

Table S5. Eukarya mRNA (and other non-rRNA) gene sequences from VS. ["n" indicates information not specified in the NCBI GenBank database.]

454 Sequence ID	Q length	Q start	Q end	e-value	%ident	%sim	GI number	Domain	Phylum	Class / Order	Genus / Species	Description
GWKQY010B20UL	80	1	80	3e-33	100%	100%	294935818	Eukaryota	Incertae sedis	Perkinsea	Perkinsus marinus	Perkinsus marinus ATCC 50983 hypothetical protein, mRNA
GWKQY010A4UTS	128	1	128	97%	97%	50657591	Eukaryota	Rhodophyta	Rhodophyta	Gracilaria tenuisipitata	Gracilaria tenuisipitata var. lili chloroplast, complete genome	
GWKQY010AVBWV	332	1	332	2e-161	98%	98%	24665761	Eukaryota	Arthropoda	Branchiopoda	Daphnia pulex	Daphnia pulex ncRNA, kairamome-isoducube transcript
GWKQY010ACUK	411	1	411	0	95%	95%	155094333	Eukaryota	Arthropoda	Insecta	Drosophila grimshawi	Drosophila grimshawi GH23874 (Dgr)(GH23874), mRNA
GWKQY010BD9Y	393	1	393	0	96%	96%	155120240	Eukaryota	Arthropoda	Insecta	Drosophila grimshawi	Drosophila grimshawi GH23709 (Dgr)(GH23709), mRNA
GWKQY010BUCFO	306	1	306	3e-144	97%	97%	25533587	Eukaryota	Arthropoda	Insecta	Euphydryas editha	Euphydryas editha clone euphyr 50 microsatellite sequence
GWKQY010BQOY	266	1	266	3e-84	88%	88%	153902485	Eukaryota	Arthropoda	Insecta	Gryllus bimaculatus	Gryllus bimaculatus mRNA, Gbcontig04373
GWKQY010B9E4	370	1	112	1e-25	88%	88%	153919432	Eukaryota	Arthropoda	Insecta	Gryllus bimaculatus	Gryllus bimaculatus mRNA, Gbcontig29766
GWKQY010AMXUS	492	1	492	6e-104	82%	82%	67540319	Eukaryota	Ascomycota	Eurotiomycetes	Emerizella nidulans	Aspergillus nidulans FGSC 4-e elongation factor 2 partial mRNA
GWKQY010A6T8	556	1	556	0	95%	95%	242805358	Eukaryota	Ascomycota	Talaromycetes	Talaromyces stipitatus	Talaromyces stipitatus ATCC 10500 hypothetical protein, mRNA
GWKQY010ASANE	72	1	72	1e-18	92%	92%	50542891	Eukaryota	Ascomycota	Yarrowia lipolytica	Yarrowia lipolytica YAL0040352p (YAL0040352g) mRNA, complete cds	
GWKQY010B8ZF	47	1	47	1e-09	94%	94%	49343205	Eukaryota	Ascomycota	Saccharomycetes	Kluyveromyces fragilis	Kluyveromyces fragilis clone euflp1 50 microsatellite sequence
GWKQY010ATCHW	494	1	494	0	99%	99%	238033210	Eukaryota	Ascomycota	Saccharomycetes	Pichia pastoris	Pichia pastoris GS115 chromosome 4, complete sequence
GWKQY010AFQK1	28	1	28	0.000005	100%	100%	259145041	Eukaryota	Ascomycota	Saccharomycetes	Saccharomyces cerevisiae	Saccharomyces cerevisiae EC1118 chromosome IV, EC1118_100 genomic scaffold, whole genome shotgun sequence
GWKQY010A475X	79	1	79	1e-32	100%	100%	259147625	Eukaryota	Ascomycota	Saccharomycetes	Saccharomyces cerevisiae	Saccharomyces cerevisiae EC1118 chromosome XI, EC1118_145 genomic scaffold, whole genome shotgun sequence
GWKQY010A8BH	115	1	115	2e-47	97%	97%	259147931	Eukaryota	Ascomycota	Saccharomycetes	Saccharomyces cerevisiae	Saccharomyces cerevisiae EC1118 chromosome XII, EC1118_110 genomic scaffold, whole genome shotgun sequence
GWKQY010AQTOM	135	1	135	4e-51	94%	94%	259149327	Eukaryota	Ascomycota	Saccharomycetes	Saccharomyces cerevisiae	Saccharomyces cerevisiae EC1118 chromosome XV, EC1118_104 genomic scaffold, whole genome shotgun sequence
GWKQY010B8ZF_2	140	1	140	2e-19	78%	78%	23630288	Eukaryota	Basidiomycota	Tremellomycetes	Cryptococcus neoformans	Cryptococcus neoformans var. grubii strain H99 mitochondrion, complete genome
GWKQY010B8ZF_3	151	1	151	5e-15	78%	78%	315465309	Eukaryota	Basidiomycota	Ustilaginomycetes	Sporisorium	Sporisorium reilianum SR22 mitochondrial DNA sequence
