

Appendix I

Mascot search results

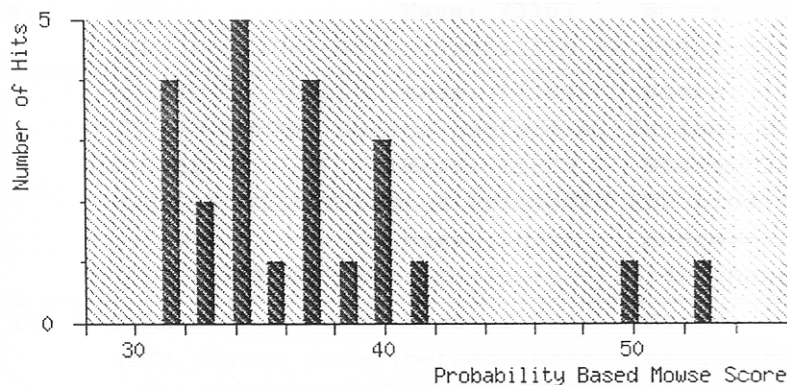
ของเมล็ดโปรตีนที่ 1

MATRIX *SCIENCE* Mascot Search Results

User : Waraporn
 Email : waraporn.p@psu.ac.th
 Search title :
 Database : SwissProt 46.6 (232847 sequences; 105413441 residues)
 Timestamp : 13 May 2005 at 11:57:03 GMT
 Top Score : 53 for Q9NH48, (KARG_ERISI) Arginine kinase (EC 2.7.3.3) (AK) Argin:

Probability Based Mowse Score

Ions score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event. Protein scores greater than 66 are significant ($p < 0.05$).



Concise Protein Summary Report

Format As

Concise Protein Summary

[Help](#)Significance threshold $p < 0.05$

Max. number of hits 20

Re-Search All

Search Unmatched

1.	<u>Q9NH48</u>	Mass: 40662	Score: 53	Expect: 1.3	Queries matched: 7
	(KARG_ERISI) Arginine kinase (EC 2.7.3.3) (AK)			Arginine kinase (EC 2.7.3.3) (AK)	
	<u>P14208</u>	Mass: 40175	Score: 39	Expect: 30	Queries matched: 6
	(KARG_HOMGA) Arginine kinase (EC 2.7.3.3) (AK)			Arginine kinase (EC 2.7.3.3) (AK)	
	<u>Q9GYX1</u>	Mass: 40565	Score: 38	Expect: 37	Queries matched: 6
	(KARG_PACMR) Arginine kinase (EC 2.7.3.3) (AK)			Arginine kinase (EC 2.7.3.3) (AK)	
	<u>Q7WHM0</u>	Mass: 22461	Score: 33	Expect: 1.1e+02	Queries matched
	(RIMM_BORBR) Probable 16S rRNA processing protein rimM			Probable 16S rRNA proces	
	<u>Q7W6N7</u>	Mass: 22461	Score: 33	Expect: 1.1e+02	Queries matched
	(RIMM_BORPA) Probable 16S rRNA processing protein rimM			Probable 16S rRNA proces	
2.	<u>P51545</u>	Mass: 40255	Score: 50	Expect: 2.3	Queries matched: 7
	(KARG_PENJP) Arginine kinase (EC 2.7.3.3) (AK)			Arginine kinase (EC 2.7.3.3) (AK)	
	<u>Q81UJ9</u>	Mass: 23067	Score: 32	Expect: 1.4e+02	Queries matched
	(3MGH_BACAN) Putative 3-methyladenine DNA glycosylase (EC 3.2.2.-)			Putative 3-m	
3.	<u>Q9I2E5</u>	Mass: 30820	Score: 42	Expect: 15	Queries matched: 5
	(UPPP_PSEAE) Undecaprenyl-diphosphatase (EC 3.6.1.27)			(Undecaprenyl pyrophospha	

