

BLAST


Basic Local Alignment Search Tool

NCBI/ [BLAST](#)/ [blastp suite](#)/ [Formatting Results - RWF78ZUZ016](#)

C. owczarzaki | query NCBI SelL (142 letters)

Query ID	lcl 38891	Database Name	nr
Description	C. owczarzaki query NCBI SelL	Description	All non-redundant GenBank CDS translations+PDB+SwissProt+PII excluding environmental samples from WGS projects
Molecule type	amino acid	Program	BLASTP 2.2.25+
Query Length	142		

Descriptions

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

<u>Accession</u>	<u>Description</u>	<u>Max score</u>	<u>Total score</u>	<u>Query coverage</u>	<u>E value</u>	<u>Links</u>
EFW45143.1	predicted protein [Capsaspora owczarzaki ATCC 30864]	303	303	100%	4e-81	
EFA74551.1	hypothetical protein PPL_00049 [Polysphondylium pallidum PN500]	73.9	73.9	96%	6e-12	
ABK23985.1	unknown [Picea sitchensis]	73.2	73.2	94%	1e-11	
XP_001752693.1	predicted protein [Physcomitrella patens subsp. patens] >gb EDQ82564.1 predicted protein [Physcomitrella patens subsp. patens]	71.6	71.6	90%	3e-11	UG
ACO13446.1	C1orf93 [Esox lucius]	71.6	71.6	99%	3e-11	
NP_030274.1	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]	71.2	71.2	91%	4e-11	UG
XP_002881499.1	hypothetical protein ARALYDRAFT_482716 [Arabidopsis lyrata subsp. lyrata] >gb EFH57758.1 hypothetical protein ARALYDRAFT_482716 [Arabidopsis lyrata subsp. lyrata]	70.5	70.5	91%	7e-11	G

Accession	Description	Max score	Total score	Query coverage	E value	Links
ACI68733.1	C1orf93 homolog [Salmo salar]	69.7	69.7	99%	1e-10	
NP_001158627.1	UPF0308 protein C9orf21 homolog [Oncorhynchus mykiss] >gb ACO08544.1 UPF0308 protein C9orf21 homolog [Oncorhynchus mykiss]	69.7	69.7	99%	1e-10	UG
ACI70082.1	C1orf93 homolog [Salmo salar]	69.3	69.3	99%	1e-10	
XP_002574167.1	PRX_like2 domain-containing protein [Schistosoma mansoni] >emb CAZ30400.1 PRX_like2 domain-containing protein [Schistosoma mansoni]	68.9	68.9	100%	2e-10	UG
ACO51726.1	C1orf93 homolog [Rana catesbeiana]	68.6	68.6	95%	3e-10	
XP_002587381.1	hypothetical protein BRAFLDRAFT_96268 [Branchiostoma floridae] >gb EEN43392.1 hypothetical protein BRAFLDRAFT_96268 [Branchiostoma floridae]	68.6	68.6	89%	3e-10	UG
AAW27591.1	SJCHGC05103 protein [Schistosoma japonicum] >emb CAX73680.1 hypothetical protein [Schistosoma japonicum]	67.8	67.8	100%	4e-10	
XP_001762195.1	predicted protein [Physcomitrella patens subsp. patens] >gb EDQ72987.1 predicted protein [Physcomitrella patens subsp. patens]	67.4	67.4	99%	6e-10	UG
ACI67539.1	C1orf93 homolog [Salmo salar]	67.4	67.4	95%	6e-10	
ADO27754.1	uncharacterized protein c1orf93-like protein [Ictalurus furcatus]	67.4	67.4	94%	7e-10	
CAX73679.1	hypothetical protein [Schistosoma japonicum]	67.0	67.0	89%	7e-10	
NP_201385.2	unknown protein [Arabidopsis thaliana] >gb AAM20692.1 unknown protein [Arabidopsis thaliana] >gb AAN15652.1 unknown protein [Arabidopsis thaliana]	67.0	67.0	86%	8e-10	UG
XP_002326159.1	predicted protein [Populus trichocarpa] >gb EEE71829.1 predicted protein [Populus trichocarpa]	66.6	66.6	89%	9e-10	UG
NP_001017220.1	hypothetical protein LOC549974 [Xenopus (Silurana) tropicalis] >sp Q28IJ3.1 CA093_XENTR RecName: Full=Uncharacterized protein C1orf93 homolog >emb CAJ83264.1 novel protein [Xenopus (Silurana) tropicalis]	66.6	66.6	94%	9e-10	UGM
CBN80880.1	Uncharacterized [Dicentrarchus labrax]	66.6	66.6	94%	1e-09	
XP_002666723.1	PREDICTED: UPF0308 protein C9orf21 homolog [Danio rerio]	65.9	65.9	97%	2e-09	UGM
XP_002866692.1	hypothetical protein ARALYDRAFT_496819 [Arabidopsis lyrata subsp. lyrata] >gb EFH42951.1 hypothetical protein	65.9	65.9	86%	2e-09	G

Accession	Description	Max score	Total score	Query coverage	E value	Links
	ARALYDRAFT_496819 [Arabidopsis lyrata subsp. lyrata]					
CAG07032.1	unnamed protein product [Tetraodon nigroviridis]	65.1	65.1	94%	3e-09	
XP_639255.1	hypothetical protein DDB_G0283129 [Dictyostelium discoideum AX4] >gb EAL65894.1 hypothetical protein DDB_G0283129 [Dictyostelium discoideum AX4]	65.1	65.1	95%	3e-09	G
XP_001493974.2	PREDICTED: similar to UPF0308 protein C9orf21 [Equus caballus]	65.1	65.1	99%	3e-09	UGM
XP_002664824.1	PREDICTED: hypothetical protein [Danio rerio]	64.7	64.7	94%	4e-09	UGM
NP_998478.1	hypothetical protein LOC406605 [Danio rerio] >sp Q6NV24.1 CA093_DANRE RecName: Full=Uncharacterized protein C1orf93 homolog >gb AAH68342.1 Zgc:85644 [Danio rerio]	64.7	64.7	94%	4e-09	UGM
XP_001784902.1	predicted protein [Physcomitrella patens subsp. patens] >gb EDQ50291.1 predicted protein [Physcomitrella patens subsp. patens]	64.3	64.3	86%	5e-09	UG
XP_001106503.1	PREDICTED: UPF0308 protein C9orf21-like [Macaca mulatta]	63.9	63.9	99%	7e-09	UGM
NP_001087128.1	chromosome 1 open reading frame 93 [Xenopus laevis] >sp Q6AZG8.1 CA093_XENLA RecName: Full=Uncharacterized protein C1orf93 homolog >gb AAH78028.1 MGC82733 protein [Xenopus laevis]	63.5	63.5	94%	7e-09	UG
EGC30304.1	hypothetical protein DICPUDRAFT_99572 [Dictyostelium purpureum]	63.2	63.2	88%	1e-08	
NP_001069904.1	hypothetical protein LOC616897 [Bos taurus] >sp Q148E0.1 C1021_BOVIN RecName: Full=UPF0308 protein C9orf21 homolog >gb AAI18426.1 Chromosome 9 open reading frame 21 ortholog [Bos taurus] >gb DAA26600.1 hypothetical protein LOC616897 [Bos taurus]	62.8	62.8	99%	1e-08	UGM
EAW92661.1	chromosome 9 open reading frame 21, isoform CRA_a [Homo sapiens]	62.8	62.8	99%	2e-08	G
XP_002928763.1	PREDICTED: UPF0308 protein C9orf21-like, partial [Ailuropoda melanoleuca]	62.8	62.8	99%	2e-08	GM
EFB20490.1	hypothetical protein PANDA_018799 [Ailuropoda melanoleuca]	62.8	62.8	99%	2e-08	
EDL84435.1	similar to UPF0308 protein C9orf21, isoform CRA_c [Rattus norvegicus]	62.4	62.4	99%	2e-08	G

Accession	Description	Max score	Total score	Query coverage	E value	Links
XP_003130904.1	PREDICTED: UPF0308 protein C9orf21 homolog isoform 1 [Sus scrofa]	62.4	62.4	99%	2e-08	UGM
BAB24662.1	unnamed protein product [Mus musculus]	62.4	62.4	99%	2e-08	G
XP_002820049.1	PREDICTED: UPF0308 protein C9orf21-like [Pongo abelii]	62.4	62.4	99%	2e-08	GM
BAE40544.1	unnamed protein product [Mus musculus]	62.4	62.4	99%	2e-08	G
XP_520707.2	PREDICTED: similar to TPA_exp: C9ORF21 isoform 2 [Pan troglodytes]	62.4	62.4	99%	2e-08	UGM
NP_079646.1	hypothetical protein LOC66129 [Mus musculus] >sp Q9D1A0.1 CI021_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]	62.4	62.4	99%	2e-08	UGM
XP_002965388.1	hypothetical protein SELMODRAFT_68006 [Selaginella moellendorffii] >gb EFJ34226.1 hypothetical protein SELMODRAFT_68006 [Selaginella moellendorffii]	62.0	62.0	94%	2e-08	UG
XP_002273449.1	PREDICTED: hypothetical protein [Vitis vinifera] >emb CBI32627.3 unnamed protein product [Vitis vinifera]	61.6	61.6	100%	3e-08	UGM
XP_002334170.1	predicted protein [Populus trichocarpa] >gb EEF07066.1 predicted protein [Populus trichocarpa]	61.2	61.2	55%	4e-08	UG
NP_714542.1	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]	61.2	61.2	99%	4e-08	UGM
ADO28366.1	upf0308 protein c9orf21-like protein [Ictalurus furcatus]	61.2	61.2	99%	4e-08	
XP_002977235.1	hypothetical protein SELMODRAFT_58056 [Selaginella moellendorffii] >gb EFJ21844.1 hypothetical protein SELMODRAFT_58056 [Selaginella moellendorffii]	60.8	60.8	93%	5e-08	UG

Accession	Description	Max score	Total score	Query coverage	E value	Links
ACU20039.1	unknown [Glycine max]	60.8	60.8	91%	5e-08	
XP_002320012.1	predicted protein [Populus trichocarpa] >gb EEE98327.1 predicted protein [Populus trichocarpa]	60.1	60.1	83%	9e-08	UG
XP_795970.1	PREDICTED: hypothetical protein [Strongylocentrotus purpuratus] >ref XP_001176073.1 PREDICTED: hypothetical protein [Strongylocentrotus purpuratus]	60.1	60.1	88%	1e-07	GM
ADE77692.1	unknown [Picea sitchensis]	59.7	59.7	86%	1e-07	
XP_001926014.1	PREDICTED: UPF0765 protein C10orf58 isoform 4 [Sus scrofa]	59.7	59.7	85%	1e-07	UGM
XP_001370596.1	PREDICTED: similar to SFLQ611 [Monodelphis domestica]	59.3	59.3	85%	2e-07	UGM
XP_001368977.1	PREDICTED: similar to C9ORF21 [Monodelphis domestica]	58.9	58.9	99%	2e-07	UGM
AAP06083.1	similar to GenBank Accession Number AK005188 putative related to F3G5 [Schistosoma japonicum]	58.9	58.9	88%	2e-07	
XP_002463942.1	hypothetical protein SORBIDRAFT_01g009350 [Sorghum bicolor] >gb EER90940.1 hypothetical protein SORBIDRAFT_01g009350 [Sorghum bicolor]	58.2	58.2	83%	4e-07	UG
YP_001529200.1	hypothetical protein Dole_1319 [Desulfococcus oleovorans Hxd3] >gb ABW67123.1 hypothetical protein Dole_1319 [Desulfococcus oleovorans Hxd3]	58.2	58.2	78%	4e-07	G
XP_546736.2	PREDICTED: hypothetical protein XP_546736 [Canis familiaris]	58.2	58.2	96%	4e-07	UGM
ACN25853.1	unknown [Zea mays]	57.8	57.8	83%	4e-07	G
NP_001170360.1	hypothetical protein LOC100384338 [Zea mays] >gb ACN36745.1 unknown [Zea mays]	57.8	57.8	83%	5e-07	UG
CAN81555.1	hypothetical protein VITISV_040397 [Vitis vinifera]	57.8	57.8	86%	5e-07	
XP_536403.1	PREDICTED: similar to R53.5 [Canis familiaris]	57.8	57.8	85%	5e-07	UGM
ACU20324.1	unknown [Glycine max]	57.4	57.4	83%	5e-07	
XP_002263959.1	PREDICTED: hypothetical protein [Vitis vinifera]	57.4	57.4	86%	5e-07	UGM
NP_001051153.1	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	57.4	57.4	83%	6e-07	UG

Accession	Description	Max score	Total score	Query coverage	E value	Links
	sativa Japonica Group] >gb EAY91738.1 hypothetical protein OsI_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]					
Q5Z134.2	RecName: Full=UPF0765 protein C10orf58 homolog; Flags: Precursor	57.4	57.4	89%	6e-07	G
NP_001180447.1	selenoprotein U [Gallus gallus] >ref NP_001180448.1 selenoprotein U [Gallus gallus]	57.4	57.4	89%	6e-07	UG
AAI14901.1	Chromosome 1 open reading frame 93 ortholog [Bos taurus] >gb DAA21137.1 hypothetical protein LOC617001 [Bos taurus]	57.4	57.4	94%	6e-07	G
ACR37670.1	unknown [Zea mays]	57.4	57.4	94%	6e-07	
XP_848380.1	PREDICTED: similar to UPF0308 protein C9orf21 [Canis familiaris]	57.4	57.4	99%	6e-07	UGM
NP_001035688.1	hypothetical protein LOC617001 [Bos taurus] >sp Q58CY6.1 CA093_BOVIN RecName: Full=Uncharacterized protein C1orf93 homolog >gb AAX46658.1 hypothetical protein MGC26818 [Bos taurus]	57.4	57.4	94%	6e-07	UGM
CBI33289.3	unnamed protein product [Vitis vinifera]	57.0	57.0	86%	7e-07	
XP_002269002.1	PREDICTED: hypothetical protein, partial [Vitis vinifera]	57.0	57.0	86%	7e-07	UGM
XP_002924432.1	PREDICTED: uncharacterized protein C1orf93-like [Ailuropoda melanoleuca]	57.0	57.0	94%	7e-07	GM
NP_001051154.1	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]	57.0	57.0	83%	9e-07	UG
XP_002525198.1	conserved hypothetical protein [Ricinus communis] >gb EEF37164.1 conserved hypothetical protein [Ricinus communis]	56.6	56.6	55%	9e-07	G
EAY91739.1	hypothetical protein OsI_13380 [Oryza sativa Indica Group]	56.6	56.6	83%	9e-07	
EEE59859.1	hypothetical protein OsJ_12441 [Oryza sativa Japonica Group]	56.6	56.6	83%	1e-06	
ZP_02926219.1	hypothetical protein VspiD_06235 [Verrucomicrobium spinosum DSM 4136]	56.6	56.6	97%	1e-06	

Accession	Description	Max score	Total score	Query coverage	E value	Links
XP_002468056.1	hypothetical protein SORBIDRAFT_01g038790 [Sorghum bicolor] >gb EER95054.1 hypothetical protein SORBIDRAFT_01g038790 [Sorghum bicolor]	56.6	56.6	94%	1e-06	UG
NP_001106912.1	prostamide/PG F synthase [Sus scrofa] >dbj BAF96021.1 prostamide/PG F synthase [Sus scrofa]	56.6	56.6	94%	1e-06	UG
NP_001092155.1	hypothetical protein LOC100049742 [Xenopus laevis] >gb AAI41728.1 LOC100049742 protein [Xenopus laevis]	56.2	56.2	89%	1e-06	UG
ACU24130.1	unknown [Glycine max]	56.2	56.2	83%	1e-06	
Q641F0.2	RecName: Full=UPF0765 protein C10orf58 homolog; Flags: Precursor	56.2	56.2	85%	1e-06	G
NP_001087861.1	chromosome 10 open reading frame 58 [Xenopus laevis] >gb AAH82387.1 MGC81827 protein [Xenopus laevis]	56.2	56.2	85%	1e-06	UG
XP_002708298.1	PREDICTED: hypothetical protein [Oryctolagus cuniculus]	56.2	56.2	99%	1e-06	UGM
ACU18761.1	unknown [Glycine max]	56.2	56.2	83%	1e-06	
CBI33071.3	unnamed protein product [Vitis vinifera]	56.2	56.2	83%	1e-06	
XP_002263922.1	PREDICTED: hypothetical protein [Vitis vinifera] >emb CAN81556.1 hypothetical protein VITISV_040398 [Vitis vinifera]	55.8	55.8	83%	2e-06	UGM
CBI33293.3	unnamed protein product [Vitis vinifera]	55.8	55.8	83%	2e-06	
ACO12061.1	C10orf58 homolog precursor [Lepeophtheirus salmonis]	55.8	55.8	90%	2e-06	
EEC75002.1	hypothetical protein OsI_11064 [Oryza sativa Indica Group]	55.8	55.8	55%	2e-06	
XP_002533575.1	conserved hypothetical protein [Ricinus communis] >gb EEF28810.1 conserved hypothetical protein [Ricinus communis]	55.8	55.8	83%	2e-06	G
XP_002191279.1	PREDICTED: hypothetical protein [Taeniopygia guttata]	55.5	55.5	98%	2e-06	UGM
NP_001180474.1	selenoprotein U [Oryzias latipes]	55.5	55.5	89%	2e-06	G
CAG09571.1	unnamed protein product [Tetraodon nigroviridis]	55.5	55.5	97%	2e-06	
Q8TBF2.1	RecName: Full=Uncharacterized protein C1orf93 >gb AAP97295.1 AF425266_1 unknown protein [Homo sapiens] >gb AAH22547.1 Chromosome 1 open reading frame 93 [Homo sapiens] >gb EAW56087.1 chromosome 1 open reading frame 93, isoform CRA_e [Homo sapiens] >emb CAX30827.1 chromosome 1 open reading frame 93 [Homo sapiens]	55.5	55.5	88%	2e-06	G

