

Table S5. Eukarya mRNA (and other non-rRNA) gene sequences from V5.

| Sequence ID     | Q length | Q start | Q end | e-value | %-ident | %-sim | Gt hit    | Domain    | Phylum           | Class/Order        | Genus/Species             | Description                                                                                                                                                                                                            | Product |
|-----------------|----------|---------|-------|---------|---------|-------|-----------|-----------|------------------|--------------------|---------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|
| GWKQY01B20DL    | 80       | 1       | 80    | 3E-33   | 1       | 1     | 294955818 |           | incertae sedis   | Perkinsea          | Perkinsus marinus         | Perkinsus marinus ATCC 50983 hypothetical 1 product="hypothetical protein"                                                                                                                                             |         |
| GWKQY01A4E1T5   | 128      | 1       | 128   | 6E-53   | 0.97    | 0.97  | 50657597  | Eukaryota | Rhodophyta       | Rhodophyceae       | Gracilaria tenuispiitata  | Gracilaria tenuispiitata var. luu chloroplast, c product="large subunit ribosomal RNA"                                                                                                                                 |         |
| GWKQY01A5B9VX   | 332      | 1       | 332   | 2E-161  | 0.98    | 0.98  | 294661761 | Eukaryota | Anthrhopoda      | Brachyopoda        | Daphnia pulex             | Daphnia pulex cDNA, salivomone-inducible 1 product="salivomone-inducible transcript"                                                                                                                                   |         |
| GWKQY01A2EVC    | 445      | 1       | 445   | 0       | 0.96    | 0.96  | 83281396  | Eukaryota | Arthropoda       | Insecta            | Culex quinquefasciatus    | Culex pipiens quinquefasciatus, clone Culex p.rpt_family="LSU-rRNA_rhA"                                                                                                                                                |         |
| GWKQY01A0BVK    | 28       | 1       | 28    | 0.002   | 1       | 1     | 113193577 | Eukaryota | Arthropoda       | Insecta            | Drosophila melanogaster   | Drosophila melanogaster chromosome X, cor note="CG32547 gene product from transcript CG32547-RC"                                                                                                                       |         |
| GWKQY01A2CJKE   | 411      | 1       | 411   | 0       | 0.95    | 0.95  | 195094123 | Eukaryota | Arthropoda       | Insecta            | Drosophila grimshawi      | Drosophila grimshawi cDNA, GH23874 gene product from transcript GH23874-RA, GH23874-PA"                                                                                                                                |         |
| GWKQY01B0B9Y    | 393      | 1       | 393   | 0       | 0.96    | 0.96  | 195102030 | Eukaryota | Arthropoda       | Insecta            | Drosophila grimshawi      | Drosophila grimshawi GH23709 (Dgr)GH23709 gene product from transcript GH23709-RA; GH23709-PA"                                                                                                                         |         |
| GWKQY01B3CYFO   | 306      | 1       | 306   | 3E-141  | 0.97    | 0.97  | 291939387 | Eukaryota | Arthropoda       | Insecta            | Euphydryas editha         | Euphydryas editha clone euphy 50 microsat rpt_type=tandem; satellite="microsatellite"                                                                                                                                  |         |
| GWKQY01B3KQD    | 266      | 1       | 266   | 3E-84   | 0.88    | 0.88  | 151902485 | Eukaryota | Arthropoda       | Insecta            | Gryllus bimaculatus       | Gryllus bimaculatus mRNA, GRCcontig4373 mol_type="mRNA"                                                                                                                                                                |         |
| GWKQY01B0B64    | 370      | 1       | 370   | 1E-25   | 0.88    | 0.88  | 151919430 | Eukaryota | Arthropoda       | Insecta            | Gryllus bimaculatus       | Gryllus bimaculatus mRNA, GRCcontig29766 mol_type="mRNA"                                                                                                                                                               |         |
| GWKQY01A1KJUG   | 492      | 1       | 492   | 6E-104  | 0.82    | 0.82  | 67540319  | Eukaryota | Ascomycota       | Euriotomycetes     | Emmericella nidulans      | Aspergillus nidulans FGSC A4 elongation factor product="EF2, NEUCR Elongation factor 2 (EF-2) [Colonial temperature-sensitive 3]"                                                                                      |         |
| GWKQY01A0A7B    | 556      | 1       | 556   | 0       | 0.95    | 0.95  | 242805358 | Eukaryota | Ascomycota       | Euriotomycetes     | Talaromyces stipitatus    | Talaromyces stipitatus ATCC 10500 hypotheit mol_type="mRNA", hypothetical                                                                                                                                              |         |
| GWKQY01A5A5AN   | 72       | 1       | 72    | 1E-18   | 0.92    | 0.92  | 56542891  | Eukaryota | Ascomycota       | Saccharomycetes    | Yarrowia lipolytica       | Yarrowia lipolytica YAL040052p (YAL04005) mol_type="mRNA"                                                                                                                                                              |         |
| GWKQY01B3KQD    | 491      | 1       | 491   | 0       | 0.95    | 0.95  | 172190    | Eukaryota | Ascomycota       | Saccharomycetes    | Saccharomyces cerevisiae  | Saccharomyces cerevisiae 2 micron circle plar product="recombinase (FLP)"                                                                                                                                              |         |
| GWKQY01B3K6Z    | 47       | 1       | 47    | 1E-09   | 0.94    | 0.94  | 49343295  | Eukaryota | Ascomycota       | Saccharomycetes    | Kluyveromyces lactis      | Kluyveromyces lactis mitochondrion, comple product="large subunit ribosomal RNA"                                                                                                                                       |         |
| GWKQY01A5ATCW   | 494      | 1       | 494   | 0       | 0.99    | 0.99  | 238032120 | Eukaryota | Ascomycota       | Saccharomycetes    | Pichia pastoris           | Pichia pastoris GS115 chromosome 4, comp="genomic DNA"                                                                                                                                                                 |         |
| GWKQY01A1F0K1   | 28       | 1       | 28    | 5E-06   | 1       | 1     | 259145041 | Eukaryota | Ascomycota       | Saccharomycetes    | Saccharomyces cerevisiae  | Saccharomyces cerevisiae EC1118 chromosom note="Similar to YDR410C STE14 