4.	<u>O05948</u>	Mass: 33409	Score: 40	Expect: 23	Queries matched: 5
	(FTSY_RICPR) Cell division protein ftsY homolog Cell division protein ftsY homo				
5.	<u>O75146-00-00-00</u>	Mass: 120011	Score: 40	Expect: 25	Queries matched: 8
	(HIP1R_HUMAN) Splice isoform Displayed; Variant Displayed; Conflict Displayed;				
	<u>P52663</u>	Mass: 32229	Score: 31	Expect: 1.8e+02	Queries matched
	(BLAN_ENTCL) Imipenem-hydrolyzing beta-lactamase precursor (EC 3.5.2.6) (Carbap				
	<u>P09255</u>	Mass: 31056	Score: 31	Expect: 2e+02	Queries matched:
	(IE68_VZVD) Gene 63/70 protein Gene 63/70 protein				
	<u>O75146-00-01-00</u>	Mass: 120043	Score: 30	Expect: 2.2e+02	Queries matched
	(HIP1R_HUMAN) Splice isoform Displayed; Variant dbSNP:2271051; Conflict Display				
6.	<u>Q9VJE4</u>	Mass: 13377	Score: 40	Expect: 25	Queries matched: 4
	(RPB11_DROME) DNA-directed RNA polymerase II 13.3 kDa polypeptide (EC 2.7.7.6)				
7.	<u>P13306</u>	Mass: 7234	Score: 38	Expect: 34	Queries matched: 3
	(Y06A_BPT4) Hypothetical 7.2 kDa protein in tk-vs intergenic region Hypothetica				
8.	<u>P55164</u>	Mass: 23163	Score: 38	Expect: 40	Queries matched: 5
	(CRBA2_CHICK) Beta crystallin A2 (Beta-A2-crystallin) Beta crystallin A2 (Beta-				
9.	<u>O75820-03-00-00</u>	Mass: 70294	Score: 37	Expect: 42	Queries matched: 7
	(ZN189_HUMAN) Splice isoform 4; Variant Displayed; Conflict Displayed; from O75				
	<u>P31548</u>	Mass: 25099	Score: 32	Expect: 1.4e+02	Queries matched
	(THIQ_ECOLI) Thiamine transport ATP-binding protein thiQ Thiamine transport ATP				
	<u>Q9Y696</u>	Mass: 28986	Score: 30	Expect: 2.2e+02	Queries matched
	(CLIC4_HUMAN) Chloride intracellular channel protein 4 (Intracellular chloride				
10.	<u>Q9ADT1</u>	Mass: 31930	Score: 37	Expect: 43	Queries matched: 5
	(FTR_METCA) Formylmethanofuran--tetrahydromethanopterin formyltransferase (EC 2				
11.	<u>Q55705</u>	Mass: 23824	Score: 36	Expect: 53	Queries matched: 4
	(Y232_SYNY3) Hypothetical protein slr0232 Hypothetical protein slr0232				
12.	<u>Q9KNM4</u>	Mass: 23318	Score: 36	Expect: 63	Queries matched: 4
	(KGUA_VIBCH) Guanylate kinase (EC 2.7.4.8) (GMP kinase) Guanylate kinase (EC 2.				
13.	<u>Q9Z7C4</u>	Mass: 48311	Score: 35	Expect: 75	Queries matched: 5
	(TOLB_CHLPN) TolB protein precursor TolB protein precursor				
14.	<u>Q8DGC6</u>	Mass: 56688	Score: 35	Expect: 81	Queries matched: 6
	(CHLB_SYNEL) Light-independent protochlorophyllide reductase subunit B (EC 1.18				
	<u>P37344</u>	Mass: 37137	Score: 31	Expect: 1.8e+02	Queries matched
	(PSPF_ECOLI) Psp operon transcriptional activator Psp operon transcriptional ac				
15.	<u>P11388-00-00-00</u>	Mass: 175030	Score: 35	Expect: 81	Queries matched: 8
	(TOP2A_HUMAN) Splice isoform 1; Variant Displayed; Conflict Displayed; from P11				
	<u>P11388-00-01-00</u>	Mass: 175002	Score: 35	Expect: 81	Queries matched: 8
	(TOP2A_HUMAN) Splice isoform 1; Variant amsacrine resistant cells; Conflict Dis				
	<u>P11388-00-02-00</u>	Mass: 175002	Score: 35	Expect: 81	Queries matched: 8
	(TOP2A_HUMAN) Splice isoform 1; Variant teniposide (VM-26) resistant cells; Con				
	<u>P11388-01-00-00</u>	Mass: 178143	Score: 34	Expect: 89	Queries matched: 8
	(TOP2A_HUMAN) Splice isoform 2; Variant Displayed; Conflict Displayed; from P11				
	<u>P11388-01-01-00</u>	Mass: 178115	Score: 34	Expect: 89	Queries matched: 8
	(TOP2A_HUMAN) Splice isoform 2; Variant amsacrine resistant cells; Conflict Dis				
	<u>P11388-01-02-00</u>	Mass: 178115	Score: 34	Expect: 89	Queries matched: 8
	(TOP2A_HUMAN) Splice isoform 2; Variant teniposide (VM-26) resistant cells; Con				
	<u>P11388-02-00-00</u>	Mass: 179412	Score: 34	Expect: 93	Queries matched: 8
	(TOP2A_HUMAN) Splice isoform 3; Variant Displayed; Conflict Displayed; from P11				
	<u>P11388-02-01-00</u>	Mass: 179384	Score: 34	Expect: 93	Queries matched: 8
	(TOP2A_HUMAN) Splice isoform 3; Variant amsacrine resistant cells; Conflict Dis				