Alignments

>**gb|EFW45143.1** | predicted protein [Capsaspora owczarzaki ATCC 30864]
Length=297

Score = 303 bits (777), Expect = 4e-81, Method: Compositional matrix adjust.
Identities = 142/142 (100%), Positives = 142/142 (100%), Gaps = 0/142 (0%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN
Sbjct 137 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 196

Query 61  GORFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWRKANKNHPNADLQGDGLQTGG 120
          GORFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWRKANKNHPNADLQGDGLQTGG
Sbjct 197  GORFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWRKANKNHPNADLQGDGLQTGG 256

Query 121  IYLVGPGADSAIHF FAFNEYDHP 142
          IYLVGPGADSAIHF FAFNEYDHP
Sbjct 257  IYLVGPGADSAIHF FAFNEYDHP 278
```

>**gb|EFA74551.1** | hypothetical protein PPL_00049 [Polysphondylium pallidum PN500]
Length=662

Score = 73.9 bits (180), Expect = 6e-12, Method: Composition-based stats.
Identities = 47/141 (34%), Positives = 75/141 (54%), Gaps = 11/141 (7%)

```
Query 2  DQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNG 61
          ++R ++ +LRRFGC +C Q + +KP+LD G+ ++ +G R E FI G
Sbjct 144  NKRCVIAVLRFRFGCLVCR LQCMDLSS LKPKLDRMGIALIAIGF-ERVGLEDFIA-----G 197

Query 62  QRFPAEVYIDPEQTAYKARGLQRVGLLHF---LSWTAISEWRK-ANKNHPNADLQGDGLQ 117
          F E+YID ++ Y+A L+R+G L +S +RK A + ++ +GDGLQ
Sbjct 198  GFFNGEIYIDRSRSVYRALSLKRMGFWDTTIGLMDPRLSVYRKEAKEKGLPSNFRGDGLQ 257

Query 118  TGGIYLVGPGADSAIHF FAFNE 138
          G +VGP A H+ F +
Sbjct 258  LGATLVVGP KPQGA-HYDFRQ 277
```

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

Score = 73.2 bits (178), Expect = 1e-11, Method: Compositional matrix adjust.
Identities = 47/142 (34%), Positives = 71/142 (50%), Gaps = 13/142 (9%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          KD++ ++ R FGC LC ++A + K Q+DAAGV +VL+G GN A+ F +
Sbjct 104  KDRKAVIGFARHFGCVLCRKRADVLASQKSQMDAAGVALVLIGPGNIEQAKAFADQT--- 160

Query 61  GORFPAEVYIDPEQTAYKARGLQRVGLLHF--LSWTAISE-WRKANKNHPNADLQGD--- 114
          +FP E+Y DP T++ A F L+ T I E + + + Q D
Sbjct 161  --KFPGEIYADPNHTSFNALKFVSGVFTTFTPLAATKIIELYVEGYRQDWGLSFQKDTMN 218

Query 115  --GLQTGGIYLVGPGADSAIHF 134
          G Q GGI + GPG D+ ++
Sbjct 219  RGGWQQGGILVAGPGGDNILYL 240
```

>**ref|XP_001752693.1** | **UG** predicted protein [Physcomitrella patens subsp. patens]

gb|EDQ82564.1 | **G** predicted protein [Physcomitrella patens subsp. patens]
Length=293

GENE ID: 5915868 PHYPADRAFT 65214 | hypothetical protein
[Physcomitrella patens subsp. patens] (10 or fewer PubMed links)

Score = 71.6 bits (174), Expect = 3e-11, Method: Compositional matrix adjust.
Identities = 43/133 (33%), Positives = 74/133 (56%), Gaps = 11/133 (8%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          +DQ V+L +LRRFGC LC Q+ + ++ QL+A V++V +G ++ E+F EN
Sbjct 138  EDQPVVHLVLRFRFGCQLCRGQSVEMAKMLSQLEANNVRVVGIGL-EKFGLEEFEEEN---- 192
```



```

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVG----LLHFLSWTAISEWRKANKNHPNADLQGDGL 116
          + +E+YID E+ +KA L +VG + + ++ E + K+ P + QGDG
Sbjct 193 -NYWKSLEYIDNEKKIHKALALTKVGVVGTFFMMLFANKSVKEAAQTKDTP-GNFQGDGR 250

Query 117 QTGGIYLVGPGAD 129
          Q G +++ G +
Sbjct 251 QLGATFVMAKGGGE 263

```

>**gb|AC013446.1** | Clorf93 [Esox lucius]
Length=225

Score = 71.6 bits (174), Expect = 3e-11, Method: Compositional matrix adjust.
Identities = 49/155 (32%), Positives = 75/155 (49%), Gaps = 21/155 (13%)

```

Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQ-LDAAGVKIVLVGTGNRYFAEKFIENVP 59
          +D++ ++I +R F C C E + I + L AG+++V++G + + + F
Sbjct 51  QDRKSVIIFVRNFLCHTCKEYVDDLRIPEVLKEAGLRLVIGQSSHHHIQSFC----- 105

Query 60  NGQRFPAEVYIDPEQTAYKARGLQ---VGLL----HFLSWTAI----SEWRKANKNHP 107
          + R+P E+Y+DPE+ YK G+ R VGL H S + S WR P
Sbjct 106 SLTRYPHEMYVDPERCIYKKGMMNRGEISVGLAQPSPHVKSGMLVGHMKSIIWRAMTS--P 163

Query 108 NADLQGDGLQTGGIYLVGPGADSAI-HFAFNEYDH 141
          D QGD Q GG + GPG++ HF N DH
Sbjct 164 IFDFQGDPRQGGGAIAGPGSEVHFAHFDMNRLDH 198

```

>**ref|NP_030274.1** | **UG** unknown protein [Arabidopsis thaliana]
sp|Q9ZUU2.2|U308_ARATH **G** RecName: Full=UPF0308 protein At2g37240, chloroplastic; f
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 = 71.2 bits (173), Expect = 4e-11, Method: Compositional matrix adjust.
Identities = 45/146 (31%), Positives = 72/146 (50%), Gaps = 29/146 (19%)

```

Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVP 60
          KD++ ++ R FGC LC ++A+++ E K +DA+GV +VL+G G+ A F+E
Sbjct 91  KDRKAVVAFARHFGCVLCRRAAYLAEEKDVMDASGVALVLIGPGSIDQANTFVEQT--- 147

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWRKANKNHPNADLQG----- 113
          +F EVY DP +Y+A L F+S +++ KA + ++G
Sbjct 148 --KFKGEVYADPNHASYE-----LEFVSGVSVTFTPKAAMKILESMEGYRQDWKL 197

Query 114 -----DGLQTGGIYLVGPGADS 130
          G Q GGI + GPG D+
Sbjct 198 SFMKDTVERGGWQGGILVAGPGKDN 223

```

>**ref|XP_002881499.1** | **G** hypothetical protein ARALYDRAFT_482716 [Arabidopsis lyrata s
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 = 70.5 bits (171), Expect = 7e-11, Method: Compositional matrix adjust.
Identities = 45/146 (31%), Positives = 71/146 (49%), Gaps = 29/146 (19%)

```

Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVP 60
          KD++ ++ R FGC LC ++A+++ E K +DA+GV +VL+G G+ A F+E
Sbjct 91  KDRKAVVAFARHFGCVLCRRAAYLAEEKDVMDASGVTTLVLIGPGSIDQANTFMEQT--- 147

```

```

Query 61  GQRFP AEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWRKANKNHPNADLQG----- 113
          +F  EVY DP  +Y+A          L F+S  ++  KA  +  ++G
Sbjct 148  --KFKGEVYADPNHASYE A-----LEFVSGVTVTFTPKAAMKILESYMEGYRQDWKL 197

Query 114 -----DGLQTGGIYLVGPGADS 130
          G Q GGI + GPG D+
Sbjct 198  SFMKD TVERGGWQQGGILVAGPGKDN 223

```

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

Score = 69.7 bits (169), Expect = 1e-10, Method: Compositional matrix adjust.
Identities = 47/155 (31%), Positives = 76/155 (50%), Gaps = 21/155 (13%)

```

Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQ-LDAAGVKIVLVGTGNRYFAEKFIENVP 59
          +D++ ++I +R F C C E  + I + L AG+++V++G + + E F ++ G
Sbjct 50  QDRKSVVIFVRNFLCHTCKEYVDDL SRIPAEVLKEAGLRLVIGQSSHHHIESFC-SLTG 108

Query 60  NGQRFP AEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAI----SEWRKANKNHP 107
          +P ++Y+DPE+ YK G++R +          H S  +  S WR  P
Sbjct 109  ----YPHDIYVDPERCIYKRLGMRRGEMSVESTKPSPHVKSGMLVGHMKS MWRAMTS--P 162

Query 108  NADLQGDGLQTGGIYLVGPGADSAI-HFAFNEYDH 141
          D QGD Q GG +VGPG++ HF N DH
Sbjct 163  IFDFQGDPRQGGAIIVGPGSEVHFAHFD MNRLDH 197

```

>**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 = 69.7 bits (169), Expect = 1e-10, Method: Compositional matrix adjust.
Identities = 47/155 (31%), Positives = 76/155 (50%), Gaps = 21/155 (13%)

```

Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQ-LDAAGVKIVLVGTGNRYFAEKFIENVP 59
          +D++ ++I +R F C C E  + I + L AG+++V++G + + E F ++ G
Sbjct 50  QDRKSVVIFVRNFLCHTCKEYVDDL SRIPAEILKEAGLRLVIGQSSHHHIESFC-SLTG 108

Query 60  NGQRFP AEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAI----SEWRKANKNHP 107
          +P ++Y+DPE+ YK G++R +          H S  +  S WR  P
Sbjct 109  ----YPHDIYVDPERCIYKRLGMRRGEMSVESAKPSPHVKSGMLVGHMKS MWRAMTS--P 162

Query 108  NADLQGDGLQTGGIYLVGPGADSAI-HFAFNEYDH 141
          D QGD Q GG +VGPG++ HF N DH
Sbjct 163  IFDFQGDPRQGGAIIVGPGSEVHFAHFD MNRLDH 197

```

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

Score = 69.3 bits (168), Expect = 1e-10, Method: Compositional matrix adjust.
Identities = 47/155 (31%), Positives = 76/155 (50%), Gaps = 21/155 (13%)

```

Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQ-LDAAGVKIVLVGTGNRYFAEKFIENVP 59
          +D++ ++I +R F C C E  + I + L AG+++V++G + + E F ++ G
Sbjct 58  QDRKSVIIFVRNFLCHTCKEYVDDL SRIPAEVLKEAGLRLVIGQSSHHHIESFC-SLTG 116

Query 60  NGQRFP AEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAI----SEWRKANKNHP 107
          +P ++Y+DPE+ YK G++R +          H S  +  S WR  P
Sbjct 117  ----YPHDMYVDPERCIYKRLGMRRGEMSVESTKPSPHVKSGMLVGHMKS MWRAMTS--P 170

Query 108  NADLQGDGLQTGGIYLVGPGADSAI-HFAFNEYDH 141
          D QGD Q GG +VGPG++ HF N DH
Sbjct 171  IFDFQGDPRQGGAIIVGPGSEVHFAHFD MNRLDH 205

```

>**ref|XP_002574167.1** | **UG** PRX_like2 domain-containing protein [Schistosoma mansoni]
emb|CAZ30400.1 | **G** PRX_like2 domain-containing protein [Schistosoma mansoni]
Length=203

GENE ID: 8355480 Smp 194940 | PRX like? domain-containing protein
[Schistosoma mansoni] (10 or fewer PubMed links)

Score = 68.9 bits (167), Expect = 2e-10, Method: Compositional matrix adjust.
Identities = 44/147 (30%), Positives = 73/147 (50%), Gaps = 11/147 (7%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          +DQ ++ RR GC C +A ++ +KP LDA +K++ + T + ++F++
Sbjct 30  RDQTCIITFFRRLGCKFCRLEAKNLSYLKPVLDARNIKLMGI-TFDEGGVKEFLD----- 83

Query 61  GQRFPAEVYIDPEQTAYKARGLQRV---GLLHFLSWTAISEWRKANKNHPNADLQGDGL 116
          G F ++Y+D E+ YKA ++V G L+ S KA + ++ GDG
Sbjct 84  GHYFDGDLYLDRERKTYKALEYKVSACSGFCSLLTKAGRSLNSKAKAANIPGNMMSGDGW 143

Query 117  QTGGIYLVGPGADSAIHFAFNE-YDHP 142
          QTGG+ +V G HF E +HP
Sbjct 144  QTGGLLVVEKGGKVLYHFEQKEVVNHP 170
```

>**gb|AC051726.1** | Clorf93 homolog [Rana catesbeiana]
Length=201

Score = 68.6 bits (166), Expect = 3e-10, Method: Compositional matrix adjust.
Identities = 43/140 (31%), Positives = 70/140 (50%), Gaps = 11/140 (7%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          KD ++ LRRFGC +C A V ++K LDA +++++ +G E F++
Sbjct 30  KDNTSVIFFLRRFGCQICRWIAKDVSQLKESLDANQIRLIGIPETVGLQE-FLD----- 83

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR---KANKNHPNADLQGDGL 116
          G+ F E+Y+D + +YK G +R L + + R KAN + + GD L
Sbjct 84  GK YFTGELYLDESKQSYKELGFKRYNALSIVPAALGKKVRDIVTKANADGVQGNFSGDLL 143

Query 117  QTGGIYLVGPGADSA-IHFA 135
          Q+GG+ +V G + A +HF
Sbjct 144  QSGGMLVVSKGGEKALLHFV 163
```

>**ref|XP_002587381.1** | **UG** hypothetical protein BRAFLDRAFT_96268 [Branchiostoma floridae]
>**gb|EEN43392.1** | **G** hypothetical protein BRAFLDRAFT_96268 [Branchiostoma floridae]
Length=185

GENE ID: 7248344 BRAFLDRAFT 96268 | hypothetical protein
[Branchiostoma floridae] (10 or fewer PubMed links)

Score = 68.6 bits (166), Expect = 3e-10, Method: Compositional matrix adjust.
Identities = 42/131 (33%), Positives = 68/131 (52%), Gaps = 10/131 (7%)

```
Query 6  LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRF 65
          +L+ LRRFGC +C A+ + ++KPQLDAA V +V VG ++F++ G+ F
Sbjct 35  VLLFLRRFGCQVCRWTATELSKLPQLDAANVNLVGVGP-EEVGDEFVQ-----GKFFA 88

Query 66  AEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR---KANKNHPNADLQGDGLQTGGI 121
          ++Y+D + YK G +R L+ + A + R KA + +GD LQ GG
Sbjct 89  GDLYVDETKQCYKDLGYRRYNALNVIPAAASKKSRDVINKAKAEGIPGNFKGDLQAGGT 148

