

## BLAST

## Basic Local Alignment Search Tool

NCBI/ [BLAST](#)/ [blastp suite](#)/ [Formatting Results - RWHM518U01S](#)

## Cmerolaegenewise gi|151559140|dbj|AP006497.2|:subseq(386500,20000).pep

<b>Query ID</b>	lcl 82391	<b>Database Name</b>	nr
<b>Description</b>	Cmerolaegenewise gi 151559140 dbj AP006497.2 :subseq(386500,20000).pep	<b>Description</b>	All non-redundant GenBank CDS translations+PDB+SwissProt+PIR+F excluding environmental samples from WGS projects
<b>Molecule type</b>	amino acid	<b>Program</b>	BLASTP 2.2.25+
<b>Query Length</b>	91		

## Descriptions

Legend for links to other resources: [U](#) UniGene [E](#) GEO [G](#) Gene [S](#) Structure [M](#) Map Viewer [P](#) PubChem BioAssay

<a href="#">Accession</a>	<a href="#">Description</a>	<a href="#">Max score</a>	<a href="#">Total score</a>	<a href="#">Query coverage</a>	<a href="#">E value</a>	<a href="#">Links</a>
<a href="#">NP_001049763.1</a>	Os03g0284600 [Oryza sativa Japonica Group] >gb ABF95344.1  UPF0308 protein, chloroplast precursor, putative, expressed [Oryza sativa Japonica Group] >dbj BAF11677.1	<a href="#">60.8</a>	60.8	95%	5e-08	<a href="#">UG</a>
	Os03g0284600 [Oryza sativa Japonica Group] >gb EEE58829.1  hypothetical protein OsJ_10400 [Oryza sativa Japonica Group]					
<a href="#">EEC75002.1</a>	hypothetical protein Osl_11064 [Oryza sativa Indica Group]	<a href="#">60.5</a>	60.5	95%	7e-08	
<a href="#">NP_001145525.1</a>	hypothetical protein LOC100278941 [Zea mays] >gb ACG48189.1  hypothetical protein [Zea mays]	<a href="#">60.5</a>	60.5	95%	8e-08	<a href="#">UG</a>
<a href="#">XP_002468056.1</a>	hypothetical protein SORBIDRAFT_01g038790 [Sorghum bicolor] >gb EER95054.1  hypothetical protein SORBIDRAFT_01g038790 [Sorghum bicolor]	<a href="#">58.2</a>	58.2	95%	3e-07	<a href="#">UG</a>
<a href="#">ACR37670.1</a>	unknown [Zea mays]	<a href="#">58.2</a>	58.2	95%	4e-07	
<a href="#">ACN35248.1</a>	unknown [Zea mays]	<a href="#">57.8</a>	57.8	95%	5e-07	<a href="#">G</a>
<a href="#">XP_002334170.1</a>	predicted protein [Populus trichocarpa] >gb EEF07066.1  predicted protein [Populus trichocarpa]	<a href="#">57.4</a>	57.4	94%	6e-07	<a href="#">UG</a>
<a href="#">XP_002326159.1</a>	predicted protein [Populus trichocarpa] >gb EEE71829.1  predicted protein [Populus trichocarpa]	<a href="#">57.0</a>	57.0	94%	7e-07	<a href="#">UG</a>
<a href="#">ACU20039.1</a>	unknown [Glycine max]	<a href="#">56.6</a>	56.6	94%	1e-06	
<a href="#">XP_002965388.1</a>	hypothetical protein SELMODRAFT_68006 [Selaginella moellendorffii] >gb EFJ34226.1  hypothetical protein SELMODRAFT_68006 [Selaginella moellendorffii]	<a href="#">55.5</a>	55.5	95%	2e-06	<a href="#">UG</a>

Accession	Description	Max score	Total score	Query coverage	E value	Links
<a href="#">XP_002977235.1</a>	hypothetical protein SELMODRAFT_58056 [Selaginella moellendorffii] >gb EFJ21844.1  hypothetical protein SELMODRAFT_58056 [Selaginella moellendorffii]	<a href="#">55.1</a>	55.1	95%	3e-06	<a href="#">UG</a>
<a href="#">ABK23985.1</a>	unknown [Picea sitchensis]	<a href="#">53.1</a>	53.1	94%	1e-05	
<a href="#">XP_002273449.1</a>	PREDICTED: hypothetical protein [Vitis vinifera] >emb CBI32627.3  unnamed protein product [Vitis vinifera]	<a href="#">53.1</a>	53.1	94%	1e-05	<a href="#">UGM</a>
<a href="#">XP_002881499.1</a>	hypothetical protein ARALYDRAFT_482716 [Arabidopsis lyrata subsp. lyrata] >gb EFH57758.1  hypothetical protein ARALYDRAFT_482716 [Arabidopsis lyrata subsp. lyrata]	<a href="#">52.0</a>	52.0	95%	2e-05	<a href="#">G</a>
<a href="#">XP_002525198.1</a>	conserved hypothetical protein [Ricinus communis] >gb EEF37164.1  conserved hypothetical protein [Ricinus communis]	<a href="#">51.6</a>	51.6	94%	3e-05	<a href="#">G</a>
<a href="#">NP_030274.1</a>	unknown protein [Arabidopsis thaliana] >sp Q9ZUU2.2 U308_ARATH RecName: Full=UPF0308 protein At2g37240, chloroplastic; Flags: Precursor >gb AAK91362.1  At2g37240/F3G5.3 [Arabidopsis thaliana] >gb AAC98045.2  expressed protein [Arabidopsis thaliana] >gb AAM67197.1  unknown [Arabidopsis thaliana] >gb AAP21145.1  At2g37240/F3G5.3 [Arabidopsis thaliana]	<a href="#">50.8</a>	50.8	95%	5e-05	<a href="#">UG</a>
<a href="#">XP_002866692.1</a>	hypothetical protein ARALYDRAFT_496819 [Arabidopsis lyrata subsp. lyrata] >gb EFH42951.1  hypothetical protein ARALYDRAFT_496819 [Arabidopsis lyrata subsp. lyrata]	<a href="#">47.4</a>	47.4	97%	7e-04	<a href="#">G</a>
<a href="#">XP_002666723.1</a>	PREDICTED: UPF0308 protein C9orf21 homolog [Danio rerio]	<a href="#">47.0</a>	47.0	89%	8e-04	<a href="#">UGM</a>
<a href="#">NP_201385.2</a>	unknown protein [Arabidopsis thaliana] >gb AAM20692.1  unknown protein [Arabidopsis thaliana] >gb AAN15652.1  unknown protein [Arabidopsis thaliana]	<a href="#">47.0</a>	47.0	97%	0.001	<a href="#">UG</a>
<a href="#">XP_002928763.1</a>	PREDICTED: UPF0308 protein C9orf21-like, partial [Ailuropoda melanoleuca]	<a href="#">45.8</a>	45.8	89%	0.002	<a href="#">GM</a>
<a href="#">EFB20490.1</a>	hypothetical protein PANDA_018799 [Ailuropoda melanoleuca]	<a href="#">45.8</a>	45.8	89%	0.002	
<a href="#">XP_002320012.1</a>	predicted protein [Populus trichocarpa] >gb EEE98327.1  predicted protein [Populus trichocarpa]	<a href="#">45.4</a>	45.4	93%	0.002	<a href="#">UG</a>
<a href="#">XP_001493974.2</a>	PREDICTED: similar to UPF0308 protein C9orf21 [Equus caballus]	<a href="#">45.4</a>	45.4	89%	0.003	<a href="#">UGM</a>
<a href="#">XP_003130905.1</a>	PREDICTED: UPF0308 protein C9orf21 homolog isoform 2 [Sus scrofa]	<a href="#">45.1</a>	45.1	95%	0.003	<a href="#">UGM</a>
<a href="#">XP_002742741.1</a>	PREDICTED: UPF0308 protein C9orf21-like [Callithrix jacchus]	<a href="#">45.1</a>	45.1	89%	0.003	<a href="#">GM</a>
<a href="#">NP_001069904.1</a>	hypothetical protein LOC616897 [Bos taurus] >sp Q148E0.1 CI021_BOVIN RecName: Full=UPF0308 protein C9orf21 homolog >gb AAI18426.1  Chromosome 9 open reading	<a href="#">45.1</a>	45.1	89%	0.003	<a href="#">UGM</a>

Accession	Description	Max score	Total score	Query coverage	E value	Links
	frame 21 ortholog [Bos taurus] >gb DAA26600.1  hypothetical protein LOC616897 [Bos taurus]					
<a href="#">XP_001153024.1</a>	PREDICTED: similar to TPA_exp: C9ORF21 isoform 1 [Pan troglodytes]	<a href="#">45.1</a>	45.1	95%	0.003	<a href="#">UGM</a>
<a href="#">XP_003130904.1</a>	PREDICTED: UPF0308 protein C9orf21 homolog isoform 1 [Sus scrofa]	<a href="#">44.7</a>	44.7	89%	0.004	<a href="#">UGM</a>
<a href="#">EAW92661.1</a>	chromosome 9 open reading frame 21, isoform CRA_a [Homo sapiens]	<a href="#">44.7</a>	44.7	100%	0.004	<a href="#">G</a>
<a href="#">XP_001588758.1</a>	hypothetical protein SS1G_10305 [Sclerotinia sclerotiorum 1980] >gb EDN94432.1  hypothetical protein SS1G_10305 [Sclerotinia sclerotiorum 1980]	<a href="#">44.7</a>	44.7	96%	0.004	<a href="#">G</a>
<a href="#">XP_002820049.1</a>	PREDICTED: UPF0308 protein C9orf21-like [Pongo abelii]	<a href="#">44.7</a>	44.7	89%	0.004	<a href="#">GM</a>
<a href="#">XP_001116318.1</a>	PREDICTED: UPF0308 protein C9orf21-like [Macaca mulatta]	<a href="#">44.7</a>	44.7	82%	0.004	<a href="#">GM</a>
<a href="#">XP_520707.2</a>	PREDICTED: similar to TPA_exp: C9ORF21 isoform 2 [Pan troglodytes]	<a href="#">44.7</a>	44.7	89%	0.004	<a href="#">UGM</a>
<a href="#">EAW92663.1</a>	chromosome 9 open reading frame 21, isoform CRA_c [Homo sapiens]	<a href="#">44.3</a>	44.3	95%	0.005	<a href="#">G</a>
<a href="#">ZP_02926219.1</a>	hypothetical protein VspiD_06235 [Verrucomicrobium spinosum DSM 4136]	<a href="#">44.3</a>	44.3	92%	0.006	
<a href="#">NP_714542.1</a>	hypothetical protein LOC195827 [Homo sapiens] >sp Q7RTV5.1 CI021_HUMAN RecName: Full=UPF0308 protein C9orf21 >tpg DAA00065.1  TPA_exp: C9ORF21 [Homo sapiens] >emb CAI40534.1  novel protein [Homo sapiens] >gb EAW92662.1  chromosome 9 open reading frame 21, isoform CRA_b [Homo sapiens] >gb AAI36504.1  Chromosome 9 open reading frame 21 [Homo sapiens]	<a href="#">43.9</a>	43.9	89%	0.006	<a href="#">UGM</a>
<a href="#">ACU20324.1</a>	unknown [Glycine max]	<a href="#">43.9</a>	43.9	93%	0.007	
<a href="#">ACN32054.1</a>	unknown [Zea mays]	<a href="#">43.5</a>	43.5	93%	0.008	<a href="#">G</a>
<a href="#">XP_001106503.1</a>	PREDICTED: UPF0308 protein C9orf21-like [Macaca mulatta]	<a href="#">43.5</a>	43.5	89%	0.008	<a href="#">UGM</a>
<a href="#">ADE77692.1</a>	unknown [Picea sitchensis]	<a href="#">43.5</a>	43.5	93%	0.008	
<a href="#">CBI33071.3</a>	unnamed protein product [Vitis vinifera]	<a href="#">43.5</a>	43.5	93%	0.009	
<a href="#">XP_002708298.1</a>	PREDICTED: hypothetical protein [Oryctolagus cuniculus]	<a href="#">43.5</a>	43.5	89%	0.009	<a href="#">UGM</a>
<a href="#">XP_002263922.1</a>	PREDICTED: hypothetical protein [Vitis vinifera] >emb CAN81556.1  hypothetical protein VITISV_040398 [Vitis vinifera]	<a href="#">43.5</a>	43.5	93%	0.010	<a href="#">UGM</a>
<a href="#">CBI33293.3</a>	unnamed protein product [Vitis vinifera]	<a href="#">43.1</a>	43.1	93%	0.011	
<a href="#">XP_002463941.1</a>	hypothetical protein SORBIDRAFT_01g009340 [Sorghum bicolor] >gb EER90939.1  hypothetical protein SORBIDRAFT_01g009340 [Sorghum bicolor]	<a href="#">43.1</a>	43.1	93%	0.011	<a href="#">UG</a>
<a href="#">XP_001547900.1</a>	hypothetical protein BC1G_13584 [Botryotinia fuckeliana B05.10] >gb EDN19735.1  hypothetical protein BC1G_13584 [Botryotinia fuckeliana]	<a href="#">42.7</a>	42.7	98%	0.014	<a href="#">G</a>

Accession	Description	Max score	Total score	Query coverage	E value	Links
	B05.10]					
<a href="#">XP_681257.1</a>	hypothetical protein AN7988.2 [Aspergillus nidulans FGSC A4] >gb EAA58791.1  hypothetical protein AN7988.2 [Aspergillus nidulans FGSC A4] >tpe CBF73658.1  TPA: conserved hypothetical protein [Aspergillus nidulans FGSC A4]	<a href="#">42.7</a>	42.7	98%	0.015	<a href="#">G</a>
<a href="#">EAY91739.1</a>	hypothetical protein Osl_13380 [Oryza sativa Indica Group]	<a href="#">42.7</a>	42.7	93%	0.017	
<a href="#">NP_001051154.1</a>	Os03g0729300 [Oryza sativa Japonica Group] >gb AAO38466.1  unknown protein [Oryza sativa Japonica Group] >gb ABF98681.1  UPF0308 protein, chloroplast precursor, putative, expressed [Oryza sativa Japonica Group] >dbj BAF13068.1  Os03g0729300 [Oryza sativa Japonica Group] >dbj BAG98418.1  unnamed protein product [Oryza sativa Japonica Group]	<a href="#">42.7</a>	42.7	93%	0.017	<a href="#">UG</a>
<a href="#">ABF98680.1</a>	expressed protein [Oryza sativa Japonica Group] >dbj BAG90239.1  unnamed protein product [Oryza sativa Japonica Group]	<a href="#">42.7</a>	42.7	93%	0.017	
<a href="#">XP_002191279.1</a>	PREDICTED: hypothetical protein [Taeniopygia guttata]	<a href="#">42.4</a>	42.4	89%	0.018	<a href="#">UGM</a>
<a href="#">EEE59859.1</a>	hypothetical protein OsJ_12441 [Oryza sativa Japonica Group]	<a href="#">42.4</a>	42.4	93%	0.019	
<a href="#">NP_001051153.1</a>	Os03g0729200 [Oryza sativa Japonica Group] >gb AAO38464.1  hypothetical protein [Oryza sativa Japonica Group] >gb ABF98679.1  expressed protein [Oryza sativa Japonica Group] >dbj BAF13067.1  Os03g0729200 [Oryza sativa Japonica Group] >gb EAY91738.1  hypothetical protein Osl_13379 [Oryza sativa Indica Group] >dbj BAG94118.1  unnamed protein product [Oryza sativa Japonica Group] >gb EEE59858.1  hypothetical protein OsJ_12440 [Oryza sativa Japonica Group]	<a href="#">42.4</a>	42.4	93%	0.019	<a href="#">UG</a>
<a href="#">XP_848380.1</a>	PREDICTED: similar to UPF0308 protein C9orf21 [Canis familiaris]	<a href="#">42.4</a>	42.4	89%	0.019	<a href="#">UGM</a>
<a href="#">XP_002463942.1</a>	hypothetical protein SORBIDRAFT_01g009350 [Sorghum bicolor] >gb EER90940.1  hypothetical protein SORBIDRAFT_01g009350 [Sorghum bicolor]	<a href="#">42.0</a>	42.0	93%	0.023	<a href="#">UG</a>
<a href="#">NP_001102584.1</a>	hypothetical protein LOC498685 [Rattus norvegicus] >gb EDL84433.1  similar to UPF0308 protein C9orf21, isoform CRA_a [Rattus norvegicus]	<a href="#">41.6</a>	41.6	89%	0.032	<a href="#">UG</a>
<a href="#">XP_002269002.1</a>	PREDICTED: hypothetical protein, partial [Vitis vinifera]	<a href="#">41.6</a>	41.6	93%	0.033	<a href="#">UGM</a>
<a href="#">ACU18761.1</a>	unknown [Glycine max]	<a href="#">41.6</a>	41.6	93%	0.034	
<a href="#">ACU24130.1</a>	unknown [Glycine max]	<a href="#">41.6</a>	41.6	93%	0.035	
<a href="#">EDL84435.1</a>	similar to UPF0308 protein C9orf21, isoform CRA_c [Rattus norvegicus]	<a href="#">41.6</a>	41.6	89%	0.036	<a href="#">G</a>
<a href="#">CBI33289.3</a>	unnamed protein product [Vitis vinifera]	<a href="#">41.6</a>	41.6	93%	0.039	