GWKQY010ANHKA	513	1	513	0	91%	91%	71483398	Eukaryota	Basidiomycota	Microbotryales	Microbotryum violaceum	Microbotryum violaceum isolate Doylestris, 9119 elongation factor 1alpha (EF1) gene, partial cds
GWKQY010B9NMN	284	1	284	7e-116	94%	94%	171191502	Eukaryota	Cercozoa	Imbricaria	Paulinella chromatophora	Paulinella chromatophora chromatophore, complete genome
GWKQY010B5D9R	103	1	100	2e-42	99%	99%	56159573	Eukaryota	Chlorophyta	Ulvophyceae	Pseudodictyonium akinetum	Pseudodictyonium akinetum chloroplast, complete genome
GWKQY010A15DF	332	1	332	2e-47	78%	78%	156307499	Eukaryota	Cnidaria	Anthozoa	Nematostella vectensis	Nematostella vectensis predicted protein (NMVEDRAFT_v1g225918) partial mRNA
GWKQY010A8687	53	1	53	3e-15	96%	96%	253981836	Eukaryota	Cryptophyta	Cryptophyceae	Cryptomonas paramecium	Cryptomonas paramecium strain CCAP977/2a plastid, complete genome
GWKQY010A7F9N	185	1	185	4e-91	100%	100%	166044520	Eukaryota	Heterokontophyta	Oomycota	Aphanomyces euteiches	Aphanomyces euteiches cDNA
GWKQY010B1T0Z	393	1	393	0	98%	98%	166044517	Eukaryota	Heterokontophyta	Oomycota	Aphanomyces euteiches	Aphanomyces euteiches cDNA
GWKQY010A1XK3	28	1	28	0.000005	100%	100%	136170404	Eukaryota	Kinetoplastida	Trypanosomatidae	Schizotrypanum	Trypanosoma cruzi clone lie_16 SMUG 5 protein gene, partial cds
GWKQY010ABPWQ	362	1	362	2e-177	98%	98%	154347362	Eukaryota	Percolozoa	Heterolobosea	Naegleria gruberi	Naegleria gruberi extrachromosomal rDNA plasmid DNA, complete sequence
GWKQY010A5AP6	96	1	96	1e-38	98%	98%	124012151	Eukaryota	Streptophyta	Chlorokybophyceae	Chlorokybus atmophyllicus	Chlorokybus atmophyllicus chloroplast, complete genome
GWKQY010A27HH	105	1	105	2e-37	94%	94%	210148086	Eukaryota	Streptophyta	Cornales	Cornus kousa	Cornus kousa clone CK083 microsatellite sequence
GWKQY010BHQDQ	218	1	218	7e-45	83%	83%	13445170	Eukaryota	Streptophyta	Liliopsida	Festuca arundinacea	Festuca arundinacea genomic FaA5 repeat element
GWKQY010A88A7	442	1	442	0	93%	93%	217799547	Eukaryota	Streptophyta	Liliopsida	Aegilops tauschii	Aegilops tauschii leucine-rich-vlae protein gene, partial cds; seed globulin (Glu-2) gene, complete cds; HMW-glutenin (Glu-Dhl) gene, Glu-Dry-T2 allele, complete cds; and HMW-glutenin (HMW) gene, HMW-Dst2 allele, complete cds
GWKQY010BPH2R	440	1	440	2e-109	83%	83%	32307244	Eukaryota	Streptophyta	Liliopsida	Aegilops tauschii	Aegilops tauschii clone 4P6-2 tandem repeat sequence
GWKQY010A1HHQ	243	1	243	3e-198	97%	97%	40849982	Eukaryota	Streptophyta	Liliopsida	Triticum turgidum	Triticum turgidum A genome HMW glutenin A gene locus, sequence
GWKQY010AAMK	81	1	81	2e-26	95%	95%	72256311	Eukaryota	Streptophyta	Liliopsida	Triticum aestivum	Triticum aestivum clone wmc3421 microsatellite sequence
GWKQY010B9HEM	334	1	334	6e-63	83%	83%	102567891	Eukaryota	Streptophyta	Liliopsida	Zea perennis	Zea perennis mitochondrion, complete genome
GWKQY010A1Q6W	300	1	299	2e-113	92%	92%	115392331	Eukaryota	Streptophyta	Liliopsida	Triticum monoccoccum	Triticum monoccoccum subsp. aegilopoides clone BAC Tb8AC5, complete sequence
GWKQY010A1Q6W_2	181	1	181	5e-70	94%	94%	115392331	Eukaryota	Streptophyta	Liliopsida	Triticum monoccoccum	Triticum monoccoccum subsp. aegilopoides clone BAC Tb8AC5, complete sequence
GWKQY010B9WVC	492	1	492	0	91%	91%	124007744	Eukaryota	Streptophyta	Liliopsida	Triticum urartu	Triticum urartu clone BAC TC18 genomic sequence