Query 122  YLVGPGADSAI 132
          +V G + +
Sbjct 149  LIVVAGGEKVL 159
```

>**gb|AAW27591.1** | SJCHGC05103 protein [Schistosoma japonicum]
>**emb|CAX73680.1** | hypothetical protein [Schistosoma japonicum]
Length=203

Score = 67.8 bits (164), Expect = 4e-10, Method: Compositional matrix adjust.
Identities = 43/147 (30%), Positives = 73/147 (50%), Gaps = 11/147 (7%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          +D+ ++ RR GC C +A ++ +KP LD +K++ + T + ++F+ +
Sbjct 30  RDRTCIVTFFRRMGCKFCRLEAKNLSYLKPALDTRNIKLIGI-TFDVGGVKEFL-----D 83

Query 61  GQRFPAEVYIDPEQTAYKARGLQRV---GLLHFLSWTAISEWRKANKNHPNADLQGDGL 116
          G F ++Y+DPE+ YKA G ++V G + S A + KA +L GDG
Sbjct 84  GHYFDGDLYLDPERMTYKALGYKKVSPCSGAI SLFSKAARALNSKAKAAKIPGNLSGDGW 143
```

Query 117 QTGGIYLVGPGADSAIHFAFNE-YDHP 142
 OTGG+ +V G ++ E HP
 Sbjct 144 QTGGLLVVEKGGKILYYEQEVRHP 170

>ref|XP_001762195.1| **UG** predicted protein [Physcomitrella patens subsp. patens]
 gb|EDQ72987.1| **G** predicted protein [Physcomitrella patens subsp. patens]
 Length=232

GENE ID: 5925481 PHYPADRAFT_163182 | hypothetical protein
 [Physcomitrella patens subsp. patens] (10 or fewer PubMed links)

Score = 67.4 bits (163), Expect = 6e-10, Method: Compositional matrix adjust.
 Identities = 49/160 (31%), Positives = 77/160 (49%), Gaps = 20/160 (12%)

Query 2 DQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLV-----GTGNR 47
 DQ VL+ +LRRFGC LC A + +I P L+A GV+I+ + R
 Sbjct 40 DQPVLIHVLRRFGCQLCRGGA VEMGKIFPDLEAHGVRIIGIVRWKSLVKDVCADVDARR 99

Query 48 YFAEKFIENVPGNGQRFPAEYIDPEQTAYKARGLQRVGLLH----FLSWTAISEWRKAN 103
 EK G + E+YID + +KA +Q+VG+L +S ++ + K
 Sbjct 100 LGIEKVGLED FQKGGFWK GELYIDNGKKIHKALNIQKVGILSSVKMMVSNKSVKDAIKKT 159

Query 104 KNHPNADLQGDGLQTGGIYLVGPGADSAIHFAFNEY-DHP 142
 K+ P D +GDG Q G +++ G ++ + F + DHP
 Sbjct 160 KDTP-GDFKGDGRQLGATFVLAKGGETLLDFRQEHFGDHP 198

>gb|ACI67539.1| Clorf93 homolog [Salmo salar]
 Length=200

Score = 67.4 bits (163), Expect = 6e-10, Method: Compositional matrix adjust.
 Identities = 44/143 (31%), Positives = 70/143 (49%), Gaps = 17/143 (11%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVG---TGNRYFAEKFIENV 57
 +D+ V+L LRRFGC +C A+ + +++P L A G+ +V +G TG + F E
 Sbjct 29 RDKPVVLFLLRRFGCQVCRWTAEEISKLEPDLTAHGIALVIGIGPEETGLKEFKE----- 82

Query 58 PGNGQRFPAEYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR----KANKNHPNADLQ 113
 G F ++YID ++ YK G +R L + + R KA + G
 Sbjct 83 ---GGFFKGDLYIDEKKQCYKDLGFKRYTALSVPVPAALGKKIREVTTKAKAQGIQGNFTG 139

Query 114 DGLQTGGIYLVGPGADSA-IHFA 135
 D LQ+GG+ +V G + +HF
 Sbjct 140 DLLQSGGMLIVAKGGEKVLHVF 162

>gb|AD027754.1| uncharacterized protein clorf93-like protein [Ictalurus furcatus]
 Length=201

Score = 67.4 bits (163), Expect = 7e-10, Method: Compositional matrix adjust.
 Identities = 43/142 (31%), Positives = 70/142 (50%), Gaps = 17/142 (11%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVG---TGNRYFAEKFIENV 57
 KD+ V++ LRRFGC +C A+ V +++ L GV ++ +G TG + F
 Sbjct 30 KDKTVVMFFLLRRFGCQICRWAAAEEVSKLEKDLRENGVALIGIGPEETGLKEFE----- 82

Query 58 PGNGQRFPAEYIDPEQTAYKARGLQRVGLLHFLSWTAISEWRKANKNHPNADLQ----G 113
 +G F E+YID ++ YK G +R ++ L + R+ N +0 G
 Sbjct 83 --DGGFFKGEIYIDEKKQCYKELGFKRYNAINVLPALGKKVREIASKASNEGIQGNFSG 140

Query 114 DGLQTGGIYLVGPGADSA-IHF 134
 D LQ+GG+ +V G + +HF
 Sbjct 141 DLLQSGGMLIVAKGGEKVLHVF 162

>emb|CAX73679.1| hypothetical protein [Schistosoma japonicum]
 Length=203

Score = 67.0 bits (162), Expect = 7e-10, Method: Compositional matrix adjust.
 Identities = 39/131 (30%), Positives = 68/131 (52%), Gaps = 10/131 (7%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
 +D+ ++ RR GC C +A ++ +KP LD +K++ + T + ++F++

```

Sbjct 30 RDRTCIVTFFRRMGCKFCRLEAKNLSYLKPALDTRNIKLIGI-TFDVGGVKEFLD----- 83
Query 61 GORFPAEVYIDPEQTAYKARGLQRV----GLLHFLSWTAISEWRKANKNHPNADLQGDGL 116
      G F ++Y+DPE+ YKA G ++V G++ S + KA +L GDG
Sbjct 84 GHYFDGDLYLDPERMTYKALGYKKVSPCSGVISLFSKAGRALNSKAKAAKIPGNLSGDGW 143
Query 117 QTGGIYLVGPG 127
      QTGG+ +V G
Sbjct 144 QTGGLLVVEKG 154

```

>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 = 67.0 bits (162), Expect = 8e-10, Method: Compositional matrix adjust.
 Identities = 48/133 (37%), Positives = 67/133 (51%), Gaps = 15/133 (11%)

```

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
      KD ++LLR FGC C E A+ + E KP+ DAAGVK++ VG G A +P
Sbjct 114 KDGIAAVVLLRHFGCVCCWELATALKEAKPRFDAAGVKLIAGVGTGTPDKARILATRLP-- 171
Query 61 GORFPAE-VYIDPEQTAYKARGLQ-RVGLLHF-----LSWTAISEWRKANKNHP--NAD 110
      FP E +Y DPE+ AY GL +G F ++ SE R+A KN+ +
Sbjct 172 ---FPMECLYADPERKAYDVLGLYFGLGRTFFNPASTKVFSEIRFSEIREATKNYTIETPE 228
Query 111 LQGDGLQTGGIYL 123
      + LQ GG ++
Sbjct 229 DRSSVLQQGGTFV 241

```

>ref|XP_002326159.1| **UG** predicted protein [Populus trichocarpa]
 gb|EEE71829.1| **G** predicted protein [Populus trichocarpa]
 Length=200

GENE ID: 7459323 POPTRDRAFT_589376 | hypothetical protein [Populus trichocarpa]
 (10 or fewer PubMed links)

Score = 66.6 bits (161), Expect = 9e-10, Method: Compositional matrix adjust.
 Identities = 44/143 (31%), Positives = 67/143 (47%), Gaps = 29/143 (20%)

```

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
      KD++ ++ R FGC LC +A ++ K +DA+GV +VL+G G+ A+ F E
Sbjct 44 KDRKAVVAFARHFGCVLCRRRADYLAAKKDIMDASGVALVLIGPGSVDQAKTFSEQT--- 100
Query 61 GORFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTA-----ISEWRKANKNHPNA 109
      +F EVY DP ++YKA L F+S + I + + +
Sbjct 101 --KFKGEVYADPSHSSYKA-----LQFVSGVSTTFTPKAGLKIISYMEGYRQDWKL 150
Query 110 DLQGD-----GLQTGGIYLVGPG 127
      +GD G Q GGI + GPG
Sbjct 151 SFEGDTVAKGGWQQGGIIVAGPG 173

```

>ref|NP_001017220.1| **UGM** hypothetical protein LOC549974 [Xenopus (Silurana) tropi
 sp|Q28IJ3.1|CA093 XENTR **G** RecName: Full=Uncharacterized protein Clorf93 homolog
 emb|CAJ83264.1| **G** novel protein [Xenopus (Silurana) tropicalis]
 Length=201

GENE ID: 549974 clorf93 | chromosome 1 open reading frame 93
 [Xenopus (Silurana) tropicalis]

Score = 66.6 bits (161), Expect = 9e-10, Method: Compositional matrix adjust.
 Identities = 43/139 (31%), Positives = 69/139 (50%), Gaps = 11/139 (7%)

```

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
      K++ +L+ LRRFGC +C A + ++K DA +++V +G E F+E
Sbjct 30 KEKTTVLLFLRRFGCQICRWIAKDIGLKLASCDAHQIRLVGIGPEEVGLKE-FLE----- 83

```

```

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR----KANKNHPNADLQGDGL 116
          G F E+YID + +YK G +R L + + R KAN + + GD L
Sbjct 84  GNFFNGELYIDESKESYKTLGFKRYSALSVIPAALGKKVIRDIVTKANADGVQGNFSGDLL 143

Query 117  QTGGIYLVGPGADSA-IHF 134
          Q+GG+ +V G + +HF
Sbjct 144  QSGGMLIVSKGGEKVLLHF 162

```

>emb|CBN80880.1| Uncharacterized [Dicentrarchus labrax]
Length=201

Score = 66.6 bits (161), Expect = 1e-09, Method: Compositional matrix adjust.
Identities = 45/139 (33%), Positives = 68/139 (49%), Gaps = 11/139 (7%)


```

Query 1  KDQRVLLILLRRFGCSLCEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          +DQ V+L LRRFGC +C AS + +++P L A+GV + VG AE F E
Sbjct 30  QDQPVVLFLLRRFGCQVCRWMASEISKLEPDLRASGVALAGVGPPEEFLAE-FKE----- 83

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR----KANKNHPNADLQGDGL 116
          G F F +Y+D + YK G +R + + + R KA + + GD L
Sbjct 84  GGFYKSGSLYVDETKKTYKDLGFKRYTAISVVPALGKKVIRDIAAKAKADGIQGNFSGDLL 143

Query 117  QTGGIYLVGPGADSA-IHF 134
          Q+GG+ +V G + +HF
Sbjct 144  QSGGMLIVAKGGEKVLLHF 162

```

>ref|XP_002666723.1|  PREDICTED: UPF0308 protein C9orf21 homolog [Danio rerio]
Length=223

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

Score = 65.9 bits (159), Expect = 2e-09, Method: Compositional matrix adjust.
Identities = 50/153 (33%), Positives = 71/153 (47%), Gaps = 23/153 (15%)


```


Query 4  RVLLILLRRFGCSLCEQASHVLEIKPQ--LDAAGVKIVLVGTGNRYFAEKFIENVPNG 61
          + ++I +R F C C E + +I PQ L + V++V++G + + F ++ G
Sbjct 53  KAIVIFVRHFLCYTCKEYVEDLGKI-PQHVLQDSNVRLLVIGQSSYSHIQGFC-SLTG-- 108

Query 62  QRFP AEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPNA 109
          FP E+Y+DPE+ YK GL+R LS + S WR P
Sbjct 109  --FPHEIYVDPERQIYKRLGLRRGETYMETPSVSPHVKSSMLSGSLKSVWRAMTS--PVF 164

Query 110  DLQGDGLQTGGIYLVGPGADSAI-HFAFNEYDH 141
          D QGD Q GG +VGPG D HF N DH
Sbjct 165  DFQGDPPQQGGALIVGPGPDVHFHAFDMNRLDH 197

```

>ref|XP_002866692.1|  hypothetical protein ARALYDRAFT_496819 [Arabidopsis lyrata ssp. lyrata]

gb|EFH42951.1 |  hypothetical protein ARALYDRAFT_496819 [Arabidopsis lyrata subsp. lyrata]
Length=265

GENE ID: 9302764 ARALYDRAFT_496819 | hypothetical protein
[Arabidopsis lyrata subsp. lyrata]

Score = 65.9 bits (159), Expect = 2e-09, Method: Compositional matrix adjust.
Identities = 47/133 (36%), Positives = 67/133 (51%), Gaps = 15/133 (11%)

```

Query 1  KDQRVLLILLRRFGCSLCEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          KD ++LLR FGC C E A+ + E KP+ DAAGVK++ VG G A +P
Sbjct 104  KDGIAAVVLLRHFGCVCCWELATALKEAKPRFDAAGVKLIAGVGTGTPDKARILATRLP-- 161

Query 61  GQRFPAE-VYIDPEQTAYKARGLQR-VGLLHF-----LSWTAISEWRKANKNHP---NAD 110
          FP E +Y DPE+ AY GL +G F ++ +E R+A KN+ +
Sbjct 162  ---FPMECLYADPERKAYDVLGLYYGLGRFTFFNPASTKVFSRFNEIREATKNYTIEATPE 218

Query 111  LQGDGLQTGGIYL 123
          + LQ GG ++
Sbjct 219  DRSSVLQQGGTFV 231

```




>emb|CAG07032.1| unnamed protein product [Tetraodon nigroviridis]
Length=191

Score = 65.1 bits (157), Expect = 3e-09, Method: Compositional matrix adjust.
Identities = 43/139 (31%), Positives = 68/139 (49%), Gaps = 11/139 (7%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          +DQ V+L LRRFGC +C A+ + +++ +L A GV LVG G K + +
Sbjct 29  RDQPVVLFLLRRFGCQICRWIAAEISKLEAELRAGGV--ALVGIGPEEVGLKEFK----D 82

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR----KANKNHPNADLQGDGL 116
          G F +YID ++ YK G +R + + + R KA + + + GD L
Sbjct 83  GGFFKGSIIYIDEKKKTYKDLGFKRYTAISVPPAAMGKKVRDVAAKAKADGVEGNFSGDLL 142

Query 117  QTGGIYLVGPGADSA-IHF 134
          Q+GG+ +V G + +HF
Sbjct 143  QSGGMLIVAKGGKGVLLHF 161
```

>ref|XP_639255.1|  hypothetical protein DDB_G0283129 [Dictyostelium discoideum AX4]
>ref|EAL65894.1|  hypothetical protein DDB_G0283129 [Dictyostelium discoideum AX4]
Length=883


GENE ID: 8623940 DDB_G0283129 | hypothetical protein
[Dictyostelium discoideum AX4] (10 or fewer PubMed links)

Score = 65.1 bits (157), Expect = 3e-09, Method: Composition-based stats.
Identities = 39/140 (28%), Positives = 75/140 (54%), Gaps = 11/140 (7%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          +++R+++ + RRFGC +C QA + +KP+LD G+++V +G F E+ +E
Sbjct 467  ENKRIVVAIFRRFGCLICRLQALDLSALKPKLDKIGIELVGIG-----FDEEGLEEFQ-Q 520

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLH----FLSWTAISEWRKANKNHPNADLQGDGL 116
          + F ++Y+D ++ Y+A L+R L FL + +R+ + + + + DG
Sbjct 521  LKFFAGKIYLDKTRSVYRALNLKRRSKLTTYELFLDPRVMVYRRIKEMGFSSNYRKDGF 580

Query 117  QTGGIYLVGPGADSAIHFAF 136
          Q G ++GP A H+ F
Sbjct 581  QLGATMVLGPKPQEA-HYDF 599
```

>ref|XP_001493974.2|  PREDICTED: similar to UPF0308 protein C9orf21 [Equus caballus]
Length=232


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

Score = 65.1 bits (157), Expect = 3e-09, Method: Compositional matrix adjust.
Identities = 46/154 (30%), Positives = 75/154 (49%), Gaps = 20/154 (12%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPG 59
          +++R +++ +R F C +C E + +I K L A V +++G + + E F + + G
Sbjct 60  RERRAVVVFMRHFLCYICKEYVEDLAKIPKSFLQEANVTLIVIGQSSYHHIEPFCK-LTG 118

Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLLHF-----LSWTAISEWRKANKNHPN 108
          + E+Y+DPE+ YK G++R + F LS + S WR P
Sbjct 119  ----YSHEIYVDPEREIYKRLGMKRGEEIAFSGKSPHIKSNILSGSIRSLWRAMTG--PL 172

Query 109  ADLQGDGLQTGGIYLVGPGAD-SAIHFNEFDH 141
          D QGD Q GG ++GPG + IH N DH
Sbjct 173  FDFQGDPAQQGGTLILGPGNNIHFHICDRNRLDH 206
```

>ref|XP_002664824.1|  PREDICTED: hypothetical protein [Danio rerio]
Length=201

GENE ID: 100332184 LOC100332184 | hypothetical protein LOC100332184
[Danio rerio]

Score = 64.7 bits (156), Expect = 4e-09, Method: Compositional matrix adjust.
Identities = 42/142 (30%), Positives = 70/142 (50%), Gaps = 17/142 (11%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVG---TGNRYFAEKFIENV 57
```



```

Sbjct 30  ++0 V+L  LRRFGC +C  A+ V +++  L A G+ +V +G  TG + F
          REQAVVLFLLRRFGCQVCRWMAAEVSKLEKDLKAHGIALVIGPEETGVKEFK----- 82
Query 58  PGNGQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR----KANKNHPNADLQG 113
          +G F ++YID + YK G +R ++ + + R KA+ + G
Sbjct 83  --DGGFFKGDIIYIDEMKQCYKDLGFKRYNAINVVPAA MGKKVREIASKASAEGIQGNFSG 140
Query 114 DGLQTGGIYLVGPGADSA-IHF 134
          D LQ+GG+ +V G + +HF
Sbjct 141 DLLQSGGMLIVAKGGEKVLLHF 162

```

>ref|NP_998478.1| **UGM** hypothetical protein LOC406605 [Danio rerio]
 sp|Q6NV24.1|CA093 **DANRE** **G** RecName: Full=Uncharacterized protein Clorf93 homolog
 gb|AAH68342.1| **G** Zgc:85644 [Danio rerio]
 Length=201

GENE ID: 406605 zgc:85644 | zgc:85644 [Danio rerio] (10 or fewer PubMed links)

Score = 64.7 bits (156), Expect = 4e-09, Method: Compositional matrix adjust.
 Identities = 42/142 (30%), Positives = 70/142 (50%), Gaps = 17/142 (11%)

```

Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVG---TGNRYFAEKFIENV 57
          ++0 V+L  LRRFGC +C  A+ V +++  L A G+ +V +G  TG + F
Sbjct 30  REQAVVLFLLRRFGCQVCRWMAAEVSKLEKDLKAHGIALVIGPEETGVKEFK----- 82
Query 58  PGNGQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR----KANKNHPNADLQG 113
          +G F ++YID + YK G +R ++ + + R KA+ + G
Sbjct 83  --DGGFFKGDIIYIDEMKQCYKDLGFKRYNAINVVPAA MGKKVREIASKASAEGIQGNFSG 140
Query 114 DGLQTGGIYLVGPGADSA-IHF 134
          D LQ+GG+ +V G + +HF
Sbjct 141 DLLQSGGMLIVAKGGEKVLLHF 162

```

>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 = 64.3 bits (155), Expect = 5e-09, Method: Compositional matrix adjust.
 Identities = 42/133 (32%), Positives = 69/133 (52%), Gaps = 15/133 (11%)

```

Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          ++ + ++ LR FGC C E A+ + E KP+ DAAG K++ +G G A+ E +P
Sbjct 26  RNGKAIVAFLRHFGCPFCWEFAAALREAKPKFDAAGFKLITIGVGPSSKAQVLSEKLP-- 83
Query 61  GQRFPAE-VYIDPEQTAYKARGLQRVGLLHFLSWTAISEWRKANKNHP-----NADLQGD 114
          FPA+ +Y DP++ AY A GL +L+ ++ + + +K N D+ D
Sbjct 84  ---FPADCLYADPDRKAYDALGLYHGVARTWLNPA SMQIFTRLDKVADAVKGWNRDVMPP 140
Query 115 G----LQTGGIYL 123
          LQ GG+Y+
Sbjct 141 NTAATLQGGGVYV 153

```

>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 = 63.9 bits (154), Expect = 7e-09, Method: Compositional matrix adjust.
 Identities = 45/154 (30%), Positives = 73/154 (48%), Gaps = 20/154 (12%)

```

Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPG 59
          +++R +++ +R F C +C E + +I K L A V ++++G + + E F
Sbjct 54  RERRAVVVFVRHFLCYICKEYVEDLAKIPKSFLEQANVTLIVIGQSSYHHIEPFCRLT-- 111
Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
          + E+Y+DPE+ YK G++R + + + LS + S WR P
Sbjct 112  ---GYSHEIYVDPEREIYKRLGMKRGEIASSGQSPHVKSNNLLSGSLQSLWRAVGT--PL 166

```

Query 109 ADLQGDGLQTGGIYLVGPGAD-SAIHFANNEYDH 141
 D QGD Q GGI ++GPG + IH N DH
 Sbjct 167 FDFQGDPAQQGGILILGPGNNIHFHRDRNRLDH 200

>ref|NP_001087128.1| **UG** chromosome 1 open reading frame 93 [Xenopus laevis]
 sp|Q6AZG8.1|CA093 XENLA **G** RecName: Full=Uncharacterized protein Clorf93 homolog
 gb|AAH78028.1| **G** MGC82733 protein [Xenopus laevis]
 Length=201

GENE ID: 447017 clorf93 | chromosome 1 open reading frame 93 [Xenopus laevis]
 (10 or fewer PubMed links)

Score = 63.5 bits (153), Expect = 7e-09, Method: Compositional matrix adjust.
 Identities = 41/139 (30%), Positives = 66/139 (48%), Gaps = 11/139 (7%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
 K+Q +L+ LRRFGC +C A + ++K D +++V +G E +
 Sbjct 30 KEQTTVLLFLRRFGCQICRWIAKDMGKLESCDVHQIRLVGIGPEEVGLKEFL-----D 83
 Query 61 GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR----KANKNHPNADLQGDGL 116
 G F E+YID + +YK G +R L + + R KAN + + GD L
 Sbjct 84 GNFFNGELYIDDSKQSYKDLGFKRYSALSVIPAALGKKVRDIVTKANADGVQGNFSGDLL 143
 Query 117 QTGGIYLVGPGADSA-IHF 134
 Q+GG+ +V G + +HF
 Sbjct 144 QSGGMLIVSKGGEKVLH 162

>gb|EGC30304.1| hypothetical protein DICPUDRAFT_99572 [Dictyostelium purpureum]
 Length=808

Score = 63.2 bits (152), Expect = 1e-08, Method: Composition-based stats.
 Identities = 37/130 (29%), Positives = 69/130 (54%), Gaps = 10/130 (7%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
 +++R+++ + RRFGC +C QA + +KP+LD G+++V +G E FI+
 Sbjct 410 ENKRIVVAIFRRFGCLICRLQALDLSLKPCLDRMGIELVIGIFDEEGIDE-FIQY---- 464
 Query 61 GQRFPAEVYIDPEQTAYKARGLQRVGLLH----FLSWTAISEWRKANKNHPNADLQGDGL 116
 + F ++YID + Y+A L+R L FL ++ +R+ + ++ + DG
 Sbjct 465 -KFFAGKIYIDKNRQVYRALNLKRRSKLTTYELFLDPRVMTYRRMKELGLPSNYRKDGF 523
 Query 117 QTGGIYLVGP 126
 Q G ++GP
 Sbjct 524 QLGATLVLGP 533

>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 = 62.8 bits (151), Expect = 1e-08, Method: Compositional matrix adjust.
 Identities = 44/154 (29%), Positives = 73/154 (48%), Gaps = 20/154 (12%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPG 59
 +++R+++ +R F C +C E + +I K L A V ++++G + + E F +
 Sbjct 56 RERRAIVVFRHFLCYICKEYVEDLAKIPKSFLQEANVTLIVIGQSSYHHIEPFCKLT-- 113
 Query 60 NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
 + E+Y+DPE+ YK G++R + + LS + S WR P
 Sbjct 114 ---GYSHEIYVDPEREIYKRLGMKRGEEIASSGQSPHVKSNIILSGSIRSLWRAVTVG--PL 168
 Query 109 ADLQGDGLQTGGIYLVGPGAD-SAIHFANNEYDH 141
 D QGD Q GG ++GPG + IH N DH
 Sbjct 169 FDFQGDPAQQGGTLILGPGNNIHFHHRDRNRLDH 202

>**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 = 62.8 bits (151), Expect = 2e-08, Method: Compositional matrix adjust.
Identities = 43/143 (31%), Positives = 71/143 (50%), Gaps = 10/143 (6%)

```
Query 1   KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPV 59
          +++R +++ +R F C +C E   + +I + L A V ++++G + + E F +
Sbjct 54   RERRAVVVVVRHFLCYICKEYVEDLAKIPRSFLQEANVTLIVIGQSSYHHIEPFCKLT-- 111

Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWRKANKNHPNADLQGDGLQTG 119
          + E+Y+DPE+ YK G++R G   S + S WR   P D QGD Q G
Sbjct 112  ---GYSHEIYVDPEREIYKRLGMKR-GEEIASSGSLQSLWRAVTG--PLDFDQGDPAQQG 165

Query 120 GIYLVGPGAD-SAIHFANNEYDH 141
          G ++GPG + IH N DH
Sbjct 166  GTLILGPGNNIHFHRDRNRLDH 188
```

>**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 = 62.8 bits (151), Expect = 2e-08, Method: Compositional matrix adjust.
Identities = 45/154 (30%), Positives = 75/154 (49%), Gaps = 20/154 (12%)

```
Query 1   KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPV 59
          +++R +++ +R F C +C E   + +I K L A V ++++G + + E F + + G
Sbjct 20   RERRAVVVVVRHFLCYICKEYVEDLAKIPKSFLQEADVTLIVIGQSSYHHIEPFCK-LTG 78

Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
          + E+Y+DPE+ YK G++R +   + LS + S WR   P
Sbjct 79   ----YSHEIYVDPEREIYKKGMRGEEIASSGKSPHIKSNILSGSIRSLWRAVTG--PL 132

Query 109 ADLQGDGLQTGGIYLVGPGAD-SAIHFANNEYDH 141
          D QGD Q GG ++GPG + IH N DH
Sbjct 133  FDFQGDPAQQGGTLILGPGNNIHFHRDRNRLDH 166
```

>**gb|EFB20490.1** | hypothetical protein PANDA_018799 [Ailuropoda melanoleuca]
Length=186

Score = 62.8 bits (151), Expect = 2e-08, Method: Compositional matrix adjust.
Identities = 45/154 (30%), Positives = 75/154 (49%), Gaps = 20/154 (12%)

```
Query 1   KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPV 59
          +++R +++ +R F C +C E   + +I K L A V ++++G + + E F + + G
Sbjct 14   RERRAVVVVVRHFLCYICKEYVEDLAKIPKSFLQEADVTLIVIGQSSYHHIEPFCK-LTG 72

Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
          + E+Y+DPE+ YK G++R +   + LS + S WR   P
Sbjct 73   ----YSHEIYVDPEREIYKKGMRGEEIASSGKSPHIKSNILSGSIRSLWRAVTG--PL 126

Query 109 ADLQGDGLQTGGIYLVGPGAD-SAIHFANNEYDH 141
          D QGD Q GG ++GPG + IH N DH
Sbjct 127  FDFQGDPAQQGGTLILGPGNNIHFHRDRNRLDH 160
```

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

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

Score = 62.4 bits (150), Expect = 2e-08, Method: Compositional matrix adjust.
Identities = 45/156 (29%), Positives = 74/156 (48%), Gaps = 24/156 (15%)

```
Query 1   KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPV 59
```

```

Sbjct 54  +++R +++ +R F C +C E   + +I K L A V ++++G + + E F +
RERRAVVVVVRHFLCYVCKEYVEDLAKIPKSVLQEADVTLVIGQSSYHHIEPFCKLT-- 111
Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
      + E+Y+DPE+ YK G++R + + LS + S WR P
Sbjct 112  ---GYSHEIYVDPEREIYKRLGMKRGEIISSSGQSPHIKSNLLSGLSLWRAVGTG--PL 166
Query 109 ADLQGDGLQTTGGIYLVGPGADSAIHFAF---NEYDH 141
      D QGD Q GG ++GPG + IHF N DH
Sbjct 167  FDFQGDPAQQGGTLILGPGNN--IHFVHRDRNRLDH 200

```

>ref|XP_003130904.1| **UGM** PREDICTED: UPF0308 protein C9orf21 homolog isoform 1 [Sus
Length=228

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

Score = 62.4 bits (150), Expect = 2e-08, Method: Compositional matrix adjust.
Identities = 45/156 (29%), Positives = 74/156 (48%), Gaps = 24/156 (15%)

```

Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVP 59
      +++R +++ +R F C +C E   + +I K L A V ++++G + + E F +
Sbjct 56  RERRAVVVVVRHFLCYICKEYVEDLAKIPKSVLQEANVTLVIGQSSYHHIEPFCKLT-- 113
Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
      + E+Y+DPE+ YK G++R + + LS + S WR P
Sbjct 114  ---GYSHEIYVDPEREIYKRLGMKRGEIASSGKSPHIKSNILSGSIRSLWRAVGTG--PL 168
Query 109 ADLQGDGLQTTGGIYLVGPGADSAIHFAF---NEYDH 141
      D QGD Q GG ++GPG + IHF N DH
Sbjct 169  FDFQGDPAQQGGTVILGPGNN--IHFVHRDRNRLDH 202

```

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

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

Score = 62.4 bits (150), Expect = 2e-08, Method: Compositional matrix adjust.
Identities = 46/156 (30%), Positives = 76/156 (49%), Gaps = 24/156 (15%)

```

Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVP 59
      +++R +++ +R F C +C E   + +I K L A V ++++G + + E F + + G
Sbjct 14  RERRAVVVVVRHFLCYVCKEYVEDLAKIPKSVLREADVTLVIGQSSYHHIEPFCK-LTG 72
Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
      + E+Y+DPE+ YK G++R + + LS + S WR P
Sbjct 73  ---YSHEIYVDPEREIYKRLGMKRGEIISSSGQSPHIKSNLLSGLSLWRAVGTG--PL 126
Query 109 ADLQGDGLQTTGGIYLVGPGADSAIHFAF---NEYDH 141
      D QGD Q GG ++GPG + IHF N DH
Sbjct 127  FDFQGDPAQQGGTLILGPGNN--IHFVHRDRNRLDH 160

```

>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 = 62.4 bits (150), Expect = 2e-08, Method: Compositional matrix adjust.
Identities = 45/156 (29%), Positives = 74/156 (48%), Gaps = 24/156 (15%)

```

Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVP 59
      +++R +++ +R F C +C E   + +I K L A V ++++G + + E F +
Sbjct 54  RERRAVVVVVRHFLCYICKEYVEDLAKIPKSVLQEANVTLVIGQSSYHHIEPFCKLT-- 111
Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
      + E+Y+DPE+ YK G++R + + LS + S WR P
Sbjct 112  ---GYSHEIYVDPEREIYKRLGMKRGEIASSGQSPHVKNLLSGLSLWRAVGTG--PL 166
Query 109 ADLQGDGLQTTGGIYLVGPGADSAIHFAF---NEYDH 141
      D QGD Q GG ++GPG + IHF N DH
Sbjct 167  FDFQGDPAQQGGTLILGPGNN--IHFVHRDRNRLDH 200

```

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

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

Score = 62.4 bits (150), Expect = 2e-08, Method: Compositional matrix adjust.
Identities = 46/156 (30%), Positives = 76/156 (49%), Gaps = 24/156 (15%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPG 59
      +++R +++ +R F C +C E + +I K L A V ++++G + + E F + + G
Sbjct 22  RERRAVVVVFRHFLCYVCKEYVEDLAKIPKSVLREADVTLIVIGQSSYHHIEPFCK-LTG 80

Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
      + E+Y+DPE+ YK G++R + + LS + S WR P
Sbjct 81  ----YSHEIYVDPEREIYKRLGMKRGEEISSSGQSPHIKSNLLSGSLQSLWRAVGT--PL 134