	<u>P11388-02-02-00</u>	Mass: 179384	Score: 34	Expect: 93	Queries matched: 8
	(TOP2A_HUMAN)	Splice isoform 3; Variant teniposide (VM-26) resistant cells; Con			
	<u>P11388-03-00-00</u>	Mass: 183495	Score: 33	Expect: 1.1e+02	Queries matched
	(TOP2A_HUMAN)	Splice isoform 4; Variant Displayed; Conflict Displayed; from P11			
	<u>P11388-03-01-00</u>	Mass: 183467	Score: 33	Expect: 1.1e+02	Queries matched
	(TOP2A_HUMAN)	Splice isoform 4; Variant amsacrine resistant cells; Conflict Dis			
	<u>P11388-03-02-00</u>	Mass: 183467	Score: 33	Expect: 1.1e+02	Queries matched
	(TOP2A_HUMAN)	Splice isoform 4; Variant teniposide (VM-26) resistant cells; Con			
	<u>P11593</u>	Mass: 56889	Score: 32	Expect: 1.5e+02	Queries matched
	(VATB_NEUCR)	Vacuolar ATP synthase subunit B (EC 3.6.3.14) (V-ATPase B subunit)			
	<u>Q9ZHD3</u>	Mass: 25447	Score: 31	Expect: 1.7e+02	Queries matched
	(SILR_SALTY)	Probable transcriptional regulatory protein silR Probable transcri			
	<u>Q9X0E6</u>	Mass: 12227	Score: 31	Expect: 1.9e+02	Queries matched
	(CUTA_THEMA)	Divalent-cation tolerance protein cutA Divalent-cation tolerance p			
16.	<u>P42345</u>	Mass: 290795	Score: 35	Expect: 81	Queries matched: 10
	(FRAP_HUMAN)	FKBP12-rapamycin complex-associated protein (FK506-binding protein			
	<u>Q8VCG4</u>	Mass: 22668	Score: 34	Expect: 91	Queries matched: 4
	(CO8G_MOUSE)	Complement component C8 gamma chain precursor Complement component			
	<u>Q9W2D9</u>	Mass: 25823	Score: 32	Expect: 1.6e+02	Queries matched
	(IF3C_DROME)	Probable eukaryotic translation initiation factor 3 subunit 12 (eI			
17.	<u>Q834E3</u>	Mass: 26194	Score: 34	Expect: 95	Queries matched: 4
	(PYRF_ENTFA)	Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxyl			
18.	<u>Q9PB58</u>	Mass: 47917	Score: 33	Expect: 1.1e+02	Queries matched
	(SYS_XYLFA)	Seryl-tRNA synthetase (EC 6.1.1.11) (Serine--tRNA ligase) (SerRS) S			
19.	<u>O88813</u>	Mass: 77226	Score: 33	Expect: 1.1e+02	Queries matched
	(ACSL5_RAT)	Long-chain-fatty-acid--CoA ligase 5 (EC 6.2.1.3) (Long-chain acyl-C			
20.	<u>Q01771</u>	Mass: 45504	Score: 32	Expect: 1.4e+02	Queries matched
	(STADS_BRANA)	Acyl-[acyl-carrier-protein] desaturase, seed specific, chloroplas			

Search Parameters

Type of search : Peptide Mass Fingerprint
 Enzyme : Trypsin
 Fixed modifications : Carboxymethyl (C)
 Variable modifications : Oxidation (M)
 Mass values : Monoisotopic
 Protein Mass : Unrestricted
 Peptide Mass Tolerance : ± 1 Da
 Peptide Charge State : 1+
 Max Missed Cleavages : 1
 Number of queries : 12

Mascot: <http://www.matrixscience.com/>

{MATRIX} {SCIENCES} Mascot Search Results

Protein View

Match to: Q9NH48 Score: 53 Expect: 1.3
 (KARG_ERISI) Arginine kinase (EC 2.7.3.3) (AK) Arginine kinase (EC 2.7.3.3) (AK)

Nominal mass (M_r): 40661; Calculated pI value: 6.34

NCBI BLAST search of Q9NH48 against nr
 Unformatted sequence string for pasting into other applications

Taxonomy: Eriocheir sinensis

Fixed modifications: Carboxymethyl (C)
 Variable modifications: Oxidation (M)
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
 Number of mass values searched: 12
 Number of mass values matched: 7
 Sequence Coverage: 22%

Matched peptides shown in Bold Red

```

1 MADAATIAKL DEGFKKLEAA TDCKSLKKY LTKDVFQQLK AKKTKLGATL
51 LDVIQSGVEN LDSGVGVYAP DAEAYTLFSP LFDPIIEDYH KGFKQTDKHP
101 NKDFGDVTQF VNVDPDGKFV ISTRVRCGRS MEGYPFNPCL TEAQYKEMES
151 KVSSTLSNLE GELKGTYPFL TGMTKEVQQK LIDDHFLFKE GDRFLQAANA
201 CRYWPAGRGI YHNDNKTFV WCNEEDHLRI ISMQMGDLG QVYRRLVSAV
251 NEIEKRVVPS HHDRLLGFLTF CPTNLGTTVR ASVHIKLPKL AANREKLEEV
301 AGRYSLQVRG TRGEHTEAEG GIYDISNKRR MGLTEFQAVK EMQDGILELI
351 KIEKEMQ
    
```