Accession	Description	Max score	Total score	Query coverage	E value	Links
<a href="#">XP_002628008.1</a>	conserved hypothetical protein [Ajellomyces dermatitidis SLH14081] >gb EEQ72686.1  conserved hypothetical protein [Ajellomyces dermatitidis SLH14081] >gb EEQ88802.1  conserved hypothetical protein [Ajellomyces dermatitidis ER-3]	<a href="#">41.2</a>	41.2	76%	0.040	<a href="#">G</a>
<a href="#">ACN25853.1</a>	unknown [Zea mays]	<a href="#">41.2</a>	41.2	93%	0.042	<a href="#">G</a>
<a href="#">NP_001170360.1</a>	hypothetical protein LOC100384338 [Zea mays] >gb ACN36745.1  unknown [Zea mays]	<a href="#">41.2</a>	41.2	93%	0.047	<a href="#">UG</a>
<a href="#">XP_001784902.1</a>	predicted protein [Physcomitrella patens subsp. patens] >gb EDQ50291.1  predicted protein [Physcomitrella patens subsp. patens]	<a href="#">40.8</a>	40.8	98%	0.052	<a href="#">UG</a>
<a href="#">YP_001543559.1</a>	hypothetical protein Haur_0783 [Herpetosiphon aurantiacus ATCC 23779] >gb ABX03431.1  hypothetical protein Haur_0783 [Herpetosiphon aurantiacus ATCC 23779]	<a href="#">40.8</a>	40.8	100%	0.053	<a href="#">G</a>
<a href="#">CBI57389.1</a>	unnamed protein product [Sordaria macrospora]	<a href="#">40.8</a>	40.8	87%	0.058	
<a href="#">CAN81555.1</a>	hypothetical protein VITISV_040397 [Vitis vinifera]	<a href="#">40.8</a>	40.8	93%	0.063	
<a href="#">XP_961259.1</a>	hypothetical protein NCU04268 [Neurospora crassa OR74A] >emb CAC28745.1  putative protein [Neurospora crassa] >gb EAA32023.1  conserved hypothetical protein [Neurospora crassa OR74A]	<a href="#">40.8</a>	40.8	87%	0.063	<a href="#">G</a>
<a href="#">AAI36502.1</a>	C9orf21 protein [Homo sapiens]	<a href="#">40.8</a>	40.8	83%	0.064	<a href="#">G</a>
<a href="#">ADO28366.1</a>	upf0308 protein c9orf21-like protein [Ictalurus furcatus]	<a href="#">40.4</a>	40.4	89%	0.068	
<a href="#">XP_002263959.1</a>	PREDICTED: hypothetical protein [Vitis vinifera]	<a href="#">40.4</a>	40.4	93%	0.078	<a href="#">UGM</a>
<a href="#">XP_002533575.1</a>	conserved hypothetical protein [Ricinus communis] >gb EEF28810.1  conserved hypothetical protein [Ricinus communis]	<a href="#">40.4</a>	40.4	93%	0.080	<a href="#">G</a>
<a href="#">CAI40533.1</a>	novel protein [Homo sapiens]	<a href="#">40.4</a>	40.4	83%	0.083	<a href="#">G</a>
<a href="#">BAE40544.1</a>	unnamed protein product [Mus musculus]	<a href="#">39.7</a>	39.7	89%	0.12	<a href="#">G</a>
<a href="#">ZP_01089021.1</a>	hypothetical protein DSM3645_00440 [Blastopirellula marina DSM 3645] >gb EAQ82136.1  hypothetical protein DSM3645_00440 [Blastopirellula marina DSM 3645]	<a href="#">39.7</a>	39.7	100%	0.12	
<a href="#">BAB24662.1</a>	unnamed protein product [Mus musculus]	<a href="#">39.7</a>	39.7	89%	0.12	<a href="#">G</a>
<a href="#">XP_001824981.1</a>	hypothetical protein AOR_1_1228084 [Aspergillus oryzae RIB40] >dbj BAE63848.1  unnamed protein product [Aspergillus oryzae]	<a href="#">39.7</a>	39.7	85%	0.13	<a href="#">G</a>
<a href="#">XP_002383597.1</a>	conserved hypothetical protein [Aspergillus flavus NRRL3357] >gb EED46061.1  conserved hypothetical protein [Aspergillus flavus NRRL3357]	<a href="#">39.7</a>	39.7	85%	0.13	<a href="#">G</a>
<a href="#">YP_001543558.1</a>	hypothetical protein Haur_0782 [Herpetosiphon aurantiacus ATCC 23779] >gb ABX03430.1  hypothetical protein Haur_0782 [Herpetosiphon aurantiacus ATCC 23779]	<a href="#">39.7</a>	39.7	89%	0.14	<a href="#">G</a>
<a href="#">BAE26709.1</a>	unnamed protein product [Mus musculus]	<a href="#">39.7</a>	39.7	89%	0.14	<a href="#">G</a>

<b>Accession</b>	<b>Description</b>	<b><u>Max score</u></b>	<b><u>Total score</u></b>	<b><u>Query coverage</u></b>	<b><u>E value</u></b>	<b>Links</b>
<a href="#">XP_001940231.1</a>	conserved hypothetical protein [Pyrenophora tritici-repentis Pt-1C-BFP] >gb EDU42950.1  conserved hypothetical protein [Pyrenophora tritici-repentis Pt-1C-BFP]	<a href="#">39.7</a>	39.7	96%	0.14	<a href="#">G</a>
<a href="#">EFQ95671.1</a>	hypothetical protein PTT_05482 [Pyrenophora teres f. teres 0-1]	<a href="#">39.7</a>	39.7	96%	0.14	
<a href="#">EDL16239.1</a>	RIKEN cDNA 1110018J18, isoform CRA_c [Mus musculus]	<a href="#">39.3</a>	39.3	89%	0.16	<a href="#">G</a>
<a href="#">NP_001158627.1</a>	UPF0308 protein C9orf21 homolog [Oncorhynchus mykiss] >gb ACO08544.1  UPF0308 protein C9orf21 homolog [Oncorhynchus mykiss]	<a href="#">39.3</a>	39.3	93%	0.17	<a href="#">UG</a>
<a href="#">NP_079646.1</a>	hypothetical protein LOC66129 [Mus musculus] >sp Q9D1A0.1 C1021_MOUSE RecName: Full=UPF0308 protein C9orf21 homolog >dbj BAB22993.1  unnamed protein product [Mus musculus] >dbj BAE28518.1  unnamed protein product [Mus musculus] >gb AAI40306.1  RIKEN cDNA 1110018J18 gene [synthetic construct] >gb EDL16238.1  RIKEN cDNA 1110018J18, isoform CRA_b [Mus musculus] >gb AAI56632.1  RIKEN cDNA 1110018J18 gene [synthetic construct]	<a href="#">39.3</a>	39.3	89%	0.17	<a href="#">UGM</a>
<a href="#">ACI68733.1</a>	C1orf93 homolog [Salmo salar]	<a href="#">38.9</a>	38.9	93%	0.23	
<a href="#">ZP_01909602.1</a>	hypothetical protein PPSIR1_24939 [Plesiocystis pacifica SIR-1] >gb EDM77490.1  hypothetical protein PPSIR1_24939 [Plesiocystis pacifica SIR-1]	<a href="#">38.5</a>	38.5	95%	0.27	
<a href="#">XP_001829670.2</a>	hypothetical protein CC1G_11406 [Coprinopsis cinerea okayama7#130] >gb EAU92121.2  hypothetical protein CC1G_11406 [Coprinopsis cinerea okayama7#130]	<a href="#">38.5</a>	38.5	83%	0.29	<a href="#">G</a>
<a href="#">NP_001177306.1</a>	selenoprotein L [Ciona intestinalis]	<a href="#">38.5</a>	38.5	100%	0.32	<a href="#">G</a>
<a href="#">XP_003050915.1</a>	hypothetical protein NECHADRAFT_69625 [Nectria haematococca mpVI 77-13-4] >gb EEU45202.1  hypothetical protein NECHADRAFT_69625 [Nectria haematococca mpVI 77-13-4]	<a href="#">38.1</a>	38.1	83%	0.36	<a href="#">G</a>
<a href="#">XP_001800743.1</a>	hypothetical protein SNOG_10473 [Phaeosphaeria nodorum SN15] >gb EAT81867.1  hypothetical protein SNOG_10473 [Phaeosphaeria nodorum SN15]	<a href="#">37.7</a>	37.7	98%	0.49	<a href="#">G</a>
<a href="#">XP_002114328.1</a>	hypothetical protein TRIADDRAFT_58024 [Trichoplax adhaerens] >gb EDV23418.1  hypothetical protein TRIADDRAFT_58024 [Trichoplax adhaerens]	<a href="#">37.7</a>	37.7	65%	0.52	<a href="#">G</a>
<a href="#">XP_003007519.1</a>	conserved hypothetical protein [Verticillium albo-atrum VaMs.102] >gb EEY15598.1  conserved hypothetical protein [Verticillium albo-atrum VaMs.102]	<a href="#">37.4</a>	37.4	87%	0.58	<a href="#">G</a>
<a href="#">ACI70082.1</a>	C1orf93 homolog [Salmo salar]	<a href="#">37.0</a>	37.0	93%	0.77	



Accession	Description	Max score	Total score	Query coverage	E value	Links
<a href="#">YP_001547656.1</a>	hypothetical protein Haur_4898 [Herpetosiphon aurantiacus ATCC 23779] >gb ABX07528.1  hypothetical protein Haur_4898 [Herpetosiphon aurantiacus ATCC 23779]	<a href="#">37.0</a>	37.0	100%	0.81	<a href="#">G</a>
<a href="#">EFQ35083.1</a>	hypothetical protein GLRG_10227 [Glomerella graminicola M1.001]	<a href="#">36.6</a>	36.6	87%	1.2	
<a href="#">XP_001368977.1</a>	PREDICTED: similar to C9ORF21 [Monodelphis domestica]	<a href="#">36.2</a>	36.2	89%	1.3	<a href="#">UGM</a>
<a href="#">XP_002544130.1</a>	predicted protein [Uncinocarpus reesii 1704] >gb EEP78801.1  predicted protein [Uncinocarpus reesii 1704]	<a href="#">36.2</a>	36.2	87%	1.5	<a href="#">G</a>
<a href="#">XP_003070048.1</a>	hypothetical protein CPC735_032390 [Coccidioides posadasii C735 delta SOWgp] >gb EER27903.1  hypothetical protein CPC735_032390 [Coccidioides posadasii C735 delta SOWgp]	<a href="#">36.2</a>	36.2	80%	1.6	<a href="#">G</a>

## Alignments

>ref|NP\_001049763.1| [UG](#) Os03g0284600 [Oryza sativa Japonica Group]  
 gb|ABF95344.1| [G](#) UPF0308 protein, chloroplast precursor, putative, expressed [Oryza sativa Japonica Group]  
 dbj|BAF11677.1| [G](#) Os03g0284600 [Oryza sativa Japonica Group]  
 gb|EEE58829.1| [G](#) hypothetical protein OsJ\_10400 [Oryza sativa Japonica Group]  
 Length=251

**GFNF TD: 4332468 0<0300284600** | Os03g0284600 [Oryza sativa Japonica Group]  
 (10 or fewer PubMed links)

Score = 60.8 bits (146), Expect = 5e-08, Method: Compositional matrix adjust.  
 Identities = 36/87 (42%), Positives = 45/87 (52%), Gaps = 6/87 (6%)

```

Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          +V + RHFGCVLCR+ LAA++ A + L++IG GT EQA F
Sbjct 99  IVAFARHFGCVLCRKRADLLAAKQDAM-EAAGVALVLIGPGTVEQAKAFYDQTKF----- 152

Query 61  SIPVYTDPLRRSYLALRFRRGILSTFN 87
          VY DP SY AL F G+ STF
Sbjct 153  KGEVYADPSHSSYNALFAFGLFSTFT 179

```

>gb|EEC75002.1| hypothetical protein OsI\_11064 [Oryza sativa Indica Group]  
 Length=239

Score = 60.5 bits (145), Expect = 7e-08, Method: Compositional matrix adjust.  
 Identities = 36/87 (42%), Positives = 45/87 (52%), Gaps = 6/87 (6%)

```

Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          +V + RHFGCVLCR+ LAA++ A + L++IG GT EQA F
Sbjct 87  IVAFARHFGCVLCRKRADLLAAKQDAM-EAAGVALVLIGPGTVEQAKAFYDQTKF----- 140

Query 61  SIPVYTDPLRRSYLALRFRRGILSTFN 87
          VY DP SY AL F G+ STF
Sbjct 141  KGEVYADPSHSSYNALFAFGLFSTFT 167

```

>ref|NP\_001145525.1| [UG](#) hypothetical protein LOC100278941 [Zea mays]

gb|ACG48189.1| [G](#) hypothetical protein [Zea mays]  
 Length=258

**GFNF TD: 100278941 100100278941** | hypothetical protein LOC100278941 [Zea mays]  
 (10 or fewer PubMed links)

Score = 60.5 bits (145), Expect = 8e-08, Method: Compositional matrix adjust.  
Identities = 36/87 (42%), Positives = 45/87 (52%), Gaps = 7/87 (8%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVGCGTPEQAAKFRTDNLGPLPA 60
          +V + RHFGCVLCR+ LAA++ Q L++IG G+ EQA F G
Sbjct 107 VVAFARHFGCVLCRKRADLLAAKQDDM-QAAGVALVLIGPGSVEQAKAFEQTKFKG---- 161