Query 109  ADLQGDGLQTTGGIYLVGPGADSAIHFAF---NEYDH 141
      D QGD Q GG ++GPG + IHF N DH
Sbjct 135  FDFQGDPAQQGGTLILGPGNN--IHFVHRDRNRLDH 168
```

>ref|XP_520707.2| **UGM** PREDICTED: similar to TPA_exp: C90RF21 isoform 2 [Pan trogl]
Length=226

GENE ID: 465250 LOC465250 | similar to TPA_exp: C90RF21 [Pan troglodytes]

Score = 62.4 bits (150), Expect = 2e-08, Method: Compositional matrix adjust.
Identities = 45/156 (29%), Positives = 74/156 (48%), Gaps = 24/156 (15%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPG 59
      +++R +++ +R F C +C E + +I K L A V ++++G + + E F +
Sbjct 54  RERRAVVVVFRHFLCYICKKEYVEDLAKIPKSFLEQANVTLIVIGQSSYHHIEPFCKLT-- 111

Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
      + E+Y+DPE+ YK G++R + + LS + S WR P
Sbjct 112  ---GYSHEIYVDPEREIYKRLGMKRGEEIASSGQSPHVKSNNLLSGSLQSLWRAVGT--PL 166

Query 109  ADLQGDGLQTTGGIYLVGPGADSAIHFAF---NEYDH 141
      D QGD Q GG ++GPG + IHF N DH
Sbjct 167  FDFQGDPAQQGGTLILGPGNN--IHFVHRDRNRLDH 200
```

>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 = 62.4 bits (150), Expect = 2e-08, Method: Compositional matrix adjust.
Identities = 45/156 (29%), Positives = 74/156 (48%), Gaps = 24/156 (15%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPG 59
      +++R +++ +R F C +C E + +I K L A V ++++G + + E F +
Sbjct 54  RERRAVVVVFRHFLCYVCKEYVEDLAKIPKSVLREADVTLIVIGQSSYHHIEPFCKLT-- 111

Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
      + E+Y+DPE+ YK G++R + + LS + S WR P
Sbjct 112  ---GYSHEIYVDPEREIYKRLGMKRGEEISSSGQSPHIKSNLLSGSLQSLWRAVGT--PL 166

Query 109  ADLQGDGLQTTGGIYLVGPGADSAIHFAF---NEYDH 141
      D QGD Q GG ++GPG + IHF N DH
Sbjct 167  FDFQGDPAQQGGTLILGPGNN--IHFVHRDRNRLDH 200
```

>ref|XP_002965388.1| **UG** hypothetical protein SELMODRAFT_68006 [Selaginella moeller
 gb|EFJ34226.1| **G** hypothetical protein SELMODRAFT_68006 [Selaginella moellendorffii]
 Length=172

GENE ID: 9633660 SELMODRAFT_68006 | hypothetical protein
 [Selaginella moellendorffii]

Score = 62.0 bits (149), Expect = 2e-08, Method: Compositional matrix adjust.
 Identities = 44/147 (30%), Positives = 63/147 (43%), Gaps = 23/147 (15%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          KD++ ++ R FGC LC ++A + K DAAGV +VLVG G A+ F
Sbjct 16  KDRtavvafarhfgcILCRKradVLASKKEVfDAAGVSLVLVPGTVDQAKAFASQT--- 72

Query 61  GQRFPAEVYIDPEQTAYKA-----RGLQRVGLLHFLSWTAISEWRKANKNHP 107
          +FP EVY DP ++ A + RV H + +W +
Sbjct 73  --QFPGEVYADPTHASFDaFQFVSGASTIFNPKAAMRVMGAHLEGYR--QDW---GLSFE 125

Query 108  NADLQDGLQTTGGIYLVGPGADSAIHF 134
          +Q G Q GGI + GPG D ++
Sbjct 126  KDTVQRGGWQGGIIVIAGPGKDRLLYI 152
```

>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 = 61.6 bits (148), Expect = 3e-08, Method: Compositional matrix adjust.
 Identities = 47/153 (31%), Positives = 70/153 (46%), Gaps = 17/153 (11%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          KD++ ++ R FGC C ++A + K ++DA+GV +VL+G G+ A+ F E
Sbjct 97  KDRKAVVAFARHFGCVFCRKRADLLASQKDRMDASGVALVLIGPGSIDQAKAFSEQT--- 153

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTA----ISEWRKANKNHPNADLQGD-- 114
          F EVY DP ++Y+ G G+L + A I + + + Q D
Sbjct 154  --NFKGEVYADPSHSSYEVLGFVS-GVLSTFTPQAGLKIIQLYMEGYRQDWGLSFQRDTV 210

Query 115  ---GLQTTGGIYLVGPGAD--SAIHFAFNEYDHP 142
          G Q GGI + GPG S IH D P
Sbjct 211  TRGGWQGGIIVAGPGKSNISYIHKDEAGDDP 243
```

>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 = 61.2 bits (147), Expect = 4e-08, Method: Compositional matrix adjust.
 Identities = 30/79 (38%), Positives = 46/79 (59%), Gaps = 5/79 (6%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          KD++ ++ R FGC LC +A ++ K +DA+GV +VL+G G+ A+ F E
Sbjct 41  KDRKAVVAFARHFGCVLCRRADYLAAKKDIMDASGVALVLIGPGSVDQAKTFSEQT--- 97

Query 61  GQRFPAEVYIDPEQTAYKA 79
          +F EVY DP ++YKA
Sbjct 98  --KFKGEVYADPSHSSYKA 114
```

>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: C9ORF21 [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 = 61.2 bits (147), Expect = 4e-08, Method: Compositional matrix adjust.
Identities = 44/156 (29%), Positives = 74/156 (48%), Gaps = 24/156 (15%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVP 59
      +++R +++ +R F C +C E + +I + L A V ++++G + + E F +
Sbjct 54  RERRAVVVVVRHFLCYICKEYVEDLAKIPRSFLQEANVTLIVIGQSSYHHIEPFCKLT-- 111

Query 60  NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
      + E+Y+DPE+ YK G++R + + LS + S WR P
Sbjct 112  ---GYSHEIYVDPEREIYKRLGMKRGEIASSGQSPHIKSNLLSGLSLWRAVVG--PL 166

Query 109  ADLQGDGLQDTGGIYLVGPGADSATHFHAF---NEYDH 141
      D QGD Q GG ++GPG + IHF N DH
Sbjct 167  FDFQGDPAQQGGTLILGPGNN--IHFIHRDRNRLDH 200
```

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

Score = 61.2 bits (147), Expect = 4e-08, Method: Compositional matrix adjust.
Identities = 44/156 (29%), Positives = 73/156 (47%), Gaps = 23/156 (14%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQ--LDAAGVKIVLVGTGNRYFAEKFIENVP 58
      + + ++I +R F C C E + +I PQ L A V ++++G E F ++
Sbjct 50  QTHKAIIFVRHFLCFTCEYVEDLSQI-PQEILLDADVRLIVIGQSGFISHIEAFC-SLT 107

Query 59  GNGQRFPAEVYIDPEQTAYKARGLQRVGLLH-----FLSWTAISEWRKANKNH 106
      G + E+Y+DPE+ Y+ G++R + L + S WR
Sbjct 108  G----YQHEIYVDPERHIYEKLGMRGEIYEETASQSPHVKSSMLVGSIKSMWRAMTS-- 161

Query 107  PNADLQGDGLQDTGGIYLVGPGADSAI-HFAFNEYDH 141
      P D QGD LQ GG ++GPG + + HF N ++H
Sbjct 162  PAFDFQGDPLQGGGALIIGPGPNIHVAHFDMNRFNH 197
```

>**ref|XP_002977235.1** | **UG** hypothetical protein SELMODRAFT_58056 [Selaginella moeller
gb|EFJ21844.1 | **G** hypothetical protein SELMODRAFT_58056 [Selaginella moellendorffii]
Length=172

GENE ID: 9638970 SELMODRAFT_58056 | hypothetical protein
[Selaginella moellendorffii]

Score = 60.8 bits (146), Expect = 5e-08, Method: Compositional matrix adjust.
Identities = 43/146 (30%), Positives = 63/146 (44%), Gaps = 23/146 (15%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVP 60
      KD+ ++ R FGC LC ++A + K D AGV +VLVG G A+ F
Sbjct 16  KDRTAVVAFARHFGCILCRKRADVLASKKEVFDGAGVSLVLVGPVTDQAKAFASQT--- 72

Query 61  GQRFPAEVYIDPEQTAYKA-----RGLQRVGLLHFLSWTAISEWRKANKNHP 107
      +FP EVY DP +++A + RV H + +W +
Sbjct 73  --QFPGEVYADPTHASFEAFQFVSGASTIFNPKAAMRVMGHLEGYR--QDW---GLSFE 125

Query 108  NADLQGDGLQDTGGIYLVGPGADSATH 133
      +Q G Q GGI + GPG D ++
Sbjct 126  KDTVQRGGWQQGGIVIAGPGKDRLLY 151
```

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

Score = 60.8 bits (146), Expect = 5e-08, Method: Compositional matrix adjust.
Identities = 40/139 (29%), Positives = 69/139 (50%), Gaps = 15/139 (10%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVP 60
      KD++ ++ R FGC LC ++A ++ K +DA+GV +VL+G G+ A+ F E
Sbjct 99  KDRKAVVAFARHFGCVLCRKRADYLSSKKDMDASGVALVLIGPGSIDQAKSFAEK---- 154

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTA----ISEWRKANKNHPNADLQGD-- 114
      +F E+Y DP ++Y+A G+L + A I + + + + D
```


Sbjct 155 -SKFEGEIYADPTHSSYEALNFVS-GVLTTFTPNAGLKIIQLYMEGYRQDWKLSFEKDTV 212
 Query 115 ---GLQTGGIYLVGPGADS 130
 G + GGI + GPG ++
 Sbjct 213 SRGGWKQGGIIVAGPGKNN 231

>ref|XP_002320012.1| **UG** predicted protein [Populus trichocarpa]

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

GENE ID: 7497309 P0PTRDRAFT_571999 | hypothetical protein [Populus trichocarpa]
 (10 or fewer PubMed links)

Score = 60.1 bits (144), Expect = 9e-08, Method: Compositional matrix adjust.
 Identities = 44/128 (35%), Positives = 64/128 (50%), Gaps = 15/128 (11%)

Query 6 LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRFP 65
 ++ LLR FGC C E AS + E K + D++GVK++ +G G A E +P FP
 Sbjct 43 VVALLRHFGCPCCWELASSLKESKEKFDSSGVKLIAGVGTGNRYFAEKFIENVPNGQRFP 97
 Query 66 AE-VYIDPEQTAYKARGLQR-VGLLHFLSWTA-----ISEWRKANKNHP---NADLQGDG 115
 + +Y DPE+ AY GL +G F +A RKA KN+ D +
 Sbjct 98 MDCLYADPERKAYDVLGLYYGLGRTFFNPASAKVFSRFDALRKAVKNYTIATPDDRSKV 157
 Query 116 LQTGGIYL 123
 LQ GG+++
 Sbjct 158 LQQGGMFV 165

>ref|XP_795970.1| **GM** PREDICTED: hypothetical protein [Strongylocentrotus purpuratus]

ref|XP_001176073.1| **UGM** PREDICTED: hypothetical protein [Strongylocentrotus purpuratus]
 Length=191

GENE ID: 591310 LOC591310 | hypothetical LOC591310
 [Strongylocentrotus purpuratus]

Score = 60.1 bits (144), Expect = 1e-07, Method: Compositional matrix adjust.
 Identities = 38/130 (30%), Positives = 61/130 (47%), Gaps = 10/130 (7%)

Query 9 LLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRFP 68
 LRRFGC +C A + +KP+LDAA V++V +G A++FIE+ G ++
 Sbjct 36 FLRRFGCPICRMGARDITHLKPRLDAANVRLVAIGQ-EETGAKEFIESGFWTG-----DL 89
 Query 69 YIDPEQTAYKARGLQRVGLLHFLSWTAISEWR----KANKNHPNADLQDGLQGGIYLV 124
 +ID ++ Y +R L ++ R KA ++ GD LQ GG ++
 Sbjct 90 FIDQQKKTGYDLKYKRYNFLTIMANLMCKMTREAVSKATSEGITGNMTGDALQMGGTLVI 149
 Query 125 GPGADSAIHF 134
 G + F
 Sbjct 150 DKGGKVLDF 159

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

Score = 59.7 bits (143), Expect = 1e-07, Method: Compositional matrix adjust.
 Identities = 44/133 (34%), Positives = 68/133 (52%), Gaps = 15/133 (11%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNG 60
 K+ ++ LLR FGC C E AS + ++ P+ D+AGVK++ +G G A E +P
 Sbjct 115 KNGTAVVALLRHFGCPCCWFASTLKDVMKPKFDSAGVKLIAIGVGTPEKARILGERLP-- 172
 Query 61 GQRFPAE-VYIDPEQTAYKARGLQR-VGLLHFLSWTA-----ISEWRKANKNHPNADLQ 113
 FP + +Y DP++ AY A GL +G F +A +KA KN+ +
 Sbjct 173 ---FPLDSLYADPDRKAYDALGLYYGLGRTFFNPASAKVLRFDLSLQKALKNYTISATPE 229
 Query 114 DG---LQTGGIYL 123
 D LQ GG+++
 Sbjct 230 DRSSVLQGGGMFV 242

>ref|XP_001926014.1| **UGM** PREDICTED: UPF0765 protein C10orf58 isoform 4 [Sus scrofa]
 Length=231

GENE ID: 100155717 LOC100155717 | SFLQ611 [Sus scrofa]
(10 or fewer PubMed links)

Score = 59.7 bits (143), Expect = 1e-07, Method: Compositional matrix adjust.
Identities = 37/125 (30%), Positives = 65/125 (52%), Gaps = 13/125 (10%)

```
Query 6  LLILLRRFGCSLCHQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGNGQRF 65
      +++ +RR GC LC E+A+ + +KP+LD GV + V E+ V F
Sbjct 78  VIMAVRRPGCFLCREEAADLSSLKPRDELGVPLYAV-----VKEQVKNEVKDFQPYFK 131

Query 66  AEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR---KANKNHPNADLQGDGLQTGGIY 122
      E+++D E+ Y G QR ++ F+ + + W +A + +L+G+G GG++
Sbjct 132  GEIFLDEEKIFY---GPQRRKMM-FMGFVRLGVWYNFFRARSGGFSGNLEGEFVLGGVF 187

Query 123  LVGPG 127
      +VGPG
Sbjct 188  VVGPG 192
```

>ref|XP_001370596.1| **UGM** PREDICTED: similar to SFLQ611 [Monodelphis domestica]
Length=229

GENE ID: 100016854 LOC100016854 | similar to SFLQ611 [Monodelphis domestica]

Score = 59.3 bits (142), Expect = 2e-07, Method: Compositional matrix adjust.
Identities = 36/125 (29%), Positives = 65/125 (52%), Gaps = 13/125 (10%)

```
Query 6  LLILLRRFGCSLCHQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGNGQRF 65
      +++ +RR GC LC E+A+ + +KPQLD GV + V EK V F
Sbjct 76  VIMAVRRPGCFLCREEAADLSALKPQLDLLGVPLYAV-----VKEKIGSEVENFQPYFK 129

Query 66  AEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR---KANKNHPNADLQGDGLQTGGIY 122
      ++++D + Y G QR ++ F+ + + W +A + +L+G+G GG+Y
Sbjct 130  GKIFLDERKKFY---GPQRRKMM-FMGFVRLGVWQNFRRARSKGFSGNLEGEFVLGGVY 185

Query 123  LVGPG 127
      ++GPG
Sbjct 186  VIGPG 190
```

>ref|XP_001368977.1| **UGM** PREDICTED: similar to C90RF21 [Monodelphis domestica]
Length=354

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

Score = 58.9 bits (141), Expect = 2e-07, Method: Compositional matrix adjust.
Identities = 43/154 (28%), Positives = 71/154 (47%), Gaps = 20/154 (12%)

```
Query 1  KDQRVLLILLRRFGCSLCHQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPG 59
      +++R +++ +R F C C E + +I K L A V +++G + E F +
Sbjct 182  RERRAIVVFRHFLCYTCKEYVEDLAKIPKSFLQDANVTLIVIGQSSFQHIPEFCKLT-- 239

Query 60  NGQRFPAEVYIDPEQTAYKARGLQV-VGLL-----HFLSWTAISEWRKANKNHPN 108
      R+ E+Y+D E+ Y+ G+ + G+ + LS + S WR P
Sbjct 240  ---RYSHEIYVDTERKIYRKLGMNKGEGIASSEQSPHVKSNNLLSGSIQSLWRAVTVG--PA 294

Query 109  ADLQGDGLQTGGIYLVGPGAD-SAIHFANEYDH 141
      D QGD Q GG ++GPG + IH N DH
Sbjct 295  FDFQGDPAQQGGTLILGPGNNIHFHLDKNRLDH 328
```