GWKQY010A1NTW	285	1	285	6e-131	97%	97%	148372275	Eukaryota	Streptophyta	Liliopsida	Triticum monoccoccum	Triticum monoccoccum subsp. aegilopoides clone Tb8AC30, complete sequence
GWKQY010B9T52	492	1	492	0	99%	99%	148910867	Eukaryota	Streptophyta	Liliopsida	Triticum turgidum	Triticum turgidum overlapping clones BAC 32652 and BAC 354M17, complete sequence
GWKQY010A1Q457	348	1	348	3e-100	87%	87%	194131647	Eukaryota	Streptophyta	Liliopsida	Triticum turgidum	Triticum turgidum subsp. dicoccoides clones BAC 391M13 and BAC 114M420 genomic sequence
GWKQY010B1U9T	262	1	262	2e-106	94%	94%	194239068	Eukaryota	Streptophyta	Liliopsida	Triticum aestivum	Triticum aestivum 38 chromosome, clone BAC TA3895C9
GWKQY010B1G6X	432	1	432	0	100%	100%	209361311	Eukaryota	Streptophyta	Liliopsida	Coix lacryma-jobi	Coix lacryma-jobi chloroplast, complete genome
GWKQY010A2CWNH	154	1	154	2e-13	77%	77%	212007811	Eukaryota	Streptophyta	Liliopsida	Triticum aestivum	Triticum aestivum clone BAC 502020, complete sequence
GWKQY010B1W1X8	383	1	383	0	99%	99%	255099160	Eukaryota	Streptophyta	Liliopsida	Dendrocalamus latiflorus	Dendrocalamus latiflorus chloroplast, complete genome
GWKQY010A1N3LX	127	1	127	2e-57	99%	99%	219819090	Eukaryota	Streptophyta	Apiales	Daucus carota	Daucus carota subsp. sativus clone BAC C034E06 genomic sequence
GWKQY010B1UC5Q	578	1	578	2e-123	83%	83%	88659961	Eukaryota	Streptophyta	Asterales	Lactuca sativa	Lactuca sativa cultivar Salinas chloroplast, complete genome
GWKQY010B1NTYY	300	273	300	0.001	100%	100%	255536188	Eukaryota	Streptophyta	Asteriales	Artemisia annua	TA: Artemisia annua strain Artemis Contig17, mRNA sequence
GWKQY010A2B8H	232	1	232	1e-107	98%	98%	194132090	Eukaryota	Streptophyta	Malvales	Gonythylus	Gonythylus bancanus chloroplast, partial genome
GWKQY010A1FTF	275	1	275	3e-139	99%	99%	170522340	Eukaryota	Streptophyta	Brassicales	Carica papaya	Carica papaya mitochondrion, complete genome
GWKQY010A2A3T	198	1	198	1e-96	99%	99%	62149314	Eukaryota	Streptophyta	Caryophyllales	Silene latifolia	Silene latifolia chloroplast rpsL, rpsC1, rpsC2, rpsL2, atpH, atpF, atpA, psbI, psbK, rps16, matK, psbA, rps12, rps12, rps12, rps7 genes, complete cds
GWKQY010A4P8N	428	1	428	0	94%	94%	193788921	Eukaryota	Streptophyta	Fabales	Trifolium subterraneum	Trifolium subterraneum chloroplast, complete genome
GWKQY010A4746	335	1	335	2e-152	96%	96%	293338622	Eukaryota	Streptophyta	Fabales	Lathyrus sativus	Lathyrus sativus cultivar Cicerchia Marchigiana chloroplast, complete genome
GWKQY010A7TGY	514	1	514	0	93%	93%	210143279	Eukaryota	Streptophyta	Fabales	Glycine max	Glycine max cDNA, clone: GmF01-21-C10
GWKQY010B6GKO	394	1	394	1e-159	93%	93%	218135405	Eukaryota	Streptophyta	Fabales	Medicago truncatula	Medicago truncatula chromosome 4 clone mth2-13520, complete sequence
GWKQY010A1C94	345	1	345	4e-178	99%	99%	29124647	Eukaryota	Streptophyta	Solanales	Solanum	Solanum tuberosum isolate DM1-3-516-R44 chloroplast, complete genome
GWKQY010B05Q3	467	1	467	0	96%	96%	147776538	Eukaryota	Streptophyta	Vitales	Vitis vinifera	Vitis vinifera contig VV78X160271.3, whole genome shotgun sequence
GWKQY010B9Y2M	52	1	52	5e-13	94%	94%	147820696	Eukaryota	Streptophyta	Vitales	Vitis vinifera	Vitis vinifera contig VV78X065077.2, whole genome shotgun sequence
GWKQY010B1TYU0	537	1	537	0	99%	99%	23976707	Eukaryota	Streptophyta	Vitales	Vitis vinifera	Vitis vinifera strain PN40024 mitochondrion, partial genome