Query 61  SIPVYTDPLRRSYLALRFRRGILSTFN 87
          VY DP SY AL F G+ STF
Sbjct 162 --EVYADPTHSSYDALEFAFGLFSTFT 186
```

>ref|XP\_002468056.1| **UG** hypothetical protein SORBIDRAFT\_01g038790 [Sorghum bicolor]  
gb|EER95054.1| **G** hypothetical protein SORBIDRAFT\_01g038790 [Sorghum bicolor]  
Length=258

**GENE ID: 8084887 SORBIDRAFT\_01g038790** | hypothetical protein [Sorghum bicolor]  
(10 or fewer PubMed links)

Score = 58.2 bits (139), Expect = 3e-07, Method: Compositional matrix adjust.  
Identities = 36/87 (42%), Positives = 44/87 (51%), Gaps = 6/87 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVGCGTPEQAAKFRTDNLGPLPA 60
          +V + RHFGCVLCR+ LAA+ Q Q L++IG G+ EQA F
Sbjct 106 VVAFARHFGCVLCRKRADLLAAK-QDVMQAAGVALVLIGPGSVEQAKAFCEQTKF----- 159

Query 61  SIPVYTDPLRRSYLALRFRRGILSTFN 87
          VY DP SY AL F G+ STF
Sbjct 160 KGEVYADPTHSSYDALEFAFGLFSTFT 186
```

>gb|ACR37670.1| unknown [Zea mays]  
Length=259

Score = 58.2 bits (139), Expect = 4e-07, Method: Compositional matrix adjust.  
Identities = 35/87 (41%), Positives = 44/87 (51%), Gaps = 6/87 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVGCGTPEQAAKFRTDNLGPLPA 60
          +V + RHFGCVLCR+ LAA++ Q Q L++IG G+ EQA F
Sbjct 107 VVAFARHFGCVLCRKRADLLAAKQDDM-QAAGVALVLIGPGSVEQAKAFCEQTKF----- 160

Query 61  SIPVYTDPLRRSYLALRFRRGILSTFN 87
          VY DP SY AL F G+ STF
Sbjct 161 KGEVYADPTHSSYDALEFAFGLFSTFT 187
```

>gb|ACN35248.1| **G** unknown [Zea mays]  
Length=224

**GENE ID: 100278941 LOC100278941** | hypothetical protein LOC100278941 [Zea mays]  
(10 or fewer PubMed links)

Score = 57.8 bits (138), Expect = 5e-07, Method: Compositional matrix adjust.  
Identities = 35/87 (41%), Positives = 44/87 (51%), Gaps = 6/87 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVGCGTPEQAAKFRTDNLGPLPA 60
          +V + RHFGCVLCR+ LAA++ Q L++IG G+ EQA F
Sbjct 107 VVAFARHFGCVLCRKRADLLAAKQDDM-QAAGVALVLIGPGSVEQAKAFCEQTKF----- 160

Query 61  SIPVYTDPLRRSYLALRFRRGILSTFN 87
          VY DP SY AL F G+ STF
Sbjct 161 KGEVYADPTHSSYDALEFAFGLFSTFT 187
```

>ref|XP\_002334170.1| **UG** predicted protein [Populus trichocarpa]  
gb|EEF07066.1| **G** predicted protein [Populus trichocarpa]  
Length=146

**GENE ID: 7497195 POPTRDRAFT\_941289** | hypothetical protein [Populus trichocarpa]  
(10 or fewer PubMed links)

Score = 57.4 bits (137), Expect = 6e-07, Method: Compositional matrix adjust.  
Identities = 33/86 (39%), Positives = 44/86 (52%), Gaps = 6/86 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVGCGTPEQAAKFRTDNLGPLPA 60
          +V + RHFGCVLCRR LAA++ L++IG G+ +QA F
Sbjct 46  VVAFARHFGCVLCRRRADYLAACKD-IMDASGVALVLIGPGSVDQAKTFSEQTKFKG--- 101
```



Query 61 SIPVYTDPLRRSYLALRFRRGILSTF 86  
 VY DP SY AL+F G+ +TF  
 Sbjct 102 --EVYADPSHSSYKALQFVSGVSTTF 125

>ref|XP\_002326159.1| **UG** predicted protein [Populus trichocarpa]  
**gb|EEE71829.1| G** predicted protein [Populus trichocarpa]  
 Length=200

**GFNF TD: 7459373 POPTRDRAFT\_589376** | hypothetical protein [Populus trichocarpa]  
 (10 or fewer PubMed links)

Score = 57.0 bits (136), Expect = 7e-07, Method: Compositional matrix adjust.  
 Identities = 33/86 (39%), Positives = 44/86 (52%), Gaps = 6/86 (6%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAQRERKWRMLMVGCGTPEQAQAFRTDNLGGLPA 60  
 +V + RHFGCVLCRR LAA++ L++IG G+ +QA F  
 Sbjct 49 VVAFARHFGCVLCRRRADYLAAKKD-IMDASGVALVLIGPGSVDAQTFSEQTKFKG--- 104

Query 61 SIPVYTDPLRRSYLALRFRRGILSTF 86  
 VY DP SY AL+F G+ +TF  
 Sbjct 105 --EVYADPSHSSYKALQFVSGVSTTF 128

>**gb|ACU20039.1|** unknown [Glycine max]  
 Length=256

Score = 56.6 bits (135), Expect = 1e-06, Method: Compositional matrix adjust.  
 Identities = 30/86 (35%), Positives = 45/86 (53%), Gaps = 6/86 (6%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAQRERKWRMLMVGCGTPEQAQAFRTDNLGGLPA 60  
 +V + RHFGCVLCR+ L++++ L++IG G+ +QA F +  
 Sbjct 104 VVAFARHFGCVLCRKRADYLSKKD-IMDASGVALVLIGPGSIDQAKSFAEKSKF----- 157

Query 61 SIPVYTDPLRRSYLALRFRRGILSTF 86  
 +Y DP SY AL F G+L+TF  
 Sbjct 158 EGEIYADPTHSSYEALNFVSGVLTTF 183

>ref|XP\_002965388.1| **UG** hypothetical protein SELMODRAFT\_68006 [Selaginella moellendorffii]  
**gb|EFJ34226.1| G** hypothetical protein SELMODRAFT\_68006 [Selaginella moellendorffii]  
 Length=172

**GENE ID: 9633660 SELMODRAFT\_68006** | hypothetical protein  
 [Selaginella moellendorffii]

Score = 55.5 bits (132), Expect = 2e-06, Method: Compositional matrix adjust.  
 Identities = 31/87 (36%), Positives = 47/87 (55%), Gaps = 6/87 (6%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAQRERKWRMLMVGCGTPEQAQAFRTDNLGGLPA 60  
 +V + RHFGC+LCR+ LA+++ F L+++G GT +QA F A+ P  
 Sbjct 21 VVAFARHFGCILCRKRADVLASKKEVF-DAAGVSLVLVGPVTVDQAKAF---ASQTQFP 76

Query 61 SIPVYTDPLRRSYLALRFRRGILSTFN 87  
 VY DP S+ A +F G + FN  
 Sbjct 77 E--VYADPTHASFDAFQFVSGASTIFN 101

>ref|XP\_002977235.1| **UG** hypothetical protein SELMODRAFT\_58056 [Selaginella moellendorffii]  
**gb|EFJ21844.1| G** hypothetical protein SELMODRAFT\_58056 [Selaginella moellendorffii]  
 Length=172

**GENE ID: 9638970 SELMODRAFT\_58056** | hypothetical protein  
 [Selaginella moellendorffii]

Score = 55.1 bits (131), Expect = 3e-06, Method: Compositional matrix adjust.  
 Identities = 31/87 (36%), Positives = 47/87 (55%), Gaps = 6/87 (6%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAQRERKWRMLMVGCGTPEQAQAFRTDNLGGLPA 60  
 +V + RHFGC+LCR+ LA+++ F L+++G GT +QA F A+ P  
 Sbjct 21 VVAFARHFGCILCRKRADVLASKKEVF-DGAGVSLVLVGPVTVDQAKAF---ASQTQFP 76

Query 61 SIPVYTDPLRRSYLALRFRRGILSTFN 87  
 VY DP S+ A +F G + FN  
 Sbjct 77 E--VYADPTHASFDAFQFVSGASTIFN 101

>gb|ABK23985.1| unknown [Picea sitchensis]  
Length=261

Score = 53.1 bits (126), Expect = 1e-05, Method: Compositional matrix adjust.  
Identities = 32/86 (38%), Positives = 45/86 (53%), Gaps = 6/86 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          ++ + RHFGCVLCR+ LA+++ L++IG G EQA F A P
Sbjct 109 VIGFARHFGCVLCRKRADVLASQKSQM-DAAGVALVLIGPGNIEQAKAF--ADQTKFPG 164

Query 61 SIPVYTDPLRRSYLALRFRRGILSTF 86
          I Y DP S+ AL+F G+ +TF
Sbjct 165 EI--YADPNHTSFNALKFVSGVFTTF 188
```

>ref|XP\_002273449.1| **UGM** PREDICTED: hypothetical protein [Vitis vinifera]  
emb|CBI32627.3| **G** unnamed protein product [Vitis vinifera]  
Length=254

**GENE ID: 100251717 LOC100251717** | hypothetical protein LOC100251717  
[Vitis vinifera] (10 or fewer PubMed links)

Score = 53.1 bits (126), Expect = 1e-05, Method: Compositional matrix adjust.  
Identities = 31/86 (37%), Positives = 42/86 (49%), Gaps = 6/86 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          +V + RHFGCV CR+ LA+++ L++IG G+ +QA F
Sbjct 102 VVAFARHFGCVFCRKRADLLASQKDRM-DASGVALVLIGPGSIDQAKAFSEQTNFKG--- 157

Query 61 SIPVYTDPLRRSYLALRFRRGILSTF 86
          VY DP SY L F G+LSTF
Sbjct 158 --EVYADPSHSSEYVLFVSGVLTSTF 181
```

>ref|XP\_002881499.1| **G** hypothetical protein ARALYDRAFT\_482716 [Arabidopsis lyrata subsp. lyrata]

gb|EFH57758.1| **G** hypothetical protein ARALYDRAFT\_482716 [Arabidopsis lyrata subsp. lyrata]  
Length=248

**GENE ID: 9315725 ARALYDRAFT\_482716** | hypothetical protein  
[Arabidopsis lyrata subsp. lyrata]

Score = 52.0 bits (123), Expect = 2e-05, Method: Compositional matrix adjust.  
Identities = 31/87 (36%), Positives = 41/87 (48%), Gaps = 6/87 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          +V + RHFGCVLCR+ LA++ L++IG G+ +QA F
Sbjct 96 VVAFARHFGCVLCRKRAYLAEKDVM-DASGVTLVLIGPGSIDQANTFMEQTKFKG--- 151

Query 61 SIPVYTDPLRRSYLALRFRRGILSTFN 87
          VY DP SY AL F G+ TF
Sbjct 152 --EVYADPNHASYEALFVSGVTVTFT 176
```

>ref|XP\_002525198.1| **G** conserved hypothetical protein [Ricinus communis]

gb|EEF37164.1| **G** conserved hypothetical protein [Ricinus communis]  
Length=249

**GENE ID: 8283839 RCOM\_0819880** | hypothetical protein [Ricinus communis]

Score = 51.6 bits (122), Expect = 3e-05, Method: Compositional matrix adjust.  
Identities = 30/86 (35%), Positives = 42/86 (49%), Gaps = 6/86 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          +V + RHFGCVLCR+ LAA++ L++IG G+ +QA F
Sbjct 97 VVAFARHFGCVLCRKRADYLAAKD-IMDASGVALVLIGPGSVDQAKTFSEQTKFKG--- 152

Query 61 SIPVYTDPLRRSYLALRFRRGILSTF 86
          VY D SY A +F G+ +TF
Sbjct 153 --EVYADTSHSSYEAFQFVSGVSTTF 176
```

>ref|NP\_030274.1| **UG** unknown protein [Arabidopsis thaliana]

sp|Q9ZUU2.2|U308\_ARATH **G** RecName: Full=UPF0308 protein At2g37240, chloroplastic; Flags: Precursor

**gb|AAK91362.1|** **G** At2g37240/F3G5.3 [Arabidopsis thaliana]  
**gb|AAC98045.2|** **G** expressed protein [Arabidopsis thaliana]  
**gb|AAM67197.1|** **G** unknown [Arabidopsis thaliana]  
**gb|AAP21145.1|** **G** At2g37240/F3G5.3 [Arabidopsis thaliana]  
 Length=248

**GENE ID: 818301 AT2G37240** | hypothetical protein [Arabidopsis thaliana]  
 (10 or fewer PubMed links)

Score = 50.8 bits (120), Expect = 5e-05, Method: Compositional matrix adjust.  
 Identities = 31/87 (36%), Positives = 41/87 (48%), Gaps = 6/87 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          +V + RHFGCVLCR+ LA ++ L++IG G+ +QA F
Sbjct 96  VVAFARHFGCVLCRKRAAYLAEKKDVM-DASGVALVLIGPGSIDQANTFVEQTKF----- 149

Query 61  SIPVYTDPLRRSYLALRFRRGILSTFN 87
          VY DP SY AL F G+ TF
Sbjct 150 KGEVYADPNHASYEALFVSGVSVTFT 176
```

>**ref|XP\_002866692.1|** **G** hypothetical protein ARALYDRAFT\_496819 [Arabidopsis lyrata subsp. lyrata]

**gb|EFH42951.1|** **G** hypothetical protein ARALYDRAFT\_496819 [Arabidopsis lyrata subsp. lyrata]  
 Length=265

**GENE ID: 9302764 ARALYDRAFT\_496819** | hypothetical protein  
 [Arabidopsis lyrata subsp. lyrata]

Score = 47.4 bits (111), Expect = 7e-04, Method: Compositional matrix adjust.  
 Identities = 33/89 (38%), Positives = 44/89 (50%), Gaps = 5/89 (5%)

```
Query 2  VIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGPLPAS 61
          +V + RHFGCV C L L + F +L+ +G GTP++A R AT P P
Sbjct 110 VVLLRHFGCVCCWELATALKEAKPRF-DAAGVKLIAVGVGTPDKA--RILATRLPFPME 165

Query 62  IPVYTDPLRRSYLALRFRRGILSTFNLPA 90
          +Y DP R++Y L G+ TF PA
Sbjct 166 C-LYADPERKAYDVLGLYGLGRTFFNPA 193
```

>**ref|XP\_002666723.1|** **UGM** PREDICTED: UPF0308 protein C9orf21 homolog [Danio rerio]  
 Length=223

**GENE ID: 100332791 LOC100332791** | UPF0308 protein C9orf21 homolog [Danio rerio]

Score = 47.0 bits (110), Expect = 8e-04, Method: Compositional matrix adjust.  
 Identities = 29/81 (36%), Positives = 39/81 (49%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          +VI+ RHF C C+ V++L 0 0+ RL+VIG + F +L P
Sbjct 55  IVIFVRHFLCYTCKEYVEDLGIQHVLDNSVRLVVIGQSSYSHIQGF---CSLTGFPH 111