>gb|AAP06083.1| similar to GenBank Accession Number AK005188 putative related
to F3G5 [Schistosoma japonicum]
Length=162

Score = 58.9 bits (141), Expect = 2e-07, Method: Compositional matrix adjust.
Identities = 37/130 (29%), Positives = 64/130 (50%), Gaps = 10/130 (7%)

```
Query 13  FGCSLCHQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGNGQRFPAEVYIDP 72
      GC C +A ++ +KP LD +K++ + T + ++F+ +G F ++Y+DP
Sbjct 1  MGCKFCRLEAKNLSYLKPALDTRNIKLIGI-TFDVGGVKEFL-----DGHYFDGDLYLDP 54

Query 73  EQTAYKARGLQRV---GLLHFLSWTAISEWRKANKNHPNADLQGDGLQTGGIYLVGPGA 128
      E+ YKA G ++V G++ S + KA +L GDG QTGG+ +V G
```

Sbjct 55 ERMTYKALGYKKVSPCSGVISLFSKAGRALNSKAKAAKIPGNLSGDGWQTGGLLVVEKGG 114
 Query 129 DSAIHFAFNE 138
 ++ E
 Sbjct 115 KILYYEYQKE 124

>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 SORBITRAFT_01g009350 | hypothetical protein [Sorghum bicolor]
 (10 or fewer PubMed links)

Score = 58.2 bits (139), Expect = 4e-07, Method: Compositional matrix adjust.
 Identities = 44/128 (35%), Positives = 64/128 (50%), Gaps = 15/128 (11%)

Query 6 LLILLRRFGCSLCHQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRFP 65
 ++ LLR FGC C E AS + + K + D+AGVK++ VG G A E +P FP
 Sbjct 104 VVALLRHFGCPCCWELASVLRDTKEKFDASAGVKLIAVGVGTPAKARILAERLP-----FP 158
 Query 66 AE-VYIDPEQTAYKARGLQ-RVGLLHFLSWTA-----ISEWRKANKNH---PNADLQGDG 115
 E +Y DP++ AY GL VG F +A ++A KN+ D +
 Sbjct 159 LEYLYADPDRKAYNLLGLYFVGVRTFFNPASAKVFSRFDLKEAVKNYTMETPDDRAGV 218
 Query 116 LQTGGIYL 123
 LQ GG+++
 Sbjct 219 LQQGGMFV 226

>ref|YP_001529200.1| **G** hypothetical protein Dole_1319 [Desulfococcus oleovorans Hxc
 gb|ABW67123.1| **G** hypothetical protein Dole_1319 [Desulfococcus oleovorans Hxd3]
 Length=144

GENE ID: 5694153 Dole_1319 | hypothetical protein
 [Desulfococcus oleovorans Hxd3]

Score = 58.2 bits (139), Expect = 4e-07, Method: Compositional matrix adjust.
 Identities = 38/112 (34%), Positives = 56/112 (50%), Gaps = 7/112 (6%)

Query 17 LCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRFPFAEVYIDPEQTA 76
 +QA+ ++ +K QLD GV +V VG+G A FIE +F EVY+DP A
 Sbjct 2 FARQQAADLMNVKKQLDEMIVALVAVGSGTPEQARVFIK-----KFEGEVYLDPSLAA 56
 Query 77 YKARGLQRVGLLHFLSWTAISE-WRKANKNHPNADLQGDGLQTGGIYLVGPG 127
 Y A L+R G L +I ++ + GD Q GG++++GPG
 Sbjct 57 YNAFRLKR-GFWRTLGPCSIGRGFKTMFRGFHQGRKAGDLWQQGGLFVMGPG 107

>ref|XP_546736.2| **UGM** PREDICTED: hypothetical protein XP_546736 [Canis familiaris
 Length=201

GENE ID: 489616 LOC489616 | hypothetical LOC489616 [Canis lupus familiaris]

Score = 58.2 bits (139), Expect = 4e-07, Method: Compositional matrix adjust.
 Identities = 45/143 (32%), Positives = 68/143 (48%), Gaps = 13/143 (9%)

Query 2 DQRVLLILLRRFGCSLCHQASHVLEIKPQLDAAGVKIVLVGTGNRYFA-EKFIENVPGN 60
 +Q ++ LRRFGCS+C A + ++ LD GV+ LVG G ++F+ +
 Sbjct 31 EQACVVAGLRRFGCSVCRWIARDLSSLRGLLDQHGVR--LVGVGPEVLGVQEF-----D 83
 Query 61 GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR---KANKNHPNADLQGDGL 116
 G F E+Y+D + Y+ G +R L L R KA + +L GD L
 Sbjct 84 GGYFAGELYLDESKQFYRELGFKRYNSLSILPAALGKPVDRDVALKAKAVGIHGNLSGDLL 143
 Query 117 QTGGIYLVGPGADSA-IHFAN 138
 Q+GG+ +V G D +HF N
 Sbjct 144 QSGGLLVVTKGGDKVLLHFVQNS 166

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

GENE ID: 100384338 pco101707a | LOC100384338 [Zea mays]

Score = 57.8 bits (138), Expect = 4e-07, Method: Compositional matrix adjust.
Identities = 44/128 (35%), Positives = 64/128 (50%), Gaps = 15/128 (11%)

```
Query 6  LLILLRRFGCSLCHQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRF 65
      ++ LLR FGC C E AS + + K + D+AGVK++ VG G A E +P FP
Sbjct 105 VVALLRHFGCPCCWELASVLRDTERFDSAGVKLIAVGVGTPAKARILAERLP-----FP 159

Query 66  AE-VYIDPEQTAYKARGLQ-RVGLLHFLSWTA-----ISEWRKANKNH---PNADLQGDG 115
      E +Y DP++ AY GL VG F +A ++A KN+ D +
Sbjct 160 LEYLYADPDRKAYNLLGLYFVGVRTFFNPASAKVFSRFDLKEAVKNYTIATPDDRAGV 219

Query 116  LQTGGIYL 123
      LQ GG+++
Sbjct 220 LQQGGMFV 227
```

>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 = 57.8 bits (138), Expect = 5e-07, Method: Compositional matrix adjust.
Identities = 44/128 (35%), Positives = 64/128 (50%), Gaps = 15/128 (11%)

```
Query 6  LLILLRRFGCSLCHQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRF 65
      ++ LLR FGC C E AS + + K + D+AGVK++ VG G A E +P FP
Sbjct 6  VVALLRHFGCPCCWELASVLRDTERFDSAGVKLIAVGVGTPAKARILAERLP-----FP 60

Query 66  AE-VYIDPEQTAYKARGLQ-RVGLLHFLSWTA-----ISEWRKANKNHP---NADLQGDG 115
      E +Y DP++ AY GL VG F +A ++A KN+ D +
Sbjct 61 LEYLYADPDRKAYNLLGLYFVGVRTFFNPASAKVFSRFDLKEAVKNYTIATPDDRAGV 120

Query 116  LQTGGIYL 123
      LQ GG+++
Sbjct 121 LQQGGMFV 128
```

>emb|CAN81555.1| hypothetical protein VITISV_040397 [Vitis vinifera]
Length=201

Score = 57.8 bits (138), Expect = 5e-07, Method: Compositional matrix adjust.
Identities = 46/134 (35%), Positives = 64/134 (48%), Gaps = 17/134 (12%)

```
Query 1  KDQRVLLILLRRFGCSLCHQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNG 60
      K+ ++ LLR FGC C E AS + E K + D+AGVK++ VG G A E +P
Sbjct 40 KEGVAVVALLRHFGCFCCWELASALKESKARFDSAGVKLIAVGVGTPNKACILAERLP-- 97

Query 61  GQRFPAE-VYIDPEQTAYKARGLQRVGLLHFL---SWTAISEWRKANKNHPNADLQGDG 115
      FP + +Y DP++ AY GL GL L S S + K N L+G
Sbjct 98 ---FPMDCLYADPDRKAYDVLGLY-YGLSRTLFSASAKVFSRFESLQKALKNYTLEGTP 153

Query 116  -----LQTGGIYL 123
      LQ GG+++
Sbjct 154 DDKSGVLQGGGMFV 167
```

>ref|XP_536403.1| **UGM** PREDICTED: similar to R53.5 [Canis familiaris]
Length=225

GENE ID: 479261 LOC479261 | similar to R53.5 [Canis lupus familiaris]

Score = 57.8 bits (138), Expect = 5e-07, Method: Compositional matrix adjust.
Identities = 36/125 (29%), Positives = 65/125 (52%), Gaps = 13/125 (10%)

```
Query 6  LLILLRRFGCSLCHQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRF 65
      +++ +RR GC LC E+A+ + +KP+LD GV + V E+ V F
Sbjct 72 VIMAVRRPGCFRCREEAADLSSLKPKLDELGVPLYAV-----VKEQIRTEVQDFQPYFK 125

Query 66  AEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR---KANKNHPNADLQGDGLQTGGIY 122
      E+++D ++ Y G QR ++ F+ + + W +A + +L+G+G GG++
Sbjct 126 GEIFLDEKFFY--GPQRRKMM-FMGFVRLGVWYNNFRARNGGFSGNLEGEFILGGVF 181
```

Query 123 LVGPG 127
 +VGPG
 Sbjct 182 VVGPG 186

>gb|ACU20324.1| unknown [Glycine max]
 Length=251

Score = 57.4 bits (137), Expect = 5e-07, Method: Compositional matrix adjust.
 Identities = 46/131 (36%), Positives = 68/131 (52%), Gaps = 21/131 (16%)

Query 6 LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRF 65
 ++ +LR FGC C E AS + E K + D+AG+K++ VG G A E +P FP
 Sbjct 95 VVAMLRHFGCICCFWELASALKESKARFDSAGIKLIAVGVGTPNKARILAERLP-----FP 149

Query 66 AE-VYIDPEQTAYKAR----GLQRVGLLHFLSWTAISEW---RKANKNH-----PNADLQ 112
 + +Y DP++ AY GL R L+ S S W +KA KN+ P+ D+
 Sbjct 150 MDCLYADPDRKAYNVLNLYFGLGRT-FLNPASAKVFSRWDALQKAAKNYITIGATPD-DIS 207

Query 113 GDGLQTGGIYL 123
 G LQ GG+++
 Sbjct 208 G-VLQQGGMFV 217

>ref|XP_002263959.1| **UGM** PREDICTED: hypothetical protein [Vitis vinifera]
 Length=255

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

Score = 57.4 bits (137), Expect = 5e-07, Method: Compositional matrix adjust.
 Identities = 46/134 (35%), Positives = 64/134 (48%), Gaps = 17/134 (12%)

Query 1 KDQVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNG 60
 K+ ++ LLR FGC C E AS + E K + D+AGVK++ VG G A E +P
 Sbjct 94 KEGVAVVALLRHFGCFCCWELASALKESKARFDSAGVKLIAVGVGTPNKACILAERLP-- 151

Query 61 GQRFPAE-VYIDPEQTAYKARGLQRVGLLHFL----SWTAISEWRKANKNHNPADLQGDG 115
 FP + +Y DP++ AY GL GL L S S + K N L+G
 Sbjct 152 ---FPMDCLYADPDRKAYDVLGLY-YGLSRTLFSASAKVFSRFESLQKALKNYTLEGTP 207

Query 116 -----LQTGGIYL 123
 LQ GG+++
 Sbjct 208 DDKSGVLQQGGMFV 221

>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| **G** 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 = 57.4 bits (137), Expect = 6e-07, Method: Compositional matrix adjust.
 Identities = 43/128 (34%), Positives = 64/128 (50%), Gaps = 15/128 (11%)

Query 6 LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRF 65
 ++ LLR FGC C E AS + + K + D+AGVK++ VG G A E +P FP
 Sbjct 102 VVALLRHFGCPCCWELASVLRDTERFDSAGVKLIAVGVGTPDKARILAERLP-----FP 156

Query 66 AE-VYIDPEQTAYKARGLQ-RVGLLHFLSWTA-----ISEWRKANKNH---PNADLQGDG 115
 + +Y DPE+ AY GL +G F +A ++A KN+ D +
 Sbjct 157 LDYLYADPERKAYDLLGLYFGIGRTFFNPASASVFSRFDLKEAVKNYTIETPDDRASV 216

Query 116 LQTGGIYL 123
 LQ GG+++
 Sbjct 217 LQQGGMFV 224

>sp|Q5ZI34.2|CJ058_CHICK  RecName: Full=UPF0765 protein C10orf58 homolog; Flags: F
Length=224


GENE ID: 423625 C10orf58 | chromosome 10 open reading frame 58 [Gallus gallus]
(10 or fewer PubMed links)


Score = 57.4 bits (137), Expect = 6e-07, Method: Compositional matrix adjust.
Identities = 37/127 (30%), Positives = 63/127 (50%), Gaps = 7/127 (5%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          K   +++ +RR G  LC E+AS +  +KPQL  GV +  V           EK   V
Sbjct 71  KKNGAVIMAVRRPGUFLCREEASELSSLPQLSKLGVPLYAV-----VKEKIGTEVEDF 124

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWRKANKNHPNADLQGDGLQTGG 120
          F  E+++D +++ Y  R  +++ L  F  + +A KN  + +L+G+G  GG
Sbjct 125  QHYFQGEIFLDEKRSFYGPRK-RKMMLSGFFRIGVWQNFRAWKNGYSGNLEGEFTLGG 183

Query 121  IYLVGPG 127
          +Y++G G
Sbjct 184  VYVIGAG 190
```

>ref|NP_001180447.1|  selenoprotein U [Gallus gallus]

ref|NP_001180448.1|  selenoprotein U [Gallus gallus]
Length=224


GENE ID: 423625 C10orf58 | chromosome 10 open reading frame 58 [Gallus gallus]
(10 or fewer PubMed links)


Score = 57.4 bits (137), Expect = 6e-07, Method: Compositional matrix adjust.
Identities = 37/127 (30%), Positives = 63/127 (50%), Gaps = 7/127 (5%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          K   +++ +RR G  LC E+AS +  +KPQL  GV +  V           EK   V
Sbjct 71  KKNGAVIMAVRRPGUFLCREEASELSSLPQLSKLGVPLYAV-----VKEKIGTEVEDF 124

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWRKANKNHPNADLQGDGLQTGG 120
          F  E+++D +++ Y  R  +++ L  F  + +A KN  + +L+G+G  GG
Sbjct 125  QHYFQGEIFLDEKRSFYGPRK-RKMMLSGFFRIGVWQNFRAWKNGYSGNLEGEFTLGG 183

Query 121  IYLVGPG 127
          +Y++G G
Sbjct 184  VYVIGAG 190
```

>gb|AAI14901.1|  Chromosome 1 open reading frame 93 ortholog [Bos taurus]

gb|DAA21137.1|  hypothetical protein LOC617001 [Bos taurus]
Length=201

GENE ID: 617001 C16H1orf93 | chromosome 1 open reading frame 93 ortholog
[Bos taurus] (10 or fewer PubMed links)

Score = 57.4 bits (137), Expect = 6e-07, Method: Compositional matrix adjust.
Identities = 44/139 (32%), Positives = 66/139 (48%), Gaps = 11/139 (7%)

```
Query 1  KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          ++Q  ++  LRRFGC +C  A  +  +K  LD  GV++V  VG  ++F+  +
Sbjct 30  QEQACVVAGLRRFGCMVCRWIARDLSNLKGLLDQHGVRVGVGP-EALGLQEFL-----D 83

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR----KANKNHPNADLQGDGL 116
          G  F  E+Y+D  +  YK  G  +R  L  L  R  KA  +L  GD  L
Sbjct 84  GGYFAGELYLDESKQFYKELGFKRYNSLSILPAALGKPVREVAAKAKAVGIQGNLSGDL 143

Query 117  QTGGIYLVGPGADSA-IHF 134
          Q+GG+ +V  G  D  +HF
Sbjct 144  QSGGLLVVAKGGDKVLLHF 162
```