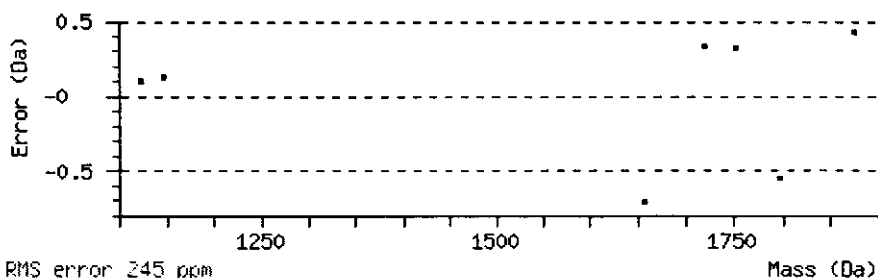
Show predicted peptides also

Sort Peptides By

Residue Number Increasing Mass Decreasing Mass

Start	End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
103	118	1753.13	1752.12	1751.80	0.32	0	K.DFGDVTQFVNVDPDGK.F
181	189	1147.74	1146.73	1146.61	0.13	0	K.LIDDHFLFK.E
265	280	1797.37	1796.36	1796.91	-0.56	0	R.LGFLTFCPTNLGTTVR.A
313	328	1720.11	1719.10	1718.77	0.33	0	R.GEHTEAEGGIYDISNK.R
313	329	1876.31	1875.30	1874.88	0.42	1	R.GEHTEAEGGIYDISNKR.R
331	340	1123.69	1122.68	1122.57	0.11	0	R.MGLTEFQAVK.E
341	354	1658.19	1657.19	1657.90	-0.71	1	K.EMQDGILELIKIEK.E

No match to: 1652.18, 2079.34, 2095.33, 2101.39, 2701.96



SWISS-PROT is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit

institutions as long as its content is in no way modified. Usage by and for commercial entities requires a license agreement. For information about the licensing scheme see: <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch.

Full entry text for [Q9NH48](#)

Mascot: <http://www.matrixscience.com/>

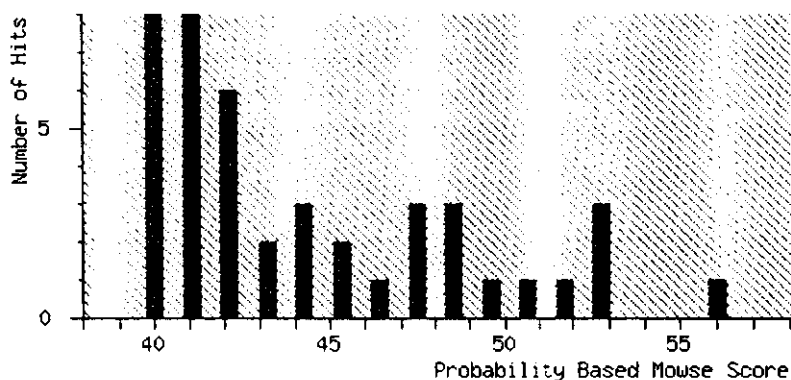
{MATRIX} Mascot Search Results

{SCIENCE}

User : Waraporn
Email : waraporn.p@psu.ac.th
Search title :
Database : NCBI nr 20050501 (2452561 sequences; 830617002 residues)
Timestamp : 13 May 2005 at 11:58:34 GMT
Top Score : 56 for gi|21754663, unnamed protein product [Homo sapiens]

Probability Based Mowse Score

Ions score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event. Protein scores greater than 76 are significant ($p < 0.05$).



Concise Protein Summary Report

[Help](#)

 Significance threshold $p <$

 Max. number of hits

1.	gi 21754663	Mass: 15180	Score: 56	Expect: 6.2	Queries matched: 5
	unnamed protein product [Homo sapiens]				
2.	gi 15896774	Mass: 29181	Score: 53	Expect: 12	Queries matched: 6
	Metallo-beta-lactamase superfamily hydrolase [Clostridium acetobutylicum ATCC 8				
3.	gi 7243759	Mass: 40662	Score: 53	Expect: 13	Queries matched: 7
	arginine kinase [Eriocheir sinensis]				
	gi 7243761	Mass: 40692	Score: 53	Expect: 13	Queries matched: 7
	arginine kinase [Chasmagnathus granulata]				
	gi 50591577	Mass: 11826	Score: 43	Expect: 1.3e+02	Queries matched
	hypothetical protein Ssui801000415 [Streptococcus suis 89/1591]				
4.	gi 54310294	Mass: 8789	Score: 52	Expect: 14	Queries matched: 4
	hypothetical protein BPRA3206 [Photobacterium profundum SS9]				
5.	gi 27463265	Mass: 40377	Score: 52	Expect: 15	Queries matched: 7
	allergen Pen m 2 [Penaeus monodon]				