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          I Y DP R+ Y L RRG
Sbjct 112 EI--YVDPERQIYKRLGLRRG 130
```

>**ref|NP\_201385.2|** **UG** unknown protein [Arabidopsis thaliana]

**gb|AAM20692.1|** **G** unknown protein [Arabidopsis thaliana]

**gb|AAN15652.1|** **G** unknown protein [Arabidopsis thaliana]  
 Length=275

**GENE ID: 836713 AT5G65840** | hypothetical protein [Arabidopsis thaliana]  
 (10 or fewer PubMed links)

Score = 47.0 bits (110), Expect = 0.001, Method: Compositional matrix adjust.  
 Identities = 33/89 (38%), Positives = 44/89 (50%), Gaps = 5/89 (5%)

```
Query 2  VIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGPLPAS 61
          +V + RHFGCV C L L + F +L+ +G GTP++A R AT P P
Sbjct 120 VVLLRHFGCVCCWELATALKEAKPRF-DAAGVKLIAVGVGTPDKA--RILATRLPFPME 175

Query 62  IPVYTDPLRRSYLALRFRRGILSTFNLPA 90
          +Y DP R++Y L G+ TF PA
```

Sbjct 176 C-LYADPERKAYDVLGLYFGLGRTFFNPA 203

>ref|XP\_002928763.1| **GM** PREDICTED: UPF0308 protein C9orf21-like, partial [Ailuropoda melanoleuca]  
Length=192

**GENE ID: 100474756 LOC100474756** | UPF0308 protein C9orf21-like [Ailuropoda melanoleuca]

Score = 45.8 bits (107), Expect = 0.002, Method: Compositional matrix adjust.  
Identities = 26/81 (33%), Positives = 41/81 (51%), Gaps = 5/81 (6%)

Query 1 LVIWPRHFGCVCRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60  
+V++ RHF C +C+ V++LA ++F QE L+VIG Q++ +  
Sbjct 25 VVVFVRHFLCYICKEYVEDLAKIPKSFLQEADVTLIVIG-----QSSYHHIEPFCKLTGY 79  
Query 61 SIPVYTDPLRRSYLALRFRRG 81  
S +Y DP R Y L +RG  
Sbjct 80 SHEIYVDPEREIYKRLGMKRG 100

>gb|EFB20490.1| hypothetical protein PANDA\_018799 [Ailuropoda melanoleuca]  
Length=186

Score = 45.8 bits (107), Expect = 0.002, Method: Compositional matrix adjust.  
Identities = 26/81 (33%), Positives = 41/81 (51%), Gaps = 5/81 (6%)

Query 1 LVIWPRHFGCVCRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60  
+V++ RHF C +C+ V++LA ++F QE L+VIG Q++ +  
Sbjct 19 VVVFVRHFLCYICKEYVEDLAKIPKSFLQEADVTLIVIG-----QSSYHHIEPFCKLTGY 73  
Query 61 SIPVYTDPLRRSYLALRFRRG 81  
S +Y DP R Y L +RG  
Sbjct 74 SHEIYVDPEREIYKRLGMKRG 94

>ref|XP\_002320012.1| **UG** predicted protein [Populus trichocarpa]

gb|EEE98327.1| **G** predicted protein [Populus trichocarpa]  
Length=199

**GENE ID: 7497309 POPTRDRAFT\_571999** | hypothetical protein [Populus trichocarpa]  
(10 or fewer PubMed links)

Score = 45.4 bits (106), Expect = 0.002, Method: Compositional matrix adjust.  
Identities = 31/85 (37%), Positives = 40/85 (48%), Gaps = 5/85 (5%)

Query 6 RHFGCVCRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPASIPVY 65  
RHFGC C L L ++ F +L+ IG GTP +A R A P P +Y  
Sbjct 48 RHFGCPCWELASSLKESEKDF-DSSGVKLIATIGVGTNKA---RLLAERLPFPMDC-LY 102  
Query 66 TDPLRRSYLALRFRRGILSTFNLPA 90  
DP R++Y L G+ TF PA  
Sbjct 103 ADPERKAYDVLGLYYGLGRTFFNPA 127

>ref|XP\_001493974.2| **UGM** PREDICTED: similar to UPF0308 protein C9orf21 [Equus caballus]  
Length=232

**GENE ID: 100062308 LOC100062308** | hypothetical protein LOC100062308 [Equus caballus]

Score = 45.4 bits (106), Expect = 0.003, Method: Compositional matrix adjust.  
Identities = 26/81 (33%), Positives = 41/81 (51%), Gaps = 5/81 (6%)

Query 1 LVIWPRHFGCVCRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60  
+V++ RHF C +C+ V++LA ++F QE L+VIG Q++ +  
Sbjct 65 VVVFMRHFLCYICKEYVEDLAKIPKSFLQEANVTIVIG-----QSSYHHIEPFCKLTGY 119  
Query 61 SIPVYTDPLRRSYLALRFRRG 81  
S +Y DP R Y L +RG  
Sbjct 120 SHEIYVDPEREIYKRLGMKRG 140

>ref|XP\_003130905.1| **UGM** PREDICTED: UPF0308 protein C9orf21 homolog isoform 2 [Sus scrofa]  
Length=184

**GENE ID: 100524271 LOC100524271** | UPF0308 protein C9orf21 homolog [Sus scrofa]

Score = 45.1 bits (105), Expect = 0.003, Method: Compositional matrix adjust.  
Identities = 29/89 (33%), Positives = 45/89 (51%), Gaps = 7/89 (7%)

```
Query 1  LVIWPRHFGVCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDATLGPLPA 60
          +V++ RHF C +C+ V++LA ++F QE L+VIG 0++ +
Sbjct 61  VVVFVRHFLCYICKEYVEDLAKIPKSFLQEANVTLIVIG-----QSSYHHIEPFCKLTGY 115

Query 61  SIPVYTDPLRRSYLALRFRRG--ILSTFN 87
          S +Y DP R Y L +RG I S+ N
Sbjct 116  SHEIYVDPEREIYKRLGMKRGEEIASSGN 144
```

>ref|XP\_002742741.1| **GM** PREDICTED: UPF0308 protein C9orf21-like [Callithrix jacchus]  
Length=I86

**GENE ID: 100415410 LOC100415410** | UPF0308 protein C9orf21-like  
[Callithrix jacchus]

Score = 45.1 bits (105), Expect = 0.003, Method: Compositional matrix adjust.  
Identities = 26/81 (33%), Positives = 42/81 (52%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGVCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDATLGPLPA 60
          +V++ RHF C +C+ V++LA ++F QE L+VIG 0++ ++
Sbjct 63  VVVFVRHFLCYICKEYVEDLARIPKSFLQEANVTLIVIG-----QSSYHHIESFCKLTGY 117

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 118  SHEIYVDPEREIYKRLGMKRG 138
```

>ref|NP\_001069904.1| **UGM** hypothetical protein LOC616897 [Bos taurus]  
sp|Q148E0.1|CI021 BOVIN **G** RecName: Full=UPF0308 protein C9orf21 homolog  
gb|AAI18426.1| **G** Chromosome 9 open reading frame 21 ortholog [Bos taurus]  
gb|DAA26600.1| **G** hypothetical protein LOC616897 [Bos taurus]  
Length=228

**GENE ID: 616897 C8H9orf21** | chromosome 9 open reading frame 21 ortholog  
[Bos taurus] (10 or fewer PubMed links)

Score = 45.1 bits (105), Expect = 0.003, Method: Compositional matrix adjust.  
Identities = 26/81 (33%), Positives = 41/81 (51%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGVCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDATLGPLPA 60
          +V++ RHF C +C+ V++LA ++F QE L+VIG 0++ +
Sbjct 61  IVVFVRHFLCYICKEYVEDLAKIPKSFLQEANVTLIVIG-----QSSYHHIEPFCKLTGY 115

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 116  SHEIYVDPEREIYKRLGMKRG 136
```

>ref|XP\_001153024.1| **UGM** PREDICTED: similar to TPA\_exp: C90RF21 isoform 1 [Pan troglod  
Length=I82

**GENE ID: 465250 LOC465250** | similar to TPA\_exp: C90RF21 [Pan troglodytes]

Score = 45.1 bits (105), Expect = 0.003, Method: Compositional matrix adjust.  
Identities = 29/89 (33%), Positives = 45/89 (51%), Gaps = 7/89 (7%)

```
Query 1  LVIWPRHFGVCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDATLGPLPA 60
          +V++ RHF C +C+ V++LA ++F QE L+VIG 0++ +
Sbjct 59  VVVFVRHFLCYICKEYVEDLAKIPKSFLQEANVTLIVIG-----QSSYHHIEPFCKLTGY 113

Query 61  SIPVYTDPLRRSYLALRFRRG--ILSTFN 87
          S +Y DP R Y L +RG I S+ N
Sbjct 114  SHEIYVDPEREIYKRLGMKRGEEIASSGN 142
```

>ref|XP\_003130904.1| **UGM** PREDICTED: UPF0308 protein C9orf21 homolog isoform 1 [Sus scr  
Length=228

**GENE ID: 100524271 LOC100524271** | UPF0308 protein C9orf21 homolog [Sus scrofa]

Score = 44.7 bits (104), Expect = 0.004, Method: Compositional matrix adjust.  
Identities = 26/81 (33%), Positives = 41/81 (51%), Gaps = 5/81 (6%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60  
 +V++ RHF C +C+ V++LA ++F QE L+VIG Q++ +  
 Sbjct 61 VVVFVRHFLCYICKEYVEDLAKIPKSFLQEANVTLIVIG-----QSSYHHIEPFCKLTGY 115

Query 61 SIPVYTDPLRRSYLALRFRRG 81  
 S +Y DP R Y L +RG  
 Sbjct 116 SHEIYVDPEREIYKRLGMKRG 136

>gb|EAW92661.1| **G** chromosome 9 open reading frame 21, isoform CRA\_a [Homo sapiens]  
 Length=214

**GENE ID: 195827 C9orf21** | chromosome 9 open reading frame 21 [Homo sapiens]  
 (10 or fewer PubMed links)

Score = 44.7 bits (104), Expect = 0.004, Method: Compositional matrix adjust.  
 Identities = 30/93 (33%), Positives = 48/93 (52%), Gaps = 7/93 (7%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60  
 +V++ RHF C +C+ V++LA ++F QE L+VIG Q++ +  
 Sbjct 59 VVVFVRHFLCYICKEYVEDLAKIPRSFLQEANVTLIVIG-----QSSYHHIEPFCKLTGY 113

Query 61 SIPVYTDPLRRSYLALRFRRG--ILSTFNLPAL 91  
 S +Y DP R Y L +RG I S+ +L +L  
 Sbjct 114 SHEIYVDPEREIYKRLGMKRGEEIASSGSLQSL 146

>ref|XP\_001588758.1| **G** hypothetical protein SS1G\_10305 [Sclerotinia sclerotiorum 1980]  
 gb|EDN94432.1| **G** hypothetical protein SS1G\_10305 [Sclerotinia sclerotiorum 1980]  
 Length=342

**GENE ID: 5484650 SS1G\_10305** | hypothetical protein  
 [Sclerotinia sclerotiorum 1980 UF-70]

Score = 44.7 bits (104), Expect = 0.004, Method: Composition-based stats.  
 Identities = 29/91 (32%), Positives = 46/91 (51%), Gaps = 12/91 (13%)

Query 1 LVIWPRHFGCVLCRRLVKELAA--REQAFROERKWRLMVIGCGTPEQAAKFRTDNLG 57  
 L+I+ RHF C C+ ++ LAA + + + V+GCG+P ++ +AT P  
 Sbjct 87 LIIFIRHFFCGNCQEFRLTLAASITEDSLQLHTPTFIAVVGCGSPSLIPMYQ-EATKCP 145

Query 58 LPASIPVYTDPLRRSYLALRFRRGILSTFNL 88  
 P+Y DP R+ Y L G++ T NL  
 Sbjct 146 F----PIYADPTRKLYDEL----GMMRTLNL 168

>ref|XP\_002820049.1| **GM** PREDICTED: UPF0308 protein C9orf21-like [Pongo abelii]  
 Length=226

**GENE ID: 100438036 LOC100438036** | UPF0308 protein C9orf21-like [Pongo abelii]

Score = 44.7 bits (104), Expect = 0.004, Method: Compositional matrix adjust.  
 Identities = 26/81 (33%), Positives = 41/81 (51%), Gaps = 5/81 (6%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60  
 +V++ RHF C +C+ V++LA ++F QE L+VIG Q++ +  
 Sbjct 59 VVVFVRHFLCYICKEYVEDLAKIPKSFLQEANVTLIVIG-----QSSYHHIEPFCKLTGY 113

Query 61 SIPVYTDPLRRSYLALRFRRG 81  
 S +Y DP R Y L +RG  
 Sbjct 114 SHEIYVDPEREIYKRLGMKRG 134

>ref|XP\_001116318.1| **GM** PREDICTED: UPF0308 protein C9orf21-like [Macaca mulatta]  
 Length=140

**GENE ID: 720662 LOC720662** | UPF0308 protein C9orf21-like [Macaca mulatta]

Score = 44.7 bits (104), Expect = 0.004, Method: Compositional matrix adjust.  
 Identities = 25/75 (34%), Positives = 37/75 (50%), Gaps = 5/75 (6%)

Query 7 HFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPASIPVYT 66  
 HF C +C+ V++LA ++F QE L+VIG Q++ + P S +Y  
 Sbjct 65 HFLCYICKEYVEDLAKIPESFLQEANVTLIVIG-----QSSYHHIEPFCKPTGYSHEIYV 119

Query 67 DPLRRSYLALRFRRG 81  
 DP R Y L +RG  
 Sbjct 120 DPEREMYKRLGMKRG 134



>ref|XP\_520707.2| **UGM** PREDICTED: similar to TPA\_exp: C90RF21 isoform 2 [Pan troglodyte  
Length=226

**GENE ID: 465250 LOC465250** | similar to TPA\_exp: C90RF21 [Pan troglodytes]

Score = 44.7 bits (104), Expect = 0.004, Method: Compositional matrix adjust.  
Identities = 26/81 (33%), Positives = 41/81 (51%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          +V++ RHF C +C+ V++LA ++F QE L+VIG Q++ +
Sbjct 59  VVVFVRHFLCYICKEYVEDLAKIPKSFLEQANVTLVIG-----QSSYHHIEPFCKLTGY 113

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 114 SHEIYVDPEREIYKRLGMKRG 134
```

>gb|EAW92663.1| **G** chromosome 9 open reading frame 21, isoform CRA\_c [Homo sapiens]  
Length=182

**GENE ID: 195827 C9orf21** | chromosome 9 open reading frame 21 [Homo sapiens]  
(10 or fewer PubMed links)

Score = 44.3 bits (103), Expect = 0.005, Method: Compositional matrix adjust.  
Identities = 29/89 (33%), Positives = 45/89 (51%), Gaps = 7/89 (7%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          +V++ RHF C +C+ V++LA ++F QE L+VIG Q++ +
Sbjct 59  VVVFVRHFLCYICKEYVEDLAKIPRSFLEQANVTLVIG-----QSSYHHIEPFCKLTGY 113

Query 61  SIPVYTDPLRRSYLALRFRRG--ILSTFN 87
          S +Y DP R Y L +RG I S+ N
Sbjct 114 SHEIYVDPEREIYKRLGMKRGEIASSGN 142
```

>ref|ZP\_02926219.1| hypothetical protein VspiD\_06235 [Verrucomicrobium spinosum DSM  
4136]  
Length=304

Score = 44.3 bits (103), Expect = 0.006, Method: Compositional matrix adjust.  
Identities = 30/84 (36%), Positives = 38/84 (46%), Gaps = 5/84 (5%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          LV++ RH GC CR + ++A A R+ ++ G P A F L LPA
Sbjct 162 LVVFLRHAGCTFCREALADIARVRPAIEASGT-RIALVHMGEPAFAAFSGQYGLADLPA 220

Query 61  SIPVYTDPLRRSYLALRFRRGILS 84
          DP RR Y L RRG LS
Sbjct 221 ---VADPSRRLYRGLGLRRGKLS 240
```