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

Score = 57.4 bits (137), Expect = 6e-07, Method: Compositional matrix adjust.
Identities = 41/144 (29%), Positives = 68/144 (48%), Gaps = 17/144 (11%)


```

Query 1   KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          K+++ ++   R FGC LC ++A + + + AAGV +VL+G G+ A+ F E
Sbjct 102  KERKAVVAFARHFGCVLCRKRADLLAAKQDDMQAAGVALVLIGPGSVEQAKAFCEQT--- 158

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTA----ISEWRKANKN-----HPNAD 110
          +F  EVY DP  ++Y A      GL  + A      I  +R+ +      N
Sbjct 159  --KFKGEVYADPTHSSYDALEFA-FGLFSTFTPAAGLKIIQLYREGYRQDWELSFENR 215

Query 111  LQGDGLQTGGIYLVGPGADSAIHF 134
          +G G  GG+ + GPG D+ ++
Sbjct 216  TKG-GWYQGGLIVAGPGIDNIIYI 238

```

>ref|XP_848380.1| **UGM** PREDICTED: similar to UPF0308 protein C9orf21 [Canis famili
Length=230

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

Score = 57.4 bits (137), Expect = 6e-07, Method: Compositional matrix adjust.
Identities = 44/156 (29%), Positives = 74/156 (48%), Gaps = 24/156 (15%)

```

Query 1   KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPG 59
          +++R +++ +R F C +C E + +I K L A + +++G + + E F +
Sbjct 58  RERRAVVVFVRHFLCYICKEYVEDLAKIPKSVLQEADITLIVIGQSSYHHIEPFCKLT-- 115

Query 60  NGQRFPAEVYIDPEQTAYKARGLQR-VGLL-----HFLSWTAISEWRKANKNHPN 108
          + E+Y+DPE+ YK G++R G+ + LS + S R P
Sbjct 116  ---GYSHEIYVDPEREIYKKGMRGEGIASGKSPHIKSNILSGSIRS LCRAVGT--PL 170

Query 109 ADLQGDGLQTGGIYLVGPGADSAIHFAF--NEYDH 141
          D QGD Q GG ++GPG + IHF N DH
Sbjct 171  FDFQGDPAQQGGTLILGPGNN--IHFIHRDRNRLDH 204

```

>ref|NP_001035688.1| **UGM** hypothetical protein LOC617001 [Bos taurus]
sp|Q58CY6.1|CA093 BOVIN **G** RecName: Full=Uncharacterized protein Clorf93 homolog
gb|AA46658.1| **G** hypothetical protein MGC26818 [Bos taurus]
Length=201

GENE ID: 617001 C16H1orf93 | chromosome 1 open reading frame 93 ortholog
[Bos taurus] (10 or fewer PubMed links)

Score = 57.4 bits (137), Expect = 6e-07, Method: Compositional matrix adjust.
Identities = 44/139 (32%), Positives = 66/139 (48%), Gaps = 11/139 (7%)

```

Query 1   KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          ++Q ++ LRRFGC +C A + +K LD GV++V VG ++F+ +
Sbjct 30  QEQACVVAGLRRFGCMVCRWIARDLSNLKGLLDQHGVRVGVGP-EALGLQEF-----D 83

Query 61  GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR---KANKNHPNADLQGDGL 116
          G F E+Y+D + YK G +R L L R KA +L GD L
Sbjct 84  GGYFAGELYLDESKQFYKELGFKRYNSLSILPAALGKPVREVAAKAKAVGIQGNLSGDLL 143

Query 117  QTGGIYLVGPGADSA-IHF 134
          Q+GG+ +V G D +HF
Sbjct 144  QSGGLLVVAKGGDKVLLHF 162

```

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

Score = 57.0 bits (136), Expect = 7e-07, Method: Compositional matrix adjust.
Identities = 46/134 (35%), Positives = 63/134 (48%), Gaps = 17/134 (12%)

```

Query 1   KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          K+ ++ LLR FGC C E AS + E K D+AGVK++ VG G A E +P
Sbjct 94  KEGVAVVALLRHFGCFCCWELASALKESKATFDSAGVKLIAVGVGTPNKACILAERLP-- 151

Query 61  GQRFPAE-VYIDPEQTAYKARGLQRVGLLHFL---SWTAISEWRKANKNHPNADLQGDG 115
          FP + +Y DP++ AY GL GL L S S + K N L+G
Sbjct 152  ---FPMDCLYADPDRKAYDVLGLY-YGLSRTLFSASAKVFSRFESLQKALKNYTLEGTP 207

```


Query 116 -----LQTGGIYL 123
 LQ GG+++
 Sbjct 208 DDKSGVLQGGMFV 221

>ref|XP_002269002.1| **UGM** PREDICTED: hypothetical protein, partial [Vitis vinifera]
 Length=223

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

Score = 57.0 bits (136), Expect = 7e-07, Method: Compositional matrix adjust.
 Identities = 46/134 (35%), Positives = 63/134 (48%), Gaps = 17/134 (12%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
 K+ ++ LLR FGC C E AS + E K D+AGVK++ VG G A E +P
 Sbjct 62 KEGVAVVALLRHFGCFCCWELASALKESKATFDSAGVKLIAVGVGTPNKACILAERLP-- 119

Query 61 GQRFPAE-VYIDPEQTAYKARGLQRVGLLHFL----SWTAISEWRKANKNHPNADLQGDG 115
 FP + +Y DP++ AY GL GL L S S + K N L+G
 Sbjct 120 ---FPMDCLYADPDRKAYDVLGLY-YGLSRTLFSASAKVFSRFESLQKALKNYTLEGTP 175

Query 116 -----LQTGGIYL 123
 LQ GG+++
 Sbjct 176 DDKSGVLQGGMFV 189

>ref|XP_002924432.1| **GM** PREDICTED: uncharacterized protein Clorf93-like [Ailuropec
 Length=217

GENE ID: 100476330 LOC100476330 | uncharacterized protein Clorf93-like
 [Ailuropec melanoleuca]

Score = 57.0 bits (136), Expect = 7e-07, Method: Compositional matrix adjust.
 Identities = 43/139 (31%), Positives = 67/139 (49%), Gaps = 11/139 (7%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
 ++Q ++ LRRFGCS+C A + +K LD GV++V VG ++F+ +
 Sbjct 30 REQACVVAGLRRFGCSVCRWIAQDLSSLKGLLDQHGVRLVGVGP-EALGLQEFL-----D 83

Query 61 GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR----KANKNHPNADLQGDGL 116
 G F E+Y+D + Y+ G +R L + R KA L +L GD L
 Sbjct 84 GGYFAGELYLDESKQCYRELGFRRYNGLSIVPAALGKPVDRDVALKAKAVGIQGNLSGDLL 143

Query 117 QTGGIYLVGPGADSA-IHF 134
 Q+GG+ +V G D +HF
 Sbjct 144 QSGGLLVVTKGGDRVLLHF 162

>ref|NP_001051154.1| **UG** Os03g0729300 [Oryza sativa Japonica Group]
gb|AA038466.1| G unknown protein [Oryza sativa Japonica Group]
gb|ABF98681.1| G UPF0308 protein, chloroplast precursor, putative, expressed [Oryz
 sativa Japonica Group]
dbj|BAF13068.1| G Os03g0729300 [Oryza sativa Japonica Group]
dbj|BAG98418.1| G unnamed protein product [Oryza sativa Japonica Group]
 Length=259

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

Score = 57.0 bits (136), Expect = 9e-07, Method: Compositional matrix adjust.
 Identities = 42/127 (34%), Positives = 63/127 (50%), Gaps = 15/127 (11%)

Query 6 LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRF 65
 ++ LLR FGC C E AS + E + DAAG K++ +G G A + +P FP
 Sbjct 106 VVALLRHFGCFCCWELASVLKESMAKFDAAGAKLIAIGVGTDPKARILADGLP-----FP 160

Query 66 AE-VYIDPEQTAYKA----RGLQRVGLLHFLSWTAISEWRKANKNH----PNADLQGDGL 116
 + +Y DPE+ AY GL R + ++ ++ +K KN+ ADL G L
 Sbjct 161 VDSLADPERKAYDVLGLYHGLGRTLISPAMYSGLSIKKVTKNYTLKGT PADLTGI-L 219

Query 117 QTGGIYL 123
 Q GG+ +

Sbjct 220 QQGGMLV 226

>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 = 56.6 bits (135), Expect = 9e-07, Method: Compositional matrix adjust.
 Identities = 28/79 (36%), Positives = 46/79 (59%), Gaps = 5/79 (6%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGN 60
 KD++ ++ R FGC LC ++A ++ K +DA+GV +VL+G G+ A+ F E
 Sbjct 92 KDRKAVVAFARHFGCVLCRKRADYLAAKKDIMDASGVALVLIGPGSVDQAKTFSEQT--- 148
 Query 61 GQRFPAEVYIDPEQTAYKA 79
 +F EVY D ++Y+A
 Sbjct 149 --KFKGEVYADTSHSSYEA 165

>gb|EAY91739.1| hypothetical protein OsI_13380 [Oryza sativa Indica Group]
 Length=259

Score = 56.6 bits (135), Expect = 9e-07, Method: Compositional matrix adjust.
 Identities = 42/127 (34%), Positives = 63/127 (50%), Gaps = 15/127 (11%)

Query 6 LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGNQRF 65
 ++ LLR FGC C E AS + E + DAAG K++ +G G A + +P FP
 Sbjct 106 VVALLRHFGCFCCWELASVLKESMAKFDAAGAKLIAIGVGTDPKARILADGLP-----FP 160
 Query 66 AE-VYIDPEQTAYKA----RGLQRVGLLHFLSWTAISEWRKANKNH----PNADLQGDGL 116
 + +Y DPE+ AY GL R + ++ ++ +K KN+ ADL G L
 Sbjct 161 VDSLADPERKAYDVLGLYHGLGRTLISPAMYSGLNSIKKVTKNYTLKGT PADLTGI-L 219
 Query 117 QTGGIYL 123
 Q GG+ +
 Sbjct 220 QQGGMLV 226

>gb|EEE59859.1| hypothetical protein OsJ_12441 [Oryza sativa Japonica Group]
 Length=283

Score = 56.6 bits (135), Expect = 1e-06, Method: Compositional matrix adjust.
 Identities = 42/127 (34%), Positives = 63/127 (50%), Gaps = 15/127 (11%)

Query 6 LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGNQRF 65
 ++ LLR FGC C E AS + E + DAAG K++ +G G A + +P FP
 Sbjct 106 VVALLRHFGCFCCWELASVLKESMAKFDAAGAKLIAIGVGTDPKARILADGLP-----FP 160
 Query 66 AE-VYIDPEQTAYKA----RGLQRVGLLHFLSWTAISEWRKANKNH----PNADLQGDGL 116
 + +Y DPE+ AY GL R + ++ ++ +K KN+ ADL G L
 Sbjct 161 VDSLADPERKAYDVLGLYHGLGRTLISPAMYSGLNSIKKVTKNYTLKGT PADLTGI-L 219
 Query 117 QTGGIYL 123
 Q GG+ +
 Sbjct 220 QQGGMLV 226

>ref|ZP_02926219.1| hypothetical protein VspiD_06235 [Verrucomicrobium spinosum DSI
 4136]
 Length=304

Score = 56.6 bits (135), Expect = 1e-06, Method: Compositional matrix adjust.
 Identities = 40/143 (28%), Positives = 63/143 (45%), Gaps = 13/143 (9%)

Query 5 VLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGNQRF 64
 VL++ LR GC+ C E + + ++P ++A+G +I LV G F +GQ
 Sbjct 161 VLVVFLRHAGCTFCREALADIARVRPAIEASGTRIALVHMGEPAFAAF-----SQQYG 214
 Query 65 PAEV--YIDPEQTAYKARGLQRVGLLHFLSWTAISEWRKAN--KNHPNADLQGDGLQGTG 119
 A++ DP + Y+ GL+R L L W WR + H +GD Q
 Sbjct 215 LADLPAVADPSRRLYRGLGLRRGKLSQLLGWRV--WWRGLQAFLRGHRVKGFEQDVTQLP 272
 Query 120 GIYLVGPGADSAIHFAFNEYDHP 142

Sbjct 273 G++L+ G +F D P
GVFLIHRGVVLRRYFHQTSADRP 295

>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 SORBITRAFT_01g038790 | hypothetical protein [Sorghum bicolor]
(10 or fewer PubMed links)

Score = 56.6 bits (135), Expect = 1e-06, Method: Compositional matrix adjust.
Identities = 41/144 (29%), Positives = 68/144 (48%), Gaps = 17/144 (11%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
K+++ ++ R FGC LC ++A + + + AAGV +VL+G G+ A+ F E
Sbjct 101 KERKAVVAFARHFGCVLCRKRADLLAAKQDVMQAAGVALVLIGPGSVEQAKAFCEQT--- 157
Query 61 GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTA----ISEWRKANKN-----HPNAD 110
+F EVY DP ++Y A GL + A I +R+ + N
Sbjct 158 --KFKGEVYADPTHSSYDALEFA-FGLFSTFTPAAGLKIIQLYREGYRQDWELSFENR 214
Query 111 LQGDGLQTGGIYLVGPGADSAIHF 134
+G G GG+ + GPG D+ ++
Sbjct 215 TKG-GWYQGGLIVAGPGIDNIIYI 237

>ref|NP_001106912.1| **UG** prostamide/PG F synthase [Sus scrofa]
dbj|BAF96021.1| **G** prostamide/PG F synthase [Sus scrofa]
Length=202

GENE ID: 100134955 LOC100134955 | prostamide/PG F synthase [Sus scrofa]
(10 or fewer PubMed links)

Score = 56.6 bits (135), Expect = 1e-06, Method: Compositional matrix adjust.
Identities = 43/139 (31%), Positives = 66/139 (48%), Gaps = 11/139 (7%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
++0 ++ LRRFGC +C A + +K LD GV++V VG ++F+ +
Sbjct 30 QEQAQCVVAGLRRFGCMVCRWIARDLSSSLKGLLDQHGVRVGVGP-EALGLQEFL-----D 83
Query 61 GQRFPAEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWR----KANKNHPNADLQGDGL 116
G F ++Y+D + YK G +R L L R KA +L GD L
Sbjct 84 GGYFAGDLYLDESKQFYKELGFKRYSSLSILPAALGKPVDRVAAKAKAAGIQGNLSGDL 143
Query 117 QTGGIYLVGPGADSA-IHF 134
Q+GG+ +V G D +HF
Sbjct 144 QSGLLVVAKGGDKVLLHF 162

>ref|NP_001092155.1| **UG** hypothetical protein LOC100049742 [Xenopus laevis]
gb|AAI41728.1| **G** LOC100049742 protein [Xenopus laevis]
Length=228

GENE ID: 100049742 c9orf21 | chromosome 9 open reading frame 21
[Xenopus laevis] (10 or fewer PubMed links)

Score = 56.2 bits (134), Expect = 1e-06, Method: Compositional matrix adjust.
Identities = 37/139 (27%), Positives = 67/139 (49%), Gaps = 19/139 (13%)

Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVP 59
++++ +++ +R F C C E + +I L+ A V+++++G + E F ++ G
Sbjct 56 RERKTIVVFRNFLCYTCKEYVEDLAKIPSSALEDANVRLVIGQSSYIHIEHFC-SLTG 114
Query 60 NGQRFPAEVYIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHPN 108
+P E+Y+D ++T Y G+++ + LS + S WR P
Sbjct 115 ----YPYEMYVDTDRITYSKLGMMKGETSTSSGRSPHVKSNIILSGSIKSIWRAMTS--PA 168
Query 109 ADLQGDGLQTGGIYLVGPG 127
D QGD Q GG +VGPG
Sbjct 169 FDFQGDPAQGGSLIVGPG 187


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

Score = 56.2 bits (134), Expect = 1e-06, Method: Compositional matrix adjust.
Identities = 42/128 (33%), Positives = 64/128 (50%), Gaps = 15/128 (11%)

```
Query 6  LLILLRRFGCSLCHQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRFP 65
++ LLR FGC C E AS + E K + D+AGVK++ VG G A E +P FP
Sbjct 95  VVALLRHFGCPCCWELASALKESKARFDSAGVKLIAVIGITPNKARMLAERLP-----FP 149

Query 66  AE-VYIDPEQTAYKARGLQRVGLLHFLSWTAISEW-----RKANKNHP---NADLQGDG 115
+ +Y DP++ AY L F + ++I + +KA KN+ D +
Sbjct 150  LDCLYADPDRKAYHVLNLYYGFGRFFNPSSIKVFSRFDALQKAVKNYTIATPDDRSVG 209

Query 116  LQTGGIYL 123
LQ GG+++
Sbjct 210  LQQGGMFV 217
```