6.	gi 12838818	Mass: 19108	Score: 51	Expect: 20	Queries matched: 5
	unnamed protein product [Mus musculus]				
7.	gi 551380	Mass: 40255	Score: 50	Expect: 24	Queries matched: 7
	arginine kinase, AK [Penaeus japonicus=shrimps, tail muscle, Peptide, 355 aa]				
8.	gi 41409471	Mass: 10185	Score: 49	Expect: 31	Queries matched: 4
	hypothetical protein MAP3373 [Mycobacterium avium subsp. paratuberculosis str.				
9.	gi 25530619	Mass: 9637	Score: 48	Expect: 36	Queries matched: 4
	transposase asl1657 [imported] - Nostoc sp. (strain PCC 7120)				
10.	gi 18957982	Mass: 24663	Score: 48	Expect: 39	Queries matched: 5
	putative transposase [Xenorhabdus nematophila]				
11.	gi 48852562	Mass: 41499	Score: 47	Expect: 46	Queries matched: 7
	COG0141: Histidinol dehydrogenase [Ferroplasma acidarmanus]				
12.	gi 42572121	Mass: 47092	Score: 47	Expect: 46	Queries matched: 6
	pseudouridine synthase family protein [Arabidopsis thaliana]				
	gi 30139293	Mass: 12331	Score: 43	Expect: 1.3e+02	Queries matched
	Transposase IS911 HTH and LZ region [Nitrosomonas europaea ATCC 19718]				
	gi 30180350	Mass: 12756	Score: 43	Expect: 1.3e+02	Queries matched
	Transposase IS911 HTH and LZ region [Nitrosomonas europaea ATCC 19718]				
13.	gi 868067	Mass: 11984	Score: 47	Expect: 48	Queries matched: 4
	envelope glycoprotein gp120				
14.	gi 25026981	Mass: 25996	Score: 46	Expect: 59	Queries matched: 5
	putative response regulator [Corynebacterium efficiens YS-314]				
15.	gi 39995542	Mass: 37778	Score: 46	Expect: 68	Queries matched: 6
	hypothetical protein GSU0434 [Geobacter sulfurreducens PCA]				
16.	gi 13195121	Mass: 43357	Score: 46	Expect: 69	Queries matched: 6
	RNase D [Agrobacterium tumefaciens]				
	gi 17935058	Mass: 43353	Score: 46	Expect: 69	Queries matched: 6
	ribonuclease D [Agrobacterium tumefaciens str. C58]				
	gi 15888489	Mass: 48150	Score: 42	Expect: 1.7e+02	Queries matched
	hypothetical protein AGR_C_2134 [Agrobacterium tumefaciens str. C58]				
17.	gi 26383198	Mass: 45723	Score: 44	Expect: 87	Queries matched: 6
	unnamed protein product [Mus musculus]				
18.	gi 1359768	Mass: 13329	Score: 44	Expect: 1e+02	Queries matched:
	immunoglobulin V(H) gene V-D-J [Homo sapiens]				
19.	gi 56962793	Mass: 14142	Score: 44	Expect: 1e+02	Queries matched:
	hypothetical protein ABC1019 [Bacillus clausii KSM-K16]				
20.	gi 54635543	Mass: 25958	Score: 44	Expect: 1.1e+02	Queries matched
	GA10234-PA [Drosophila pseudoobscura]				

Search Parameters

Type of search : Peptide Mass Fingerprint
 Enzyme : Trypsin
 Fixed modifications : Carboxymethyl (C)
 Variable modifications : Oxidation (M)
 Mass values : Monoisotopic
 Protein Mass : Unrestricted
 Peptide Mass Tolerance : ± 1 Da
 Peptide Charge State : 1+
 Max Missed Cleavages : 1
 Number of queries : 12

{MATRIX} {SCIENCES} Mascot Search Results

Protein View

Match to: **gi|21754663** Score: **56** Expect: **6.2**
unnamed protein product [Homo sapiens]

Nominal mass (M_r): **15180**; Calculated pI value: **9.26**

NCBI BLAST search of **gi|21754663** against nr
 Unformatted sequence string for pasting into other applications

Taxonomy: Homo sapiens

Fixed modifications: Carboxymethyl (C)
 Variable modifications: Oxidation (M)
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
 Number of mass values searched: 12
 Number of mass values matched: 5
 Sequence Coverage: 33%

Matched peptides shown in **Bold Red**

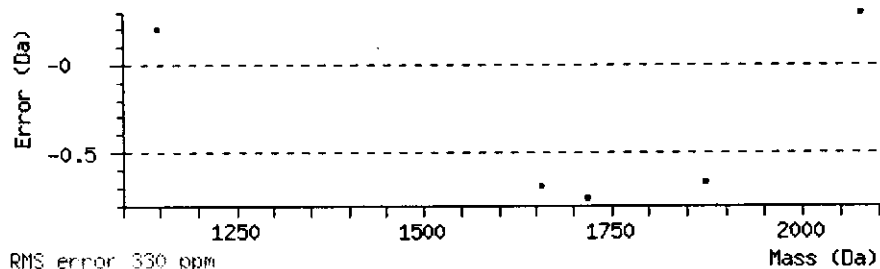
1 MVRGWPGSIS HGCNLNLCIQL LIVTHMSTGL THFSSVSHST TKEHHTQDSG
 51 ESHRLEEPGA WPHLGLISSV GCRSCARSSW VLASCCMRRK **WRAAPISSTV**
 101 **TMAGWVSSRR** **VTCEQSPGLR** **VIEVPNVRGL** **AGSYQG**

Show predicted peptides also

Sort Peptides By: Residue Number Increasing Mass Decreasing Mass

Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence	Ox:
91 - 109	2079.34	2078.33	2078.04	0.30	1	K.WRAAPISSTVTMAGWVSSR.R	
93 - 109	1720.11	1719.10	1719.86	-0.76	0	R.AAPISSTVTMAGWVSSR.R	
93 - 110	1876.31	1875.30	1875.96	-0.66	1	R.AAPISSTVTMAGWVSSRR.V	
111 - 120	1147.74	1146.73	1146.53	0.20	0	R.VTCEQSPGLR.V	
121 - 136	1658.19	1657.19	1657.88	-0.69	1	R.VIEVPNVRGLAGSYQG.-	

No match to: 1123.69, 1652.18, 1753.13, 1797.37, 2095.33, 2101.39, 2701.96



LOCUS BAC04544 136 aa linear PRI 30-JAN-2004
 DEFINITION unnamed protein product [Homo sapiens].
 ACCESSION BAC04544
 VERSION BAC04544.1 GI:21754663
 DBSOURCE accession AK095412.1
 KEYWORDS .
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Hominidae; Homo.