>ref|NP\_714542.1| **UGM** hypothetical protein LOC195827 [Homo sapiens]

sp|Q7RTV5.1|CI021 **HUMAN G** RecName: Full=UPF0308 protein C9orf21

tpg|DAA00065.1| **G** TPA\_exp: C90RF21 [Homo sapiens]

emb|CAI40534.1| **G** novel protein [Homo sapiens]

gb|EAW92662.1| **G** chromosome 9 open reading frame 21, isoform CRA\_b [Homo sapiens]

gb|AAI36504.1| **G** Chromosome 9 open reading frame 21 [Homo sapiens]  
Length=226

**GENE ID: 195827 C9orf21** | chromosome 9 open reading frame 21 [Homo sapiens]  
(10 or fewer PubMed links)

Score = 43.9 bits (102), Expect = 0.006, Method: Compositional matrix adjust.  
Identities = 26/81 (33%), Positives = 41/81 (51%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          +V++ RHF C +C+ V++LA ++F QE L+VIG Q++ +
Sbjct 59  VVVFVRHFLCYICKEYVEDLAKIPRSFLEQANVTLVIG-----QSSYHHIEPFCKLTGY 113

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 114 SHEIYVDPEREIYKRLGMKRG 134
```

>gb|ACU20324.1| unknown [Glycine max]

Length=251

Score = 43.9 bits (102), Expect = 0.007, Method: Compositional matrix adjust.  
Identities = 29/85 (35%), Positives = 39/85 (46%), Gaps = 5/85 (5%)

```
Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDATLGPLPASIPVY 65
          RHFGC+ C L L F +L+ +G GTP +A R A P P +Y
Sbjct 100 RHFGCICCFWELASALKESKARF-DSAGIKLIAVGVGTPNKA---RILAERLPPFMD-LY 154

Query 66 TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y L G+ TF PA
Sbjct 155 ADPDRKAYNVNLNYFGLGRTFLNPA 179
```

>**gb|ACN32054.1|** **G** unknown [Zea mays]  
Length=258

**GFNF TD: 100275311 LOC100275311** | hypothetical protein LOC100275311 [Zea mays]  
(10 or fewer PubMed links)

Score = 43.5 bits (101), Expect = 0.008, Method: Compositional matrix adjust.  
Identities = 31/85 (37%), Positives = 38/85 (45%), Gaps = 5/85 (5%)

```
Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDATLGPLPASIPVY 65
          RHFGC C L L F +L+ IG GTPE+A R A P P +Y
Sbjct 105 RHFGCFCCWELASVLKDAMAEFDSAGA-KLIAIGVGTPEKA---RILADRLPPFMD-SLY 159

Query 66 TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y L G+ T PA
Sbjct 160 ADPERKAYSVLGLYHGLGRTLFSFA 184
```

>**ref|XP\_001106503.1|** **UGM** PREDICTED: UPF0308 protein C9orf21-like [Macaca mulatta]  
Length=226

**GENE ID: 710488 LOC710488** | similar to UPF0308 protein C9orf21 [Macaca mulatta]

Score = 43.5 bits (101), Expect = 0.008, Method: Compositional matrix adjust.  
Identities = 26/81 (33%), Positives = 41/81 (51%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDATLGPLPA 60
          +V++ RHF C +C+ V++LA ++F QE L+VIG Q++ +
Sbjct 59 VVVFVRHFLCYICKEYVEDLAKIPKSFLQEANVTLIVIG-----QSSYHHIEPFCLRTGY 113

Query 61 SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 114 SHEIYVDPEREIYKRLGMKRG 134
```

>**gb|ADE77692.1|** unknown [Picea sitchensis]  
Length=276

Score = 43.5 bits (101), Expect = 0.008, Method: Composition-based stats.  
Identities = 31/85 (37%), Positives = 38/85 (45%), Gaps = 5/85 (5%)

```
Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDATLGPLPASIPVY 65
          RHFGC C L L F +L+ IG GTPE+A R P P +Y
Sbjct 125 RHFGCPCCWFASTLKDVMKPF-DSAGVKLIAIGVGTPEKA---RILGERLPPFLD-SLY 179

Query 66 TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y AL G+ TF PA
Sbjct 180 ADPDRKAYDALGLYYGLGRTFFNPA 204
```

>**emb|CBI33071.3|** unnamed protein product [Vitis vinifera]  
Length=159

Score = 43.5 bits (101), Expect = 0.009, Method: Compositional matrix adjust.  
Identities = 30/85 (36%), Positives = 40/85 (48%), Gaps = 5/85 (5%)

```
Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDATLGPLPASIPVY 65
          RHFGC C L L ++ F +L+ +G GTP++A R A P P +Y
Sbjct 8  RHFGCPCCWDLASALKESKERF-DSAGVKLIAVGVGTPDKA---RILAERLPPFLDC-LY 62

Query 66 TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y L G TF PA
Sbjct 63 ADPDRKAYDVLGLYYGFGRTFFNPA 87
```

>ref|XP\_002708298.1| **UGM** PREDICTED: hypothetical protein [Oryctolagus cuniculus]  
Length=226

**GENE ID: 100342762 LOC100342762** | hypothetical protein LOC100342762  
[Oryctolagus cuniculus]

Score = 43.5 bits (101), Expect = 0.009, Method: Compositional matrix adjust.  
Identities = 26/81 (33%), Positives = 41/81 (51%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDNLGPLPA 60
          +V++ RHF C +C+ V++LA 0+F +E L+VIG 0++ +
Sbjct 59  VVVFVRHFLCYVCKEYVEDLAKVPQSFLREADVTLIVIG-----QSSYHHIEPFCKLTGY 113

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 114  SHEIYVDPEREIYKRLGMKRG 134
```

>ref|XP\_002263922.1| **UGM** PREDICTED: hypothetical protein [Vitis vinifera]  
**emb|CAN81556.1| G** hypothetical protein VITISV\_040398 [Vitis vinifera]  
Length=256

**GENE ID: 100256614 LOC100256614** | hypothetical protein [Vitis vinifera]  
(10 or fewer PubMed links)

Score = 43.5 bits (101), Expect = 0.010, Method: Compositional matrix adjust.  
Identities = 30/85 (36%), Positives = 40/85 (48%), Gaps = 5/85 (5%)

```
Query 6  RHFVCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFVC C L L ++ F +L+ +G GTP++A R A P P +Y
Sbjct 105  RHFVCPCWDLASALKESKERF-DSAGVKLIAVGVGTPDKA---RILAERLPFPLDC-LY 159

Query 66  TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y L G TF PA
Sbjct 160  ADPDRKAYDVLGLYYGFGRTFFNPA 184
```

>emb|CBI33293.3| unnamed protein product [Vitis vinifera]  
Length=256

Score = 43.1 bits (100), Expect = 0.011, Method: Compositional matrix adjust.  
Identities = 30/85 (36%), Positives = 40/85 (48%), Gaps = 5/85 (5%)

```
Query 6  RHFVCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFVC C L L ++ F +L+ +G GTP++A R A P P +Y
Sbjct 105  RHFVCPCWDLASALKESKERF-DSAGVKLIAVGVGTPDKA---RILAERLPFPLDC-LY 159

Query 66  TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y L G TF PA
Sbjct 160  ADPDRKAYDVLGLYYGFGRTFFNPA 184
```

>ref|XP\_002463941.1| **UG** hypothetical protein SORBIDRAFT\_01g009340 [Sorghum bicolor]  
**gb|EER90939.1| G** hypothetical protein SORBIDRAFT\_01g009340 [Sorghum bicolor]  
Length=259

**GENE ID: 8060536 SORBIDRAFT\_01g009340** | hypothetical protein [Sorghum bicolor]  
(10 or fewer PubMed links)

Score = 43.1 bits (100), Expect = 0.011, Method: Compositional matrix adjust.  
Identities = 31/85 (37%), Positives = 38/85 (45%), Gaps = 5/85 (5%)

```
Query 6  RHFVCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFVC C L L F +L+ IG GTPE+A R A P P +Y
Sbjct 108  RHFVCFCWELASVLKDSIAKFDSDAGA-KLIAIGVGTPEKA---RILADRLPFPLD-SLY 162

Query 66  TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y L G+ T PA
Sbjct 163  ADPERKAYNVLGLYHGLGRTLFNPA 187
```




>ref|XP\_001547900.1| **G** hypothetical protein BC1G\_13584 [Botryotinia fuckeliana B05.10]  
**gb|EDN19735.1| G** hypothetical protein BC1G\_13584 [Botryotinia fuckeliana B05.10]  
Length=364

**GENE ID: 5428362 BC1G\_13584** | hypothetical protein  
[Botryotinia fuckeliana B05.10]

Score = 42.7 bits (99), Expect = 0.014, Method: Composition-based stats.  
Identities = 29/93 (32%), Positives = 46/93 (50%), Gaps = 12/93 (12%)

```
Query 1  LVIWPRHFGCVLCRRLVKELA---AREQAFRQERKWRMLVIGCGTPEQAAKFRTDNLGP 57
          LVI+ RHF C C+ ++ LA + + + V+GCG+P ++ +AT P
Sbjct 110 LVIFIRHFFCGNCQEYLRTLATSITEDSLQLHTPTFIAVVGCGSPSLIPMYQ-EATNCP 168

Query 58 LPASIPVYTDPLRRSYLALRFRRGILSTFNLPA 90
          P+Y DP ++ Y L G++ T NL A
Sbjct 169 F----PIYADPTKKLYDEL----GMMRTLNLGA 193
```

>ref|XP\_681257.1|  hypothetical protein AN7988.2 [Aspergillus nidulans FGSC A4]  
gb|EAA58791.1|  hypothetical protein AN7988.2 [Aspergillus nidulans FGSC A4]  
tpe|CBF73658.1|  TPA: conserved hypothetical protein [Aspergillus nidulans FGSC A4]  
Length=226

**GENE ID: 2869289 AN7988.2** | hypothetical protein [Aspergillus nidulans FGSC A4]  
(10 or fewer PubMed links)

Score = 42.7 bits (99), Expect = 0.015, Method: Compositional matrix adjust.  
Identities = 26/96 (28%), Positives = 47/96 (49%), Gaps = 8/96 (8%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAR-----EQAFRQERKWRMLVIGCGTPEQAAKFRTDAT 54
          ++I+ RHF C+ + V+ ++ + + +LM+IGCG P + ++ T
Sbjct 49  VIIFIRHFFCIYDQDYVRTVSHHLTDSVLQTISKTG GPIQLMIIGCGDPSLIVPVYSETT 108






Query 55 LGPLPASIPVYTDPLRRSYLALRFRRGILSTFNLPA 90
          G + A PVYTD P + Y L +R + + + P
Sbjct 109 -GDI-AQFPVYTDPDGKLYEKLHMKRTLNTIMHPPV 142
```

>gb|EAY91739.1| hypothetical protein OsI\_13380 [Oryza sativa Indica Group]  
Length=259

Score = 42.7 bits (99), Expect = 0.017, Method: Compositional matrix adjust.  
Identities = 30/85 (36%), Positives = 38/85 (45%), Gaps = 5/85 (5%)

```
Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRMLVIGCGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFGC C L L F +L+ IG GTP++A R A P P +Y
Sbjct 111 RHFGCFCCWELASVLKESMAKFDAAGA-KLIAIGVGTDPKA---RILADGLPFPVD-SLY 165

Query 66 TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y L G+ T PA
Sbjct 166 ADPERKAYDVLGLYHGLGRTLISPA 190
```

>ref|NP\_001051154.1|  Os03g0729300 [Oryza sativa Japonica Group]  
gb|AA038466.1|  unknown protein [Oryza sativa Japonica Group]  
gb|ABF98681.1|  UPF0308 protein, chloroplast precursor, putative, expressed [Oryza sativa Japonica Group]  
dbj|BAF13068.1|  Os03g0729300 [Oryza sativa Japonica Group]  
dbj|BAG98418.1|  unnamed protein product [Oryza sativa Japonica Group]  
Length=259

**GENE ID: 4333988 Os03g0729300** | Os03g0729300 [Oryza sativa Japonica Group]  
(10 or fewer PubMed links)

Score = 42.7 bits (99), Expect = 0.017, Method: Compositional matrix adjust.  
Identities = 30/85 (36%), Positives = 38/85 (45%), Gaps = 5/85 (5%)

```
Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRMLVIGCGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFGC C L L F +L+ IG GTP++A R A P P +Y
Sbjct 111 RHFGCFCCWELASVLKESMAKFDAAGA-KLIAIGVGTDPKA---RILADGLPFPVD-SLY 165

Query 66 TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y L G+ T PA
Sbjct 166 ADPERKAYDVLGLYHGLGRTLISPA 190
```

>gb|ABF98680.1| expressed protein [Oryza sativa Japonica Group]  
dbj|BAG90239.1| unnamed protein product [Oryza sativa Japonica Group]  
Length=192

Score = 42.7 bits (99), Expect = 0.017, Method: Compositional matrix adjust.