>sp|Q641F0.2|CJ058_XENLA  RecName: Full=UPF0765 protein C10orf58 homolog; Flags: F
Length=227



GENE ID: 447722 c10orf58 | chromosome 10 open reading frame 58 [Xenopus laevis]
(10 or fewer PubMed links)

Score = 56.2 bits (134), Expect = 1e-06, Method: Compositional matrix adjust.
Identities = 37/124 (30%), Positives = 63/124 (51%), Gaps = 11/124 (8%)

```
Query 6  LLILLRRFGCSLCHQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRFP 65
+++ +RR GC LC E+AS + +KPQLD GV + + E V F
Sbjct 75  VIMAVRRPGCFCLCREEASGLSTLKPQLDQLGVPLYAI-----VKENIGNEVEHFQPYFN 128

Query 66  AEVYIDPEQTAY--KARGLQRVGLLHFLSWTAISEWRKANKNHPNADLQGDGLQTGGIYL 123
+V++D + Y + R + +GL+ W +R+A K +L+G+GL GG+++
Sbjct 129  GKVFLDAKQFYGPQKRKMMLLGLVRLGVW---QNFRRAWKGGFEGNLEGEGLILGGMFV 185

Query 124  VGPG 127
+G G
Sbjct 186  IGSG 189
```

>ref|NP_001087861.1|  chromosome 10 open reading frame 58 [Xenopus laevis]
>gb|AAH82387.1|  MGC81827 protein [Xenopus laevis]
Length=210


GENE ID: 447722 c10orf58 | chromosome 10 open reading frame 58 [Xenopus laevis]
(10 or fewer PubMed links)

Score = 56.2 bits (134), Expect = 1e-06, Method: Compositional matrix adjust.
Identities = 37/124 (30%), Positives = 63/124 (51%), Gaps = 11/124 (8%)

```
Query 6  LLILLRRFGCSLCHQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRFP 65
+++ +RR GC LC E+AS + +KPQLD GV + + E V F
Sbjct 65  VIMAVRRPGCFCLCREEASGLSTLKPQLDQLGVPLYAI-----VKENIGNEVEHFQPYFN 118

Query 66  AEVYIDPEQTAY--KARGLQRVGLLHFLSWTAISEWRKANKNHPNADLQGDGLQTGGIYL 123
+V++D + Y + R + +GL+ W +R+A K +L+G+GL GG+++
Sbjct 119  GKVFLDAKQFYGPQKRKMMLLGLVRLGVW---QNFRRAWKGGFEGNLEGEGLILGGMFV 175

Query 124  VGPG 127
+G G
Sbjct 176  IGSG 179
```

>ref|XP_002708298.1|  PREDICTED: hypothetical protein [Oryctolagus cuniculus]
Length=226

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

Score = 56.2 bits (134), Expect = 1e-06, Method: Compositional matrix adjust.
Identities = 44/155 (29%), Positives = 73/155 (48%), Gaps = 22/155 (14%)

```
Query 1  KDQRVLLILLRRFGCSLCHQASHVLEIKPQ--LDAAGVKIVLVGTGNRYFAEKFIENVP 58
```

```

Sbjct 54 +D+R +++ +R F C +C E + ++ PQ L A V ++++G + + E F +
RDRRAVVVVRHFLCYVCKEYVEDLAKV-PQSFLREADVTLIVIGQSSYHHIEPFCKLT- 111
Query 59 GNGQRFPAEYVIDPEQTAYKARGLQRVGLL-----HFLSWTAISEWRKANKNHP 107
          + E+Y+DPE+ YK G++R + LS + S WR P
Sbjct 112 ----GYSHEIYVDPEREIYKRLGMRGEEIASSGRSPHIKSSLLSGSLRSLWRAVGTG--P 165
Query 108 NADLQGDGLQTGGIYLVGPGAD-SAIHFANNEYDH 141
          D QGD Q GG ++GPG + +H N DH
Sbjct 166 LFDLQGDPAQQGGTLILGPGNNIHFMLDRNRLDH 200

```

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

Score = 56.2 bits (134), Expect = 1e-06, Method: Compositional matrix adjust.
Identities = 42/128 (33%), Positives = 64/128 (50%), Gaps = 15/128 (11%)

```

Query 6 LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGNGQRF 65
      ++ LLR FGC C E AS + E K + D+AGVK++ VG G A E +P FP
Sbjct 95 VVALLRHFGCPCCWELASALKESKARFDSAGVKLIAVGIGTPNKARMLAERLP-----FP 149
Query 66 AE-VYIDPEQTAYKARGLQRVGLLHFLSWTAISEW-----RKANKNHP---NADLQGDG 115
      + +Y DP++ AY L F + ++I + +KA KN+ D +
Sbjct 150 LDCLYADPDRKAYHVLNLYYGFGRFFNPSSIKVFSRFDALQKAVKNYTIATPDDRSV 209
Query 116 LQTGGIYL 123
          LQ GG+++
Sbjct 210 LQQGGMFV 217

```

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

Score = 56.2 bits (134), Expect = 1e-06, Method: Compositional matrix adjust.
Identities = 43/128 (34%), Positives = 63/128 (50%), Gaps = 15/128 (11%)

```

Query 6 LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGNGQRF 65
      ++ LLR FGC C + AS + E K + D+AGVK++ VG G A E +P FP
Sbjct 3 VVALLRHFGCPCCWDLASALKESKERFDSAGVKLIAVGVGTPDKARILAERLP-----FP 57
Query 66 AE-VYIDPEQTAYKARGLQR-VGLLHFLSWTA-----ISEWRKANKNHP---NADLQGDG 115
      + +Y DP++ AY GL G F +A +KA KN+ D +
Sbjct 58 LDCLYADPDRKAYDVLGLYYGFGRFFNPASAKVLLRFEALQKAVKNYTIKATPDDKSSV 117
Query 116 LQTGGIYL 123
          LQ GG+++
Sbjct 118 LQQGGMFV 125

```

>**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 | NC100256614 | hypothetical protein [Vitis vinifera]
(10 or fewer PubMed links)

Score = 55.8 bits (133), Expect = 2e-06, Method: Compositional matrix adjust.
Identities = 43/128 (34%), Positives = 63/128 (50%), Gaps = 15/128 (11%)

```

Query 6 LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGNGQRF 65
      ++ LLR FGC C + AS + E K + D+AGVK++ VG G A E +P FP
Sbjct 100 VVALLRHFGCPCCWDLASALKESKERFDSAGVKLIAVGVGTPDKARILAERLP-----FP 154
Query 66 AE-VYIDPEQTAYKARGLQR-VGLLHFLSWTA-----ISEWRKANKNHP---NADLQGDG 115
      + +Y DP++ AY GL G F +A +KA KN+ D +
Sbjct 155 LDCLYADPDRKAYDVLGLYYGFGRFFNPASAKVLLRFEALQKAVKNYTIKATPDDKSSV 214
Query 116 LQTGGIYL 123
          LQ GG+++
Sbjct 215 LQQGGMFV 222

```

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

Score = 55.8 bits (133), Expect = 2e-06, Method: Compositional matrix adjust.
Identities = 43/128 (34%), Positives = 63/128 (50%), Gaps = 15/128 (11%)

```
Query 6 LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRF 65
      ++ LLR FGC C + AS + E K + D+AGVK++ VG G A E +P FP
Sbjct 100 VVALLRHFGCPCCWDLASALKESKERFDSAGVKLIAVGVGTPDKARILAERLP-----FP 154

Query 66 AE-VYIDPEQTAYKARGLQR-VGLLHFLSWTA-----ISEWRKANKNHP---NADLQGDG 115
      + +Y DP++ AY GL G F +A +KA KN+ D +
Sbjct 155 LDCLYADPDRKAYDVLGLYYGFGRFFNPASAKVLLRFEALQKAVKNYTIKATPDDKSSV 214

Query 116 LQTGGIYL 123
      LQ GG+++
Sbjct 215 LQQGGMFV 222
```

>**gb|AC012061.1** | C10orf58 homolog precursor [Lepeophtheirus salmonis]
Length=216

Score = 55.8 bits (133), Expect = 2e-06, Method: Compositional matrix adjust.
Identities = 40/136 (30%), Positives = 67/136 (50%), Gaps = 23/136 (16%)

```
Query 6 LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGT-----GNRYFAEKFIENVPGN 60
      +++++RR GC LC E+A ++IK L A + I LVG G FA F +
Sbjct 69 VIMVVRPFGCILCREEALFEMKIKSDLSA--LDIPLVGIVHEEEGAEFFASNFFTS---- 122

Query 61 GQRFPAEVYIDPEQTAY--KARGLQRVGLLHF-LSWTAISEWRKANKNHPNADLQGDGLQ 117
      ++VY D + + K R + GLL+F W+K + +L+GDG
Sbjct 123 -----SDVYFDINKKFFGPKERRIMLTGLLNFRFILKTFGAWKKG----VSGNLEGDGSL 173


Query 118 TGGIYLVGPGADSAIH 133
      GG +++GPG++ ++
Sbjct 174 LGGTFVMGPGSEGVLY 189
```


>**gb|EEC75002.1** | hypothetical protein OsI_11064 [Oryza sativa Indica Group]
Length=239

Score = 55.8 bits (133), Expect = 2e-06, Method: Compositional matrix adjust.
Identities = 27/79 (35%), Positives = 44/79 (56%), Gaps = 5/79 (6%)

```
Query 1 KDQRVLLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
      KD++ ++ R FGC LC ++A + + ++AAGV +VL+G G A+ F +
Sbjct 82 KDRKAIVAFARHFGCVLCRKRADLLAAKQDAMEAAGVALVLIGPGTVEQAKAFYDQT--- 138

Query 61 GQRFPAEVYIDPEQTAYKA 79
      +F EVY DP ++Y A
Sbjct 139 --KFKGEVYADPSHSSYNA 155
```

>**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 = 55.8 bits (133), Expect = 2e-06, Method: Compositional matrix adjust.
Identities = 42/128 (33%), Positives = 62/128 (49%), Gaps = 15/128 (11%)

```
Query 6 LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRF 65
      ++ LLR FGC C E AS + E K + D+AGVK++ +G G A + +P FP
Sbjct 99 VVALLRHFGCPCCWELASVLKEAKSKFDSAGVKLIAIGVGAPNKARMLADRLP-----FP 153

Query 66 AE-VYIDPEQTAYKARGLQR-VGLLHFLSWTA-----ISEWRKANKNHP---NADLQGDG 115
      + +Y DP + AY GL G F +A R+A KN+ D +
Sbjct 154 MDCLYADPNREAYNVLGLYYGFGRFFNPASAKVFSRFDLSLRQAVKNYTIATPDDRSV 213

Query 116 LQTGGIYL 123
      LQ GG+++
Sbjct 214 LQQGGMFV 221
```

>**ref|XP_002191279.1** |  PREDICTED: hypothetical protein [Taeniopygia guttata]
Length=222

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

Score = 55.5 bits (132), Expect = 2e-06, Method: Compositional matrix adjust.
Identities = 42/153 (28%), Positives = 73/153 (48%), Gaps = 20/153 (13%)

```
Query 2   DQVLLILLRRFGCSLCHEQASHVLEI-KPQLDAAGVKIVLVGTGNRYFAEKFIENVPGN 60
          +Q+ +++ +R F C C E + ++ K L + V+++++G + + + F ++ G
Sbjct 51  EQKAIVLVFNFLCYTCKEYVEDLAKVPKAFQESNVRLIVIGQSSYHHIKPFC-SLTG- 108

Query 61  GQRFPAEVYIDPEQTAYKARGLQR-----VGLLHFLSWTAI----SEWRKANKNHPNA 109
          + E+Y+DP + YK G++R V H S T + S WR P
Sbjct 109 ---YTHEMYVDPREIYKILGMKRGEGNKASVRSPhVKSNTFLGSIRSIRWRAMTG--PAF 163

Query 110 DLQGDGLQTGGIYLVGPGADSA-IHFANEYDH 141
          D QGD Q GG ++GPG + +H N DH
Sbjct 164 DFQGDPAQQGGALIIGPGNEVHFLHLDKNRLDH 196
```

>ref|NP_001180474.1| **G** selenoprotein U [Oryzias latipes]
Length=212

GENE ID: 100499408 selU | selenoprotein U [Oryzias latipes]
(10 or fewer PubMed links)

Score = 55.5 bits (132), Expect = 2e-06, Method: Compositional matrix adjust.
Identities = 37/127 (30%), Positives = 62/127 (49%), Gaps = 7/127 (5%)

```
Query 6   LLILLRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRFP 65
          +++ +RR G LC E+AS + +KPQL+ GV +V V E + F
Sbjct 65  VVMAVRRPGUFLCREEASELSSLKPQLEELGVPLVAV-----VKENLGSEIQDFRPHFA 118

Query 66  AEVYIDPEQTAYKARGLQRVGLLHFLSWTAISEWRKANKNHPNADLQGDGLQTGGIYLVG 125
          ++YID E+ Y +R+G L F+ + +A K+ ++ G+G GG+Y++G
Sbjct 119 GDIYIDEEKRFYGPL-QRRMGGLGFIIRIGVWQNFIRAWKSGYQGNMNGEGFILGGVYVIG 177

Query 126 PGADSAI 132
          G I
Sbjct 178 AGEQGII 184
```

>emb|CAG09571.1| unnamed protein product [Tetraodon nigroviridis]
Length=189

Score = 55.5 bits (132), Expect = 2e-06, Method: Compositional matrix adjust.
Identities = 43/152 (29%), Positives = 66/152 (44%), Gaps = 21/152 (13%)

```
Query 4   RVLLILLRRFGCSLCHEQASHVLEIKPQ-LDAAGVKIVLVGTGNRYFAEKFIENVPNGNQ 62
          + ++I +R F C C E + +I L AG+++V++G E F ++ G
Sbjct 19  KSVIIFVFNFLCYTCKEYVEDLSKIPEDVLKDAGIRLVVIGQSLHRHIEAFC-SLTG--- 74

Query 63  RFPAEVYIDPEQTAYKARGLQRVGLLH-----FLSWTAISEWRKANKNHPNAD 110
          +P E+Y+DP++ Y+ G++R + S WR P D
Sbjct 75 -YPYEMYVDPDRYIYQKLGMRREETFTDSAQSPHVKSGVFAGQMKSIWRAMTG--PIFD 131

Query 111 LQGDGLQTGGIYLVGPGADSAI-HFANEYDH 141
          QGD Q GG + GPGA HF N DH
Sbjct 132 FQGDHLHQGGAIITGPGAQVLFCHFDTNRLDH 163
```

>sp|Q8TBF2.1|CA093_HUMAN **G** RecName: Full=Uncharacterized protein Clorf93
gb|AAP97295.1|AF425266_1 **G** unknown protein [Homo sapiens]
gb|AAH22547.1 **G** Chromosome 1 open reading frame 93 [Homo sapiens]
gb|EAW56087.1 **G** chromosome 1 open reading frame 93, isoform CRA_e [Homo sapiens]
emb|CAX30827.1 **G** chromosome 1 open reading frame 93 [Homo sapiens]
Length=198

GENE ID: 127281 Clorf93 | chromosome 1 open reading frame 93 [Homo sapiens]
(10 or fewer PubMed links)

Score = 55.5 bits (132), Expect = 2e-06, Method: Compositional matrix adjust.
Identities = 42/130 (33%), Positives = 57/130 (44%), Gaps = 11/130 (8%)


```
Query 10  LRRFGCSLCHEQASHVLEIKPQLDAAGVKIVLVGTGNRYFAEKFIENVPNGQRFPAEVY 69
          LRRFGC +C  A  +  +  LD  GV++V VG      E      +G  F  E+Y
Sbjct 39  LRRFGCVVCRWIAQDLSSLAGLLDQHGVRVLVGVGPEALGLQEF-----DGDYFAGELY 92

Query 70  IDPEQTAYKARGLQRVGLLHFLSWTAISEWR----KANKNHPNADLQGDGLQTGGIYLVG 125
          +D  +  YK  G +R  L  L      R  KA      +L  GD  LQ+GG+ +V
Sbjct 93  LDESKQLYKELGFKRYNSLSILPAALGKPVVDVAAKAKAVGIQGNLSGDLLQSGGLLVVS 152

Query 126 PGADSA-IHF 134
          G D  +HF
Sbjct 153 KGGDKVLLHF 162
```