REFERENCE 1

AUTHORS

Ota,T., Suzuki,Y., Nishikawa,T., Otsuki,T., Sugiyama,T., Irie,R., Wakamatsu,A., Hayashi,K., Sato,H., Nagai,K., Kimura,K., Makita,H., Sekine,M., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T., Ishii,S., Yamamoto,J., Saito,K., Kawai,Y., Isono,Y., Nakamura,Y., Nagahari,K., Murakami,K., Yasuda,T., Iwayanagi,T., Wagatsuma,M., Shiratori,A., Sudo,H., Hosoiri,T., Kaku,Y., Kodaira,H., Kondo,H., Sugawara,M., Takahashi,M., Kanda,K., Yokoi,T., Furuya,T., Kikkawa,E., Omura,Y., Abe,K., Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M., Ninomiya,K., Ishibashi,T., Yamashita,H., Murakawa,K., Fujimori,K., Tanai,H., Kimata,M., Watanabe,M., Hiraoka,S., Chiba,Y., Ishida,S., Ono,Y., Takiguchi,S., Watanabe,S., Yosida,M., Hotuta,T., Kusano,J., Kanehori,K., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Imose,N., Musashino,K., Yuuki,H., Oshima,A., Sasaki,N., Aotsuka,S., Yoshikawa,Y., Matsunawa,H., Ichihara,T., Shiohata,N., Sano,S., Moriya,S., Momiyama,H., Satoh,N., Takami,S., Terashima,Y., Suzuki,O., Nakagawa,S., Senoh,A., Mizoguchi,H., Goto,Y., Shimizu,F., Wakebe,H., Hishigaki,H., Watanabe,T., Sugiyama,A., Takemoto,M., Kawakami,B., Yamazaki,M., Watanabe,K., Kumagai,A., Itakura,S., Fukuzumi,Y., Fujimori,Y., Komiyama,M., Tashiro,H., Tanigami,A., Fujiwara,T., Ono,T., Yamada,K., Fujii,Y., Ozaki,K., Hirao,M., Ohmori,Y., Kawabata,A., Hikiji,T., Kobatake,N., Inagaki,H., Ikema,Y., Okamoto,S., Okitani,R., Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T., Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Sasaki,M., Togashi,T., Oyama,M., Hata,H., Watanabe,M., Komatsu,T., Mizushima-Sugano,J., Satoh,T., Shirai,Y., Takahashi,Y., Nakagawa,K., Okumura,K., Nagase,T., Nomura,N., Kikuchi,H., Masuho,Y., Yamashita,R., Nakai,K., Yada,T., Nakamura,Y., Ohara,O., Isogai,T. and Sugano,S.

TITLE

Complete sequencing and characterization of 21,243 full-length human cDNAs

JOURNAL

Nat. Genet. 36 (1), 40-45 (2004)

PUBMED

14702039

REFERENCE 2

AUTHORS

Kanehori,K., Ishibashi,T., Chiba,Y., Fujimori,K., Hiraoka,S., Tanai,H., Watanabe,S., Ishida,S., Ono,Y., Hotuta,T., Watanabe,M., Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K., Wagatsuma,M., Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y., Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isogai,T.

TITLE

NEDO human cDNA sequencing project

JOURNAL

Unpublished

REFERENCE 3 (residues 1 to 136)

AUTHORS

Isogai,T. and Yamamoto,J.

TITLE

Direct Submission

JOURNAL

Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan

COMMENT

(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers

source

1..136
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CTONG2027361"
 /tissue_type="tongue, tumor tissue"
 /clone_lib="CTONG2"
 /note="cloning vector: pME18SFL3"

Protein

1..136
 /name="unnamed protein product"

CDS

1..136

/coded_by="AK095412.1:513..923"

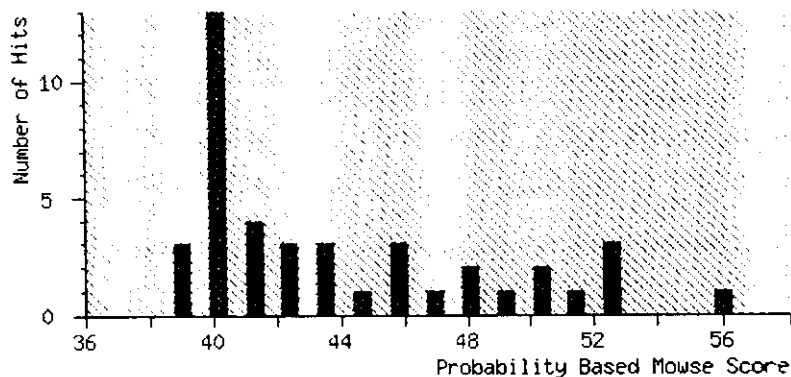
Mascot: <http://www.matrixscience.com/>

**{MATRIX}
{SCIENCES} Mascot Search Results**

User : Waraporn
 Email : waraporn.p@psu.ac.th
 Search title :
 Database : MSDB 20050227 (1942918 sequences; 629040812 residues)
 Timestamp : 13 May 2005 at 11:26:06 GMT
 Top Score : 56 for Q8N1N6_HUMAN, Hypothetical protein FLJ38093.- Homo sapiens (1)