Identities = 31/85 (37%), Positives = 41/85 (49%), Gaps = 5/85 (5%)

```
Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFGC  C  L  L  ++ F  +L+ +G GTP++A  R  A  P  P  +Y
Sbjct 107 RHFGCPCCWELASVLRDTERKRF-DSAGVKLIAVGVGTPDKA---RILAERLPFPLDY-LY 161

Query 66  TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y  L  GI  TF  PA
Sbjct 162  ADPERKAYDLLGLYFGIGRTFFNPA 186
```

>ref|XP\_002191279.1| **UGM** PREDICTED: hypothetical protein [Taeniopygia guttata]  
Length=222

**GENE ID: 100224761 LOC100224761** | hypothetical protein LOC100224761  
[Taeniopygia guttata]

Score = 42.4 bits (98), Expect = 0.018, Method: Compositional matrix adjust.  
Identities = 26/81 (33%), Positives = 40/81 (50%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDNLGPLPA 60
          +V++ R+F C  C+  V++LA  +AF OE  RL+VIG  +  F  +
Sbjct 55  IVLFVRNFLCYTCKEYVEDLAKVPKAFQESNVRLIVIGQSSYHHIKPFCSLTGY----- 109

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          +  +Y DP R  Y  L  +RG
Sbjct 110  THEMVDPPREIYKILGMKRG 130
```

>gb|EEE59859.1| hypothetical protein OsJ\_12441 [Oryza sativa Japonica Group]  
Length=283

Score = 42.4 bits (98), Expect = 0.019, Method: Compositional matrix adjust.  
Identities = 30/85 (36%), Positives = 38/85 (45%), Gaps = 5/85 (5%)

```
Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFGC  C  L  L  F  +L+ IG GTP++A  R  A  P  P  +Y
Sbjct 111 RHFGCFCCWELASVLKESMAKFDAAGA-KLIAIGVGVGTPDKA---RILADGLPFPVD-SLY 165

Query 66  TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y  L  G+  T  PA
Sbjct 166  ADPERKAYDVLGLYHGLGRTLISPA 190
```

>ref|NP\_001051153.1| **UG** Os03g0729200 [Oryza sativa Japonica Group]  
**gb|AA038464.1|** **G** hypothetical protein [Oryza sativa Japonica Group]  
**gb|ABF98679.1|** **G** expressed protein [Oryza sativa Japonica Group]  
**dbj|BAF13067.1|** **G** Os03g0729200 [Oryza sativa Japonica Group]  
**gb|EAY91738.1|** hypothetical protein OsI\_13379 [Oryza sativa Indica Group]  
**dbj|BAG94118.1|** **G** unnamed protein product [Oryza sativa Japonica Group]  
**gb|EEE59858.1|** **G** hypothetical protein OsJ\_12440 [Oryza sativa Japonica Group]  
Length=258

**GENE ID: 4333987 Os03g0729200** | Os03g0729200 [Oryza sativa Japonica Group]  
(10 or fewer PubMed links)

Score = 42.4 bits (98), Expect = 0.019, Method: Compositional matrix adjust.  
Identities = 31/85 (37%), Positives = 41/85 (49%), Gaps = 5/85 (5%)

```
Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFGC  C  L  L  ++ F  +L+ +G GTP++A  R  A  P  P  +Y
Sbjct 107 RHFGCPCCWELASVLRDTERKRF-DSAGVKLIAVGVGTPDKA---RILAERLPFPLDY-LY 161

Query 66  TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y  L  GI  TF  PA
Sbjct 162  ADPERKAYDLLGLYFGIGRTFFNPA 186
```

>ref|XP\_848380.1| **UGM** PREDICTED: similar to UPF0308 protein C9orf21 [Canis familiaris]  
Length=230

**GENE ID: 606835 LOC606835** | similar to UPF0308 protein C9orf21  
[Canis lupus familiaris]

Score = 42.4 bits (98), Expect = 0.019, Method: Compositional matrix adjust.  
Identities = 25/81 (31%), Positives = 40/81 (50%), Gaps = 5/81 (6%)

```

Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          +V++ RHF C +C+ V++LA ++ QE L+VIG Q++ +
Sbjct 63  VVVFVRHFLCYICKEYVEDLAKIPKSVLQEADITLVIG-----QSSYHHIEPFCKLTGY 117

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 118  SHEIYVDPEREIYKRLGMKRG 138

```

>ref|XP\_002463942.1| **UG** hypothetical protein SORBIDRAFT\_01g009350 [Sorghum bicolor]  
**gb|EER90940.1| G** hypothetical protein SORBIDRAFT\_01g009350 [Sorghum bicolor]  
 Length=260

**GENE ID: 8059958 SORBDRAFT\_01g009350** | hypothetical protein [Sorghum bicolor]  
 (10 or fewer PubMed links)

Score = 42.0 bits (97), Expect = 0.023, Method: Compositional matrix adjust.  
 Identities = 31/85 (37%), Positives = 40/85 (48%), Gaps = 5/85 (5%)

```

Query 6  RHFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFGC C L L + F +L+ +G GTP AK R A P P +Y
Sbjct 109 RHFGCPCWELASVLRDRTKEKF-DSAGVKLIAVGVGTP---AKARILAERLPPFLEY-LY 163

Query 66  TDPLRRSYLALRFRRGILSTFNLP 90
          DP R++Y L G+ TF PA
Sbjct 164 ADPDRKAYNLLGLYFGVGRFFNPA 188

```

>ref|NP\_001102584.1| **UG** hypothetical protein LOC498685 [Rattus norvegicus]  
**gb|EDL84433.1| G** similar to UPF0308 protein C9orf21, isoform CRA\_a [Rattus norvegicus]  
 Length=147

**GENE ID: 498685 LOC498685** | similar to UPF0308 protein C9orf21  
 [Rattus norvegicus]

Score = 41.6 bits (96), Expect = 0.032, Method: Compositional matrix adjust.  
 Identities = 25/81 (31%), Positives = 40/81 (50%), Gaps = 5/81 (6%)

```

Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPA 60
          +V++ RHF C +C+ V++LA ++ QE L+VIG Q++ +
Sbjct 59  VVVFVRHFLCYVCKEYVEDLAKIPKSVLQEADITLVIG-----QSSYHHIEPFCKLTGY 113

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 114  SHEIYVDPEREIYKRLGMKRG 134

```

>ref|XP\_002269002.1| **UGM** PREDICTED: hypothetical protein, partial [Vitis vinifera]  
 Length=223

**GENE ID: 100245543 LOC100245543** | hypothetical protein LOC100245543  
 [Vitis vinifera]

Score = 41.6 bits (96), Expect = 0.033, Method: Compositional matrix adjust.  
 Identities = 28/85 (33%), Positives = 37/85 (44%), Gaps = 5/85 (5%)

```

Query 6  RHFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFGC C L L + F +L+ +G GTP +A R A P P +Y
Sbjct 72  RHFGCFCCWELASALKESKATF-DSAGVKLIAVGVGTPNKACIL---AERLPPFMDL-LY 126

Query 66  TDPLRRSYLALRFRRGILSTFNLP 90
          DP R++Y L G+ T PA
Sbjct 127 ADPDRKAYDVLGLYYGLSRTLFSFA 151

```

>**gb|ACU18761.1|** unknown [Glycine max]  
 Length=251

Score = 41.6 bits (96), Expect = 0.034, Method: Compositional matrix adjust.  
 Identities = 29/85 (35%), Positives = 38/85 (45%), Gaps = 5/85 (5%)

```

Query 6  RHFGCVLCRRLVKELAAREQAFROERKWRLMVIGCGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFGC C L L + F +L+ +G GTP +A R A P P +Y
Sbjct 100 RHFGCPCWELASALKESKARF-DSAGVKLIAVGIPTNKA---RMLAERLPPFLDC-LY 154

Query 66  TDPLRRSYLALRFRRGILSTFNLP 90
          DP R++Y L G+ T P+
Sbjct 155 ADPDRKAYHVLNLYYGFGRFFNPS 179

```




>**gb|ACU24130.1** | unknown [Glycine max]  
Length=251

Score = 41.6 bits (96), Expect = 0.035, Method: Compositional matrix adjust.  
Identities = 29/85 (35%), Positives = 38/85 (45%), Gaps = 5/85 (5%)

```
Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFGC C L L + F +L+ +G GTP +A R A P P +Y
Sbjct 100 RHFGCPCCWELASALKESKARF-DSAGVKLIAVGIGTPNKA---RMLAERLPFPLDC-LY 154

Query 66  TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y L G TF P+
Sbjct 155 ADPDRKAYHVLNLYYGFGRFFNPS 179
```

>**gb|EDL84435.1** |  similar to UPF0308 protein C9orf21, isoform CRA\_c [Rattus norvegicus]  
Length=226

**GENE ID: 498685 LOC498685** | similar to UPF0308 protein C9orf21  
[Rattus norvegicus]

Score = 41.6 bits (96), Expect = 0.036, Method: Compositional matrix adjust.  
Identities = 25/81 (31%), Positives = 40/81 (50%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFROERKWRLMVICGTPEQAAKFRTDNLGPLPA 60
          +V++ RHF C +C+ V++LA ++ QE L+VIG 0++ +
Sbjct 59  VVVFVRHFLCYVCKEYVEDLAKIPKSVLQEADVTLIVIG-----QSSYHHIEPFCKLTGY 113


Query 61  SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 114 SHEIYVDPEREIYKRLGMKRG 134
```


>**emb|CBI33289.3** | unnamed protein product [Vitis vinifera]  
Length=255

Score = 41.6 bits (96), Expect = 0.039, Method: Compositional matrix adjust.  
Identities = 28/85 (33%), Positives = 37/85 (44%), Gaps = 5/85 (5%)

```
Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDNLGPLPASIPVY 65
          RHFGC C L L + F +L+ +G GTP +A A P P +Y
Sbjct 104 RHFGCFCCWELASALKESKATF-DSAGVKLIAVGVGTPNKACIL---AERLPFPMDC-LY 158

Query 66  TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y L G+ T PA
Sbjct 159 ADPDRKAYDVLGLYYGLSRTLFSFA 183
```

>**ref|XP\_002628008.1** |  conserved hypothetical protein [Ajellomyces dermatitidis SLH14081]

**gb|EEQ72686.1** |  conserved hypothetical protein [Ajellomyces dermatitidis SLH14081]

**gb|EEQ88802.1** | conserved hypothetical protein [Ajellomyces dermatitidis ER-3]  
Length=225

**GENE ID: 8507079 BDBG 00916** | hypothetical protein  
[Ajellomyces dermatitidis SLH14081]

Score = 41.2 bits (95), Expect = 0.040, Method: Compositional matrix adjust.  
Identities = 25/71 (36%), Positives = 36/71 (51%), Gaps = 6/71 (8%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQ-ERKWRLMVICGTPEQAAKFRTDNLGPLP 59
          L+I+ RHF C C+ VK ++ A 0 +++GCG A +R D T P
Sbjct 51  LIIFIRHFFCGSCQDYVKAIVESIPAPSQLPNTNTTIIIVGCGASSLIATYR-DTTKCPF- 108

Query 60  ASIPVYTDPLR 70
          P+YTDP R
Sbjct 109 ---PIYTDPTR 116
```

>**gb|ACN25853.1** |  unknown [Zea mays]  
Length=261

**GENE ID: 100384338 pco101707a** | LOC100384338 [Zea mays]

Score = 41.2 bits (95), Expect = 0.042, Method: Compositional matrix adjust.  
Identities = 31/85 (37%), Positives = 40/85 (48%), Gaps = 5/85 (5%)

```
Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVICGTPEQAAKFRTDNLGPLPASIPVY 65
```

```

Sbjct 110 RHFGC C L L ++ F +L+ +G GTP AK R A P P +Y
RHFGCPCCWELASVLRDTERF-DSAGVKLIAVGVGTP---AKARILAERLPPFLEY-LY 164
Query 66 TDPLRRSYLALRFRRGILSTFNLPA 90
DP R++Y L G+ TF PA
Sbjct 165 ADPDRKAYNLLGLYFGVGRFFNPA 189

```

>ref|NP\_001170360.1| **UG** hypothetical protein LOC100384338 [Zea mays]  
**gb|ACN36745.1| G** unknown [Zea mays]  
Length=162

**GENE ID: 100384338 pco101707a** | LOC100384338 [Zea mays]

Score = 41.2 bits (95), Expect = 0.047, Method: Compositional matrix adjust.  
Identities = 31/85 (37%), Positives = 40/85 (48%), Gaps = 5/85 (5%)

```

Query 6 RHFVCVLCRRLVKELAAREQAFRQERKWRMLVIGCGTPEQAAKFRTDGLPLPASIPVY 65
RHFGC C L L ++ F +L+ +G GTP AK R A P P +Y
Sbjct 11 RHFGCPCCWELASVLRDTERF-DSAGVKLIAVGVGTP---AKARILAERLPPFLEY-LY 65
Query 66 TDPLRRSYLALRFRRGILSTFNLPA 90
DP R++Y L G+ TF PA
Sbjct 66 ADPDRKAYNLLGLYFGVGRFFNPA 90

```

>ref|XP\_001784902.1| **UG** predicted protein [Physcomitrella patens subsp. patens]  
**gb|EDQ50291.1| G** predicted protein [Physcomitrella patens subsp. patens]  
Length=187

**GENE ID: 5948108 PHYPADRAFT\_154541** | hypothetical protein  
[Physcomitrella patens subsp. patens] (10 or fewer PubMed links)

Score = 40.8 bits (94), Expect = 0.052, Method: Compositional matrix adjust.  
Identities = 29/90 (33%), Positives = 43/90 (48%), Gaps = 5/90 (5%)

```

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRMLVIGCGTPEQAAKFRTDGLPLPA 60
+V + RHFGC C L + F ++L+ IG G P A+ ++ P PA
Sbjct 31 IVAFLRHFGCPCWEFAAALREAKPKF-DAAGFKLITIGVG-PSSKAQVLSEKL--PFPA 86
Query 61 SIPVYTDPLRRSYLALRFRRGILSTFNLPA 90
+Y DP R++Y AL G+ T+ PA
Sbjct 87 DC-LYADPDRKAYDALGLYHGVRTWLNPA 115

```

>ref|YP\_001543559.1| **G** hypothetical protein Haur\_0783 [Herpetosiphon aurantiacus ATCC 23779]

**gb|ABX03431.1| G** hypothetical protein Haur\_0783 [Herpetosiphon aurantiacus ATCC 23779]  
Length=184

**GENE ID: 5732667 Haur 0783** | hypothetical protein  
[Herpetosiphon aurantiacus ATCC 23779]

Score = 40.8 bits (94), Expect = 0.053, Method: Compositional matrix adjust.  
Identities = 31/91 (35%), Positives = 42/91 (47%), Gaps = 6/91 (6%)

```

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRMLVIGCGTPEQAAKFRTDGLPLPA 60
L++ RH GC LCR+ + L +Q F + VI G + A FR L LP
Sbjct 32 LLMLMRHVGCGLCRQQLYRLREYQRFHHAH-CEIAVIIMGDAKMAHGFR---QLNRLP- 86
Query 61 SIPVYTDPLRRSYLALRFRRGILSTFNLPA 91
PVY+DP + Y A +G L T P +
Sbjct 87 -FPVYSDPKQHVEAFEIGQGSGLWTVAGPHV 116

```

>emb|CBI57389.1| unnamed protein product [Sordaria macrospora]  
Length=411

Score = 40.8 bits (94), Expect = 0.058, Method: Composition-based stats.  
Identities = 28/83 (34%), Positives = 38/83 (46%), Gaps = 8/83 (9%)

