEAVTVHIT	ELELQFRRV KCKYP
		Sequence position	5-31	182-191	194-202	93-117	153-166
8	AAS 48882 .1	Sequence	SSSVLLVAAVL AAVVCG AHGIAKVPPG	YLALVIKFL	KTVVDDVIPK	DGDVVGVDI K	CGSCFELKC TKPEACSG
		Sequence position	5-31	186-194	247-256	197-206	96-112
9	ABF 81662 .1	Sequence	VSIMWSLVQVQ VLVAVAL SFLVGGAWCGP PKVPPG	CGNVPIFKDGLGC GSCFEIKCDKPAE CSGKPVVVYIT	PNYLALLVKY V	DIVAVDI	KTVYDDVIP T
		Sequence position	4-38	88-124	186-196	201-207	249-258
10	CAC 40805 .1	Sequence	HSGIIDIQFRRVP CNFPGLKINFHV VDGSNAVYLAV LIE	VLSVKVAALAGL IFSVLAASAAAK	CGSCYQIRCS	YYPVAQYHF DLSGT	CGFKHVNQ YPFS
		Sequence position	152-90	4-27	95-104	123-136	67-78
11	ABB 83474 .1	Sequence	GCGACYQVKCT	LNPSIVFTYFTTFS LLTITCSCLHPKR F	SGKPVRVVIT DSCPGGPCLS ESAHFDLSGT	GSSLVAPFSL KLT	NAGVIQIQY KRVECNYP GVKLTFHV DS
		Sequence position	95-105	8-35	111-140	232-244	156-182
12	AAZ 08315 .1	Sequence	GCGACYQVKCT ENAA	SGNPVTVVITDE	YRSVVN	PGGPCVAES AHFDLSGT	NAGVLQIQY QKVKCNFP GAKVAFHV DS
		Sequence position	29-43	45-56	202-207	58-74	90-116
13	NP_1 90 182.2	Sequence	GCGACFQVRCK NPKLCNSKGTIV MVT	AKPVVGVDKYLL KQGIVDVEYQRV PCNY	PNYLAIKLLY QG	TEVVGIDIAP VGSS	TDLVLSSR
		Sequence position	23-48	69-96	112-123	126-139	56-63
14	ABK 93417 .1	Sequence	FICFLFLAISYAT ACDRCVHQSKV AYFSRDSALSSG ACGYG	YLAIKLLYQG	CGACFQIRCK DTTLCSR	TEVVAIDFAK VG	GHLAAAVSS LYK
		Sequence position	4-44	157-166	69-85	169-180	53-64
15	AAV 85475 .1	Sequence	ATISFISLVLLLS LVEARIPGVYT G	AGDLVKVSVK	NSVLVGQSLS FRVT	SCGACFEIKC ANEPQWCHS GSPSIFIT	YFNLVLITN V
		Sequence position	5-29	186-195	215-228	80-106	174-183

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 4(1): 12-18 (2009)
© 2009 Biomedical Informatics

16	ACB45301.1	Sequence	SANAVALAALVSVLLTYGCCAQSPLNYTSLAKAS	NYLAVLVEF	CGACYRIRCVNNK	KTLVAKQVIPA	YYPVAKYHFDLSGT	
		Sequence position	5-38	183-191	95-107	248-258	123-136	
17	ABJ90221.1	Sequence	NPPQHHFDLSQPVFQHIAQYKAGVVPVSYRRVPCR	GCGSCYEIRCVNDPKWCLPGTIAVTATNFC	AGDVHSVSVK	YFNLVLITNV	MAGLLAMLVASAHAY	
		Sequence position	122->156	78->107	182->191	170->179	11->25	
18	AAT11859.2	Sequence	VGLSMACILSLRCLMWM	YFNLVLIADVA	NTVLVGQALSFRVR	SGDIVKVSVK	PMFLKIAEYRAGIVPVSYRRVPCR	
		Sequence position	4-20	176-186	217-230	188-197	139-162	
19	BAC67192.1	Sequence	NPPQHHFDLSQPVFQHIAQYKAGVVPVSYRRVPCM	GCGSCYEIRCVNDPKWCLPGSIVVTATNFC	AGDVHSVSVK	YFNLVLITNV	MAGLLAMLVASAHAY	
		Sequence position	122-156	78-107	182-191	170-179	11-25	
20	BAC66787.1	Sequence	VSVSACLASLLISLMWV	YFNLVLVSNVA	NAVLVGQSLSFRV	AGDIVRVSVKGS	SCGACFEIKCA	
		Sequence position	4-20	176-186	217-229	188-199	82-92	

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 4(1): 12-18 (2009)

© 2009 Biomedical Informatics