Probability Based Mowse Score

Ions score is $-10 \cdot \log(P)$, where P is the probability that the observed match is a random event.
 Protein scores greater than 75 are significant ($p < 0.05$).

**Concise Protein Summary Report**

Format AS

Concise Protein Summary [Help](#)Significance threshold $p <$ Max. number of hits Re-Search All Search Unmatched

1.	<u>Q8N1N6_HUMAN</u> Hypothetical protein FLJ38093.- Homo sapiens (Human).	Mass: 15180	Score: 56	Expect: 4.9	Queries matched: 5
2.	<u>D97334</u> metallo-beta-lactamase superfamily hydrolase [imported] - Clostridium acetobuty	Mass: 29181	Score: 53	Expect: 9.7	Queries matched: 6
3.	<u>Q9NH47_9EUCA</u> Arginine kinase.- Chasmagnathus granulata.	Mass: 40692	Score: 53	Expect: 10	Queries matched: 7
	<u>AAF43437</u> AF233356 NID: - Eriocheir sinensis	Mass: 40662	Score: 53	Expect: 10	Queries matched: 7
4.	<u>Q6LMG5_PHOPR</u> Hypothetical protein VV2784.- Photobacterium profundum (Photobacterium sp. (str	Mass: 8789	Score: 52	Expect: 11	Queries matched: 4
5.	<u>Q8I9P7_PENMO</u> Allergen Pen m 2.- Penaeus monodon (Penaeid shrimp).	Mass: 40377	Score: 52	Expect: 12	Queries matched: 7
6.	<u>Q9DAC2_MOUSE</u> Mus musculus adult male testis cDNA, RIKEN full-length enriched library, clone:	Mass: 19108	Score: 51	Expect: 16	Queries matched: 5

7.	<u>KARG_PENJP</u>	Mass: 40255	Score: 50	Expect: 19	Queries matched: 7
	Arginine kinase (EC 2.7.3.3) (AK).- <i>Penaeus japonicus</i> (Kuruma prawn).				
8.	<u>Q73UJ5_MYCPA</u>	Mass: 10185	Score: 49	Expect: 24	Queries matched: 4
	Hypothetical protein.- <i>Mycobacterium paratuberculosis</i> .				
9.	<u>AC2013</u>	Mass: 9637	Score: 48	Expect: 29	Queries matched: 4
	transposase asl1657 [imported] - <i>Nostoc</i> sp. (strain PCC 7120)				
10.	<u>Q8RLV2_XENNE</u>	Mass: 24663	Score: 48	Expect: 31	Queries matched: 5
	Putative transposase.- <i>Xenorhabdus nematophilus</i> .				
11.	<u>Q71268_9HIV1</u>	Mass: 11984	Score: 47	Expect: 38	Queries matched: 4
	Envelope glycoprotein gp120 (Fragment).- Human immunodeficiency virus 1.				
12.	<u>Q8FSG8_COREF</u>	Mass: 25996	Score: 46	Expect: 47	Queries matched: 5
	Putative response regulator.- <i>Corynebacterium efficiens</i> .				
	<u>CAF17619</u>	Mass: 6491	Score: 40	Expect: 2.1e+02	Queries matched
	AX972511 NID: - <i>Homo sapiens</i>				
13.	<u>Q74G18_GEOSL</u>	Mass: 37778	Score: 46	Expect: 54	Queries matched: 6
	Hypothetical protein.- <i>Geobacter sulfurreducens</i> .				
14.	<u>AF2718</u>	Mass: 43353	Score: 46	Expect: 55	Queries matched: 6
	ribonuclease D [imported] - <i>Agrobacterium tumefaciens</i> (strain C58, Dupont)				
	<u>Q9AF17_9RHIZ</u>	Mass: 43357	Score: 46	Expect: 55	Queries matched: 6
	RNase D.- <i>Agrobacterium tumefaciens</i> .				
	<u>B97500</u>	Mass: 48150	Score: 42	Expect: 1.3e+02	Queries matched
	ribonuclease D (PA1294) [imported] - <i>Agrobacterium tumefaciens</i> (strain C58, Cer				
15.	<u>Q8BMW2_MOUSE</u>	Mass: 45723	Score: 44	Expect: 69	Queries matched: 6
	<i>Mus musculus</i> 8 days embryo whole body cDNA, RIKEN full-length enriched library,				
16.	<u>CAA61289</u>	Mass: 13329	Score: 44	Expect: 79	Queries matched: 4
	IMMUNOGLOBULIN V(H) GENE V-D-J (FRAGMENT).- <i>Homo sapiens</i> (Human).				
17.	<u>Q5WJ97_BACSK</u>	Mass: 14142	Score: 44	Expect: 79	Queries matched: 4
	Hypothetical protein.- <i>Bacillus clausii</i> (strain KSM-K16).				
18.	<u>Q8XRW7_RALSO</u>	Mass: 31278	Score: 43	Expect: 93	Queries matched: 6
	PUTATIVE D-ALANINE AMINOTRANSFERASE (D-ASPARTATE AMINOTRANSFERASE) PROTEIN (EC				
19.	<u>Q81ZQ0_NITEU</u>	Mass: 12331	Score: 43	Expect: 1e+02	Queries matched:
	Transposase IS911 HTH and LZ region.- <i>Nitrosomonas europaea</i> .				
	<u>Q82WD9_NITEU</u>	Mass: 12756	Score: 43	Expect: 1e+02	Queries matched:
	Transposase IS911 HTH and LZ region.- <i>Nitrosomonas europaea</i> .				
20.	<u>F83401</u>	Mass: 30820	Score: 42	Expect: 1.2e+02	Queries matched
	bacitracin resistance protein PA1959 [imported] - <i>Pseudomonas aeruginosa</i> (strai				