```

Query 1 LVIWPRHFGCVLCRRLVKELAAR---EQAFRQERKWRMLVIGCGTPEQAAKFRTDGLP 57
L+I+ RHF C C+ ++ L+A + R + VIGCG P Q AT P
Sbjct 124 LIIFVRHFFCGCQEYLRTLSASITPDALLRLPLSTFIAVIGCGDP-QLIDMYAQATNCP 182
Query 58 LPASIPVYTDPLRRSYLALRFRR 80
P+Y DP R+ Y L R



```

Sbjct 183 F----PIYADPTRKLYQELGMVR 201

>emb|CAN81555.1| hypothetical protein VITISV\_040397 [Vitis vinifera]  
Length=201

Score = 40.8 bits (94), Expect = 0.063, Method: Compositional matrix adjust.  
Identities = 28/85 (33%), Positives = 37/85 (44%), Gaps = 5/85 (5%)


Query 6 RHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRDATLGPLPASIPVY 65  
RHFGC C L L + F +L+ +G GTP +A A P P +Y  
Sbjct 50 RHFGFCCEWELASALKESKARF-DSAGVKLIAVGVGTPNKACIL---AERLPFPMDC-LY 104  
Query 66 TDPLRRSYLALRFRRGILSTFNLPA 90  
DP R++Y L G+ T PA  
Sbjct 105 ADPDRKAYDVLGLYYGLSRTLFSFA 129

>ref|XP\_961259.1|  hypothetical protein NCU04268 [Neurospora crassa OR74A]  
emb|CAC28745.1| putative protein [Neurospora crassa]  
gb|EAA32023.1|  conserved hypothetical protein [Neurospora crassa OR74A]  
Length=411

GENE ID: 3877391 NCU04268 | hypothetical protein [Neurospora crassa OR74A]  
(10 or fewer PubMed links)

Score = 40.8 bits (94), Expect = 0.063, Method: Composition-based stats.  
Identities = 27/83 (33%), Positives = 38/83 (46%), Gaps = 8/83 (9%)

Query 1 LVIWPRHFGCVLCRRLVKELAAR---EQAFRQERKWRLMVIGCGTPEQAAKFRDATLGP 57  
L+I+ RHF C C+ ++ L+A + R + V+GCG P Q AT P  
Sbjct 124 LIIFVRHFFCGCQEQEYLRTLSASITPDALLRLPLSTFIAVVGCGNP-QLIDMYAQATNCP 182  
Query 58 LPASIPVYTDPLRRSYLALRFRR 80  
P+Y DP R+ Y L R  
Sbjct 183 F----PIYADPTRKLYQELGMVR 201

>gb|AAI36502.1|  C9orf21 protein [Homo sapiens]  
Length=172

GENE ID: 195827 C9orf21 | chromosome 9 open reading frame 21 [Homo sapiens]  
(10 or fewer PubMed links)


Score = 40.8 bits (94), Expect = 0.064, Method: Compositional matrix adjust.  
Identities = 24/76 (32%), Positives = 37/76 (49%), Gaps = 5/76 (6%)

Query 6 RHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRDATLGPLPASIPVY 65  
+HF C +C+ V++LA ++F OE L+VIG 0++ + S +Y  
Sbjct 10 QHFLCYICKEYVEDLAKIPRSFLQEVNVTLIVIG-----QSSYHHIEPFCKLTGYSHEIY 64  
Query 66 TDPLRRSYLALRFRRG 81  
DP R Y L +RG  
Sbjct 65 VDPEREIYKRLGMKRG 80

>gb|AD028366.1| upf0308 protein c9orf21-like protein [Ictalurus furcatus]  
Length=223

Score = 40.4 bits (93), Expect = 0.068, Method: Compositional matrix adjust.  
Identities = 24/81 (30%), Positives = 37/81 (46%), Gaps = 5/81 (6%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRDATLGPLPA 60  
++I+ RHF C C+ V++L+ 0 + RL+VIG 0+ +A  
Sbjct 55 IIIFVRHFLCFTCQEQYVEDLSQIPQEILLDADVRLIVIG-----QSGFSHIEAFCSLTGY 109  
Query 61 SIPVYTDPLRRSYLALRFRRG 81  
+Y DP R Y L +RG  
Sbjct 110 QHEIYVDPERHIYEKLGKMRG 130

>ref|XP\_002263959.1|  PREDICTED: hypothetical protein [Vitis vinifera]  
Length=255

GENE ID: 100249728 LOC100249728 | hypothetical protein LOC100249728  
[Vitis vinifera] (10 or fewer PubMed links)



Score = 40.4 bits (93), Expect = 0.078, Method: Compositional matrix adjust.  
Identities = 28/85 (33%), Positives = 37/85 (44%), Gaps = 5/85 (5%)

```

Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRDATLGPLPASIPVY 65
          RHFGC C L L + F +L+ +G GTP +A A P P +Y
Sbjct 104 RHFGCFCCWELASALKESKARF-DSAGVKLIAVGVGTPNKACIL---AERLPPFMDCLY 158

Query 66 TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R++Y L G+ T PA
Sbjct 159 ADPDRKAYDVLGLYYGLSRTLFSPLA 183

```

>ref|XP\_002533575.1|  conserved hypothetical protein [Ricinus communis]  
 gb|EEF28810.1|  conserved hypothetical protein [Ricinus communis]  
 Length=255

**GENE ID: 8276075 RCOM\_0366250** | hypothetical protein [Ricinus communis]


Score = 40.4 bits (93), Expect = 0.080, Method: Compositional matrix adjust.  
 Identities = 30/85 (36%), Positives = 36/85 (43%), Gaps = 5/85 (5%)

```

Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRDATLGPLPASIPVY 65
          RHFGC C L L + F +L+ IG G P +A R A P P +Y
Sbjct 104 RHFGCPCCWELASVLKEAKSKF-DSAGVKLIAIGVGVGAPNKA---RMLADRLPPFMDCLY 158

Query 66 TDPLRRSYLALRFRRGILSTFNLPA 90
          DP R +Y L G TF PA
Sbjct 159 ADPNREAYNVLGLYYGFGRTFFNPA 183

```

>emb|CAI40533.1|  novel protein [Homo sapiens]  
 Length=179

**GENE ID: 195827 C9orf21** | chromosome 9 open reading frame 21 [Homo sapiens]  
 (10 or fewer PubMed links)


Score = 40.4 bits (93), Expect = 0.083, Method: Compositional matrix adjust.  
 Identities = 24/76 (32%), Positives = 37/76 (49%), Gaps = 5/76 (6%)

```

Query 6  RHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRDATLGPLPASIPVY 65
          +HF C +C+ V++LA ++F QE L+VIG Q++ + S +Y
Sbjct 17 QHFLCYICKEYVEDLAKIPRSFLQEANVTLIVIG-----QSSYHHIEPFCKLTGYSHEIY 71

Query 66 TDPLRRSYLALRFRRG 81
          DP R Y L +RG
Sbjct 72 VDPEREIYKRLGMKRG 87

```

>dbj|BAE40544.1|  unnamed protein product [Mus musculus]  
 Length=194

**GENE ID: 66129 1110018J18Rik** | RIKEN cDNA 1110018J18 gene [Mus musculus]  
 (Over 10 PubMed links)

Score = 39.7 bits (91), Expect = 0.12, Method: Compositional matrix adjust.  
 Identities = 24/81 (30%), Positives = 40/81 (50%), Gaps = 5/81 (6%)

```

Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRDATLGPLPA 60
          +V++ RHF C +C+ V++LA ++ +E L+VIG Q++ +
Sbjct 27 VVVFVRHFLCYVCKEYVEDLAKIPKSVLREADVTLIVIG-----QSSYHHIEPFCKLTGY 81

Query 61 SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 82 SHEIYVDPEREIYKRLGMKRG 102

```

>ref|ZP\_01089021.1| hypothetical protein DSM3645\_00440 [Blastopirellula marina DSM  
 3645]  
 gb|EAQ82136.1| hypothetical protein DSM3645\_00440 [Blastopirellula marina DSM  
 3645]  
 Length=184

Score = 39.7 bits (91), Expect = 0.12, Method: Compositional matrix adjust.  
 Identities = 27/91 (30%), Positives = 39/91 (43%), Gaps = 6/91 (6%)


```

Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRDATLGPLPA 60
          L+ + RH GC CR K+L Q+ ++ I GT E+A + + L
Sbjct 41 LIQFARHLGCTFCRDRAKQLKLDYPEI-QQHNGDVVFITMGThERAQQLQDGMQL----- 94

Query 61 SIPVYTDPLRRSYLALRFRRGILSTFNLPAL 91
          V DP R +Y A R RG +S P +

```

Sbjct 95 PFDVLVDPHRAAYQAFRVPRGNVSQVAGPQV 125


>dbj|BAB24662.1|  unnamed protein product [Mus musculus]  
Length=186

**GFNF TD: 66179 1110018J18Rik** | RIKEN cDNA 1110018J18 gene [Mus musculus]  
(Over 10 PubMed links)

Score = 39.7 bits (91), Expect = 0.12, Method: Compositional matrix adjust.  
Identities = 24/81 (30%), Positives = 40/81 (50%), Gaps = 5/81 (6%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRMLVIGCGTPEQAAKFRDTATLGPLPA 60  
+V++ RHF C +C+ V++LA ++ +E L+VIG Q++ +  
Sbjct 19 VVVFVRHFLCYCKEYVEDLAKIPKSVLREADVTLIVIG-----QSSYHHIEPFCKLTGY 73

Query 61 SIPVYTDPLRRSYLALRFRRG 81  
S +Y DP R Y L +RG  
Sbjct 74 SHEIYVDPEREIYKRLGMKRG 94



>ref|XP\_001824981.1|  hypothetical protein AOR\_1\_1228084 [Aspergillus oryzae RIB40]  
dbj|BAE63848.1| unnamed protein product [Aspergillus oryzae]  
Length=250

**GENE ID: 5997067 AOR\_1\_1228084** | hypothetical protein  
[Aspergillus oryzae RIB40] (10 or fewer PubMed links)

Score = 39.7 bits (91), Expect = 0.13, Method: Compositional matrix adjust.  
Identities = 24/81 (30%), Positives = 36/81 (45%), Gaps = 8/81 (9%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAFRQE---RKWRLMVICGCGTPEQAAKFRDTATLGP 57  
LVI+ RHF C C+ ++ L+A E +++IGCG P + +  
Sbjct 75 LVIFVRHFFCGSCQEFRLTSLASITPKALEPLATSTSVVIIGCGDPGLIEMYEKETN--- 131

Query 58 LPASIPVYTDPLRRSYLALRF 78  
P+YTDP R+ Y L  
Sbjct 132 --CQFPIYDTPTRQLYQDLDM 150


>ref|XP\_002383597.1|  conserved hypothetical protein [Aspergillus flavus NRRL3357]  
gb|EED46061.1|  conserved hypothetical protein [Aspergillus flavus NRRL3357]  
Length=250


**GENE ID: 7910923 AFLA\_097240** | hypothetical protein  
[Aspergillus flavus NRRL3357]

Score = 39.7 bits (91), Expect = 0.13, Method: Compositional matrix adjust.  
Identities = 24/81 (30%), Positives = 36/81 (45%), Gaps = 8/81 (9%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAFRQE---RKWRLMVICGCGTPEQAAKFRDTATLGP 57  
LVI+ RHF C C+ ++ L+A E +++IGCG P + +  
Sbjct 75 LVIFVRHFFCGSCQEFRLTSLASITPKALEPLATSTSVVIIGCGDPGLIEMYEKETN--- 131

Query 58 LPASIPVYTDPLRRSYLALRF 78  
P+YTDP R+ Y L  
Sbjct 132 --CQFPIYDTPTRQLYQDLDM 150

>ref|YP\_001543558.1|  hypothetical protein Haur\_0782 [Herpetosiphon aurantiacus ATCC 23779]

gb|ABX03430.1|  hypothetical protein Haur\_0782 [Herpetosiphon aurantiacus ATCC 23779]  
Length=183

**GENE ID: 5732666 Haur\_0782** | hypothetical protein  
[Herpetosiphon aurantiacus ATCC 23779]

Score = 39.7 bits (91), Expect = 0.14, Method: Compositional matrix adjust.  
Identities = 21/81 (26%), Positives = 40/81 (50%), Gaps = 6/81 (7%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRMLVIGCGTPEQAAKFRDTATLGPLPA 60  
L+ + R+ GC +CR+ + L +QAF+ + W+++V+ G E ++FR+ +  
Sbjct 37 LIFFMRNIGCGICRQTLNLRDYOQAFK-DAGWQIVVLMGMGNVELVSRFRSMYNI----- 90

Query 61 SIPVYTDPLRRSYLALRFRRG 81  
P+Y D + Y G  
Sbjct 91 PFPIYVDQSLQVYDYFEIVEG 111

>dbj|BAE26709.1| **G** unnamed protein product [Mus musculus]  
Length=182

**GFNF TD: 66129 1110018J18Rik** | RIKEN cDNA 1110018J18 gene [Mus musculus]  
(Over 10 PubMed links)

Score = 39.7 bits (91), Expect = 0.14, Method: Compositional matrix adjust.  
Identities = 24/81 (30%), Positives = 40/81 (50%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGGLPA 60
          +V++ RHF C +C+ V++LA ++ +E L+VIG 0++ +
Sbjct 59  VVVFVRHFLCYVCKEYVEDLAKIPKSVLREADVTLVIG-----QSSYHHIEPFCKLTGY 113

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 114 SHEIYVDPEREIYKRLGMKRG 134
```

>ref|XP\_001940231.1| **G** conserved hypothetical protein [Pyrenophora tritici-repentis  
Pt-1C-BFP]

**gb|EDU42950.1** **G** conserved hypothetical protein [Pyrenophora tritici-repentis  
Pt-1C-BFP]  
Length=342

**GENE ID: 6348196 PTRG\_09899** | hypothetical protein  
[Pyrenophora tritici-repentis Pt-1C-BFP]

Score = 39.7 bits (91), Expect = 0.14, Method: Composition-based stats.  
Identities = 29/91 (32%), Positives = 44/91 (49%), Gaps = 12/91 (13%)

```
Query 1  LVIWPRHFGVLCRRLVKELAAR--EQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGGLP 57
          L+I+ RHF C C+ ++ LA+ E + VIGCG P+ + +AT P
Sbjct 120 LIIFIRHFFCGNCQEYIRTLASSVKPEDLLALPTPTSITVIGCGRPDLIPMY-IEATGCP 178

Query 58  LPASIPVYTDPLRRSYLALRFRRGILSTFNL 88
          P+Y +P R+ Y L G+ T+NL
Sbjct 179 F----PIYAEPTRKLYDHL----GMTRTYNL 201
```

>gb|EFQ95671.1| hypothetical protein PTT\_05482 [Pyrenophora teres f. teres 0-1]  
Length=342

Score = 39.7 bits (91), Expect = 0.14, Method: Composition-based stats.  
Identities = 29/91 (32%), Positives = 44/91 (49%), Gaps = 12/91 (13%)

```
Query 1  LVIWPRHFGVLCRRLVKELAAR--EQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGGLP 57
          L+I+ RHF C C+ ++ LA+ E + VIGCG P+ + +AT P
Sbjct 120 LIIFIRHFFCGNCQEYIRTLASSVKPEDLLALPTPTSITVIGCGRPDLIPMY-IEATSCP 178

Query 58  LPASIPVYTDPLRRSYLALRFRRGILSTFNL 88
          P+Y +P R+ Y L G+ T+NL
Sbjct 179 F----PIYAEPTRKLYDHL----GMTRTYNL 201
```

>gb|EDL16239.1| **G** RIKEN cDNA 1110018J18, isoform CRA\_c [Mus musculus]  
Length=141

**GFNF TD: 66129 1110018J18Rik** | RIKEN cDNA 1110018J18 gene [Mus musculus]  
(Over 10 PubMed links)

Score = 39.3 bits (90), Expect = 0.16, Method: Compositional matrix adjust.  
Identities = 24/81 (30%), Positives = 40/81 (50%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGGLPA 60
          +V++ RHF C +C+ V++LA ++ +E L+VIG 0++ +
Sbjct 59  VVVFVRHFLCYVCKEYVEDLAKIPKSVLREADVTLVIG-----QSSYHHIEPFCKLTGY 113

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 114 SHEIYVDPEREIYKRLGMKRG 134
```

>ref|NP\_001158627.1| **UG** UPF0308 protein C9orf21 homolog [Oncorhynchus mykiss]

**gb|AC008544.1** **G** UPF0308 protein C9orf21 homolog [Oncorhynchus mykiss]  
Length=224



**GENE ID: 100305250 ci021** | UPF0308 protein C9orf21 homolog  
[Oncorhynchus mykiss]

Score = 39.3 bits (90), Expect = 0.17, Method: Compositional matrix adjust.  
Identities = 28/85 (33%), Positives = 39/85 (46%), Gaps = 5/85 (5%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDATLGPLPA 60
          +VI+ R+F C C+ V +L+ +E RL+VIG + F +L P
Sbjct 55  VVIFVRNFLCHTCKEYVDDL SRIPAEILKEAGLRLVVIGQSSHHHIESF---CSLTGYPH 111

Query 61  SIPVYTDPLRRSYLALRFRRGILST 85
          I Y DP R Y L RRG +S
Sbjct 112  DI--YVDPERCIYKRLGMRRGEMSV 134
```