Search Parameters

Type of search : Peptide Mass Fingerprint
 Enzyme : Trypsin
 Fixed modifications : Carboxymethyl (C)
 Variable modifications : Oxidation (M)
 Mass values : Monoisotopic
 Protein Mass : Unrestricted
 Peptide Mass Tolerance : ± 1 Da
 Peptide Charge State : 1+
 Max Missed Cleavages : 1
 Number of queries : 12

Mascot: <http://www.matrixscience.com/>

{MATRIX} {SCIENCE} Mascot Search Results

Protein View

Match to: Q8N1N6_HUMAN Score: 56 Expect: 4.9
 Hypothetical protein FLJ38093.- Homo sapiens (Human).

Nominal mass (M_r): 15180; Calculated pI value: 9.26

NCBI BLAST search of Q8N1N6_HUMAN against nr
 Unformatted sequence string for pasting into other applications

Taxonomy: Homo sapiens

Links to retrieve other entries containing this sequence from NCBI Entrez:
BAC04544 from Homo sapiens

Fixed modifications: Carboxymethyl (C)
 Variable modifications: Oxidation (M)
 Cleavage by Trypsin: cuts C-term side of KR unless next residue is P
 Number of mass values searched: 12
 Number of mass values matched: 5
 Sequence Coverage: 33%

Matched peptides shown in Bold Red

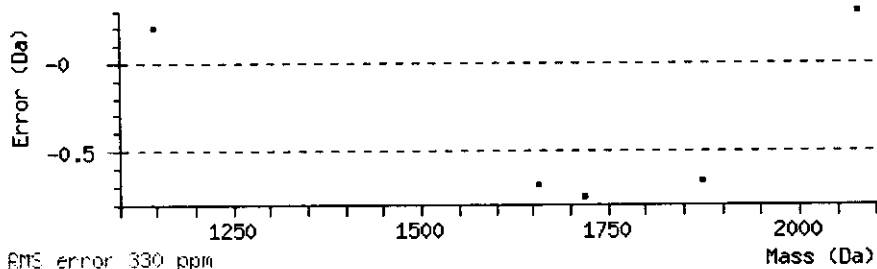
1 MVRGWPGSIS HGCNLCIQL LIVTHMSTGL THFSSVSHST TKEHHTQDSG
 51 ESHRLEEPGA WPHLGLISSV GCRSCARSSW VLASCCMRK **WRAAPISSTV**
 101 **TMAGWSSRR** **VTCEQSPGLR** **VIEVPNVRGL** **AGSYQG**

Show predicted peptides also

Sort Peptides By Residue Number Increasing Mass Decreasing Mass

Start - End	Observed	Mr(expt)	Mr(calc)	Delta	Miss	Sequence
91 - 109	2079.34	2078.33	2078.04	0.30	1	K.WRAAPISSTVTMAGWSSR.R Ox
93 - 109	1720.11	1719.10	1719.86	-0.76	0	R.AAPISSTVTMAGWSSR.R
93 - 110	1876.31	1875.30	1875.96	-0.66	1	R.AAPISSTVTMAGWSSRR.V
111 - 120	1147.74	1146.73	1146.53	0.20	0	R.VTCEQSPGLR.V
121 - 136	1658.19	1657.19	1657.88	-0.69	1	R.VIEVPNVRGLAGSYQG.-

No match to: 1123.69, 1652.18, 1753.13, 1797.37, 2095.33, 2101.39, 2701.96



>P1;Q8N1N6_HUMAN
 Hypothetical protein FLJ38093.- Homo sapiens (Human).
 C;Species Q8N1N6_HUMAN: Homo sapiens (Human).
 C;Species BAC04544: Homo sapiens
 C;Family: Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
 C;Accession: Q8N1N6;
 C;Reference [1]
 C;NUCLEOTIDE SEQUENCE.