>ref|NP\_079646.1| **UGM** hypothetical protein LOC66129 [Mus musculus]  
sp|Q9D1A0.1|**CI021 MOUSE G** RecName: Full=UPF0308 protein C9orf21 homolog  
dbj|BAB22993.1| **G** unnamed protein product [Mus musculus]  
dbj|BAE28518.1| **G** unnamed protein product [Mus musculus]  
gb|AAI40306.1| **G** RIKEN cDNA 1110018J18 gene [synthetic construct]  
gb|EDL16238.1| **G** RIKEN cDNA 1110018J18, isoform CRA\_b [Mus musculus]  
gb|AAI56632.1| **G** RIKEN cDNA 1110018J18 gene [synthetic construct]  
Length=226

**GENE ID: 66129 1110018J18Rik** | RIKEN cDNA 1110018J18 gene [Mus musculus]  
(Over 10 PubMed links)

Score = 39.3 bits (90), Expect = 0.17, Method: Compositional matrix adjust.  
Identities = 24/81 (30%), Positives = 40/81 (50%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDATLGPLPA 60
          +V++ RHF C +C+ V++LA ++ +E L+VIG Q++ +
Sbjct 59  VVVFVRHFLCYVCKEYVEDLAKIPKSVLREADVTLIVIG-----QSSYHHIEPFCKLTGY 113

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          S +Y DP R Y L +RG
Sbjct 114  SHEIYVDPEREIYKRLGMKRG 134
```

>gb|ACI68733.1| Clorf93 homolog [Salmo salar]  
Length=224

Score = 38.9 bits (89), Expect = 0.23, Method: Compositional matrix adjust.  
Identities = 28/85 (33%), Positives = 39/85 (46%), Gaps = 5/85 (5%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDATLGPLPA 60
          +VI+ R+F C C+ V +L+ +E RL+VIG + F +L P
Sbjct 55  VVIFVRNFLCHTCKEYVDDL SRIPAEVLKEAGLRLVVIGQSSHHHIESF---CSLTGYPH 111

Query 61  SIPVYTDPLRRSYLALRFRRGILST 85
          I Y DP R Y L RRG +S
Sbjct 112  DI--YVDPERCIYKRLGMRRGEMSV 134
```

>ref|ZP\_01909602.1| hypothetical protein PPSIR1\_24939 [Plesiocystis pacifica SIR-1]  
gb|EDM77490.1| hypothetical protein PPSIR1\_24939 [Plesiocystis pacifica SIR-1]  
Length=237

Score = 38.5 bits (88), Expect = 0.27, Method: Compositional matrix adjust.  
Identities = 28/87 (33%), Positives = 39/87 (45%), Gaps = 3/87 (3%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDATLGPLPA 60
          L+++ RH+ CV C R V+EL+ R R + IG G EQ A F L
Sbjct 76  LLLFLRHACVGC SRQVEELSPRVLEL-SALGVRVFIGSGDAEQMAAFAQRHAL--TDK 132

Query 61  SIPVYTDPLRRSYLALRFRRGILSTFN 87
          ++ + DP SY A RR + S
Sbjct 133  AVVMVCDPALGSYAAADMRRSLWSAVE 159
```


>ref|XP\_001829670.2| **G** hypothetical protein CC1G\_11406 [Coprinopsis cinerea okayama7#130]  
gb|EAU92121.2| **G** hypothetical protein CC1G\_11406 [Coprinopsis cinerea okayama7#130]  
Length=222

**GENE ID: 6006103 CC1G\_11406** | hypothetical protein  
[Coprinopsis cinerea okayama7#130] (10 or fewer PubMed links)

Score = 38.5 bits (88), Expect = 0.29, Method: Compositional matrix adjust.  
Identities = 21/76 (28%), Positives = 37/76 (49%), Gaps = 5/76 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDATALGPLA 60
          +VI+ RHF C C+ V+ +A + ++ +++V+GCG + +
Sbjct 55  VVIFIRHFFCGSCQMYVEHVAKVPASNLEQAGVQIVVVGCGDWKAIPMYHETTGF----- 109

Query 61  SIPVYTDPLRRSYLAL 76
          + P+Y DP R Y AL
Sbjct 110 TGPIYADPNRDLYFAL 125
```


>ref|NP\_001177306.1|  selenoprotein L [Ciona intestinalis]  
Length=303


**GENE ID: 100175934** [cell](#) | selenoprotein L [Ciona intestinalis]  
(10 or fewer PubMed links)

Score = 38.5 bits (88), Expect = 0.32, Method: Composition-based stats.  
Identities = 29/91 (32%), Positives = 43/91 (48%), Gaps = 6/91 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDATALGPLA 60
          LVI RHF + R+ V E+ ++ F Q K ++++ G E+A D L
Sbjct 150 LVILLRHFALPURKHVTEIQEKQDEFSQ-LKCSIILVSFG--EEAG---ADLWLKETKF 203

Query 61  SIPVYTDPLRRSYLALRFRRGILSTFNLPAL 91
          S P+Y D R+ Y AL R + +N AL
Sbjct 204 SFPMYLDKQRKIYSALGLGRSVAKVYVNCNAL 234
```

>ref|XP\_003050915.1|  hypothetical protein NECHADRAFT\_69625 [Nectria haematococca mpVI  
77-13-4]


**gb|EEU45202.1|**  hypothetical protein NECHADRAFT\_69625 [Nectria haematococca mpVI  
77-13-4]  
Length=386


**GENE ID: 9668148 NECHADRAFT\_69625** | hypothetical protein  
[Nectria haematococca mpVI 77-13-4]

Score = 38.1 bits (87), Expect = 0.36, Method: Composition-based stats.  
Identities = 26/79 (33%), Positives = 35/79 (45%), Gaps = 8/79 (10%)

```
Query 1  LVIWPRHFGCVLCRRLVKELA---AREQAFRQERKWRLMVIGCGTPEQAAKFRTDATALGP 57
          L+I+ RHF C C+ ++ L+ E R + VIGCG P D +
Sbjct 148 LIIFVRHFFCGNCQEFRLSLSDAVTPEALLRPLVSTFIAVIGCGDPAL-----IDMYVNE 202

Query 58  LPASIPVYTDPLRRSYLAL 76
          PVYTD P R + AL
Sbjct 203 TNCRFPVYTDPTRSLFHAL 221
```

>ref|XP\_001800743.1|  hypothetical protein SNOG\_10473 [Phaeosphaeria nodorum SN15]


**gb|EAT81867.1|**  hypothetical protein SNOG\_10473 [Phaeosphaeria nodorum SN15]  
Length=338


**GENE ID: 5977650 SNOG\_10473** | hypothetical protein [Phaeosphaeria nodorum SN15]

Score = 37.7 bits (86), Expect = 0.49, Method: Composition-based stats.  
Identities = 29/93 (32%), Positives = 42/93 (46%), Gaps = 12/93 (12%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAR---EQAFRQERKWRLMVIGCGTPEQAAKFRTDATALGP 57
          L+I+ RHF C C+ ++ L++ E R + VIGCG P + + T P
Sbjct 122 LIIFIRHFFCGNCQEYLRSLSSINAEALLALPTPTSITVIGCGKPALIPMY-AETTKCP 180

Query 58  LPASIPVYTDPLRRSYLALRFRRGILSTFNLPA 90
          P+Y DP R+ Y L G+ TF L A
Sbjct 181 F----PIYADPTRKLYDLL----GMTRTFQLGA 205
```

>ref|XP\_002114328.1|  hypothetical protein TRIADDRAFT\_58024 [Trichoplax adhaerens]

**gb|EDV23418.1|**  hypothetical protein TRIADDRAFT\_58024 [Trichoplax adhaerens]  
Length=218

**GENE ID: 6755541 TRTADDRAFT\_58024** | hypothetical protein [Trichoplax adhaerens]  
(10 or fewer PubMed links)

Score = 37.7 bits (86), Expect = 0.52, Method: Compositional matrix adjust.  
Identities = 22/60 (37%), Positives = 28/60 (47%), Gaps = 5/60 (8%)

Query 17 VKELAAREQAFRQERKWRMLVIGCGTPEQAAKFRDTATLGPLPASIPVYTDPLRRSYLAL 76  
V + + R AF +E RL+VIGCG P KFR + +Y DP R Y L  
Sbjct 93 VYKYSKRSYAFLEKANVRLVIGCGKPNFIKFRQETGYNH-----DMYCDPQRTIYSKL 147

>ref|XP\_003007519.1| **G** conserved hypothetical protein [Verticillium albo-atrum VaMs.102]  
>gb|EEY15598.1| **G** conserved hypothetical protein [Verticillium albo-atrum VaMs.102]  
Length=412

**GENE ID: 9533848 VDBG\_01707** | hypothetical protein  
[Verticillium albo-atrum VaMs.102]

Score = 37.4 bits (85), Expect = 0.58, Method: Composition-based stats.  
Identities = 28/83 (34%), Positives = 34/83 (41%), Gaps = 8/83 (9%)

Query 1 LVIWPRHFGCVLCR---RLVKELAAREQAFRQERKWRMLVIGCGTPEQAAKFRDTATLGP 57  
L I+ RHF C C+ R V E E R + VIGCG P + + T P  
Sbjct 171 LFIFVRHFFCGNCQEFRAVSESITTESLLRLPVSTFITVIGCGDPGLIQMY-AETTNC 229  
Query 58 LPASIPVYTDPLRRSYLALRFRR 80  
P+YTD P R Y R  
Sbjct 230 F----PIYDPTRSLYKFFAMTR 248

>gb|ACI70082.1| Clorf93 homolog [Salmo salar]  
Length=232

Score = 37.0 bits (84), Expect = 0.77, Method: Compositional matrix adjust.  
Identities = 26/85 (31%), Positives = 39/85 (46%), Gaps = 5/85 (5%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRMLVIGCGTPEQAAKFRDTATLGPLPA 60  
++I+ R+F C C+ V +L+ +E RL+VIG + F +L P  
Sbjct 63 VIIFVRNFLCHTCKEYVDDLRSIPAEVLKEAGRLVWVIGQSSHHHIESF---CSLTGYPH 119  
Query 61 SIPVYTDPLRRSYLALRFRRGILST 85  
+ Y DP R Y L RRG +S  
Sbjct 120 DM--YVDPERCIYKRLGMRRGEMSV 142

>ref|YP\_001547656.1| **G** hypothetical protein Haur\_4898 [Herpetosiphon aurantiacus ATCC 23779]

>gb|ABX07528.1| **G** hypothetical protein Haur\_4898 [Herpetosiphon aurantiacus ATCC 23779]  
Length=193

**GENE ID: 5736734 Haur\_4898** | hypothetical protein  
[Herpetosiphon aurantiacus ATCC 23779]

Score = 37.0 bits (84), Expect = 0.81, Method: Compositional matrix adjust.  
Identities = 33/91 (37%), Positives = 46/91 (51%), Gaps = 6/91 (6%)

Query 1 LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRMLVIGCGTPEQAAKFRDTATLGPLPA 60  
L+++ RH GC +CRR + +L AFR L VI T + A + A L LP  
Sbjct 30 LLVFLRHGCAVCRNLLDLYEYTTAFRM-LDINLAVI---TMAEPAAAQAFARLYRLP- 84  
Query 61 SIPVYTDPLRRSYLALRFRRGILSTFNLPAL 91  
IP+Y+DP R+ Y A F G L + P +  
Sbjct 85 -IPIYSDPERQIYRAAGFSEGLYSVASPQV 114

>gb|EFQ35083.1| hypothetical protein GLRG\_10227 [Glomerella graminicola M1.001]  
Length=428

Score = 36.6 bits (83), Expect = 1.2, Method: Composition-based stats.  
Identities = 26/83 (32%), Positives = 35/83 (43%), Gaps = 8/83 (9%)

Query 1 LVIWPRHFGCVLCRRLVKELAAR--EQAFRQERKWRMLVIGCGTPEQAAKFRDTATLGP 57  
L+I+ RHF C C+ ++ L+ + + VIGCG P D L  
Sbjct 184 LIIFVRHFFCGNCQDFLRTLSESITPDALLSLPMSTFIAVIGCGNPGL-----IDMYLQE 238  
Query 58 LPASIPVYTDPLRRSYLALRFRR 80  
PVYTD P RR + L R  
Sbjct 239 TGCPFPVYDPTRRFLDFTLGMTR 261

>ref|XP\_001368977.1| **UGM** PREDICTED: similar to C90RF21 [Monodelphis domestica]  
Length=354

**GENE ID: 100014720 LOC100014720** | similar to C90RF21 [Monodelphis domestica]

Score = 36.2 bits (82), Expect = 1.3, Method: Composition-based stats.  
Identities = 23/81 (29%), Positives = 38/81 (47%), Gaps = 5/81 (6%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAREQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGGLPA 60
          +V++ RHF C C+ V++LA ++F Q+ L+VIG + + F
Sbjct 187 IVVFVRHFLCYTCKEYVEDLAKIPKSFLQDANVTLIVIGQSSFQHIEPFCKLTRY----- 241

Query 61  SIPVYTDPLRRSYLALRFRRG 81
          S +Y D R+ Y L +G
Sbjct 242 SHEIYVDTERKIYRKLGMNKG 262
```

>ref|XP\_002544130.1| **G** predicted protein [Uncinocarpus reesii 1704]

**gb|EEP78801.1| G** predicted protein [Uncinocarpus reesii 1704]  
Length=245

**GENE ID: 8443780 UREG\_03647** | hypothetical protein [Uncinocarpus reesii 1704]

Score = 36.2 bits (82), Expect = 1.5, Method: Compositional matrix adjust.  
Identities = 23/81 (29%), Positives = 34/81 (42%), Gaps = 6/81 (7%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAR-EQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGGLP 59
          +V++ RHF C C+ V+ L++ L+VIGCG + +
Sbjct 80  MVVFVRHFFCGSCQDYVQTLSSIPSPASLPTGTSLVVIGCGATSLIPMYAKTTS----- 134

Query 60  ASIPVYTDPLRRSYLALRFRR 80
          S P+YTD P R Y R
Sbjct 135 CSFPIYTDPTSRLYTIFGMTR 155
```

>ref|XP\_003070048.1| **G** hypothetical protein CPC735\_032390 [Coccidioides posadasii C735 delta 50Wgp]

**gb|EER27903.1| G** hypothetical protein CPC735\_032390 [Coccidioides posadasii C735 delta 50Wgp]  
Length=206

**GENE ID: 9695543 CPC735\_032390** | hypothetical protein  
[Coccidioides posadasii C735 delta 50Wgp]

Score = 36.2 bits (82), Expect = 1.6, Method: Compositional matrix adjust.  
Identities = 22/74 (30%), Positives = 33/74 (45%), Gaps = 6/74 (8%)

```
Query 1  LVIWPRHFGCVLCRRLVKELAAR-EQAFRQERKWRLMVIGCGTPEQAAKFRTDNLGGLP 59
          +V++ RHF C C+ V L++ L+VIGCG + + +
Sbjct 41  MVVFVRHFFCGSCQDYVATLSSIPSPASLPPGVNLVVIGCGASSLISMYANTTS----- 95

Query 60  ASIPVYTDPLRRSY 73
          S P+YTD P R Y
Sbjct 96  CSFPIYTDPTGRLY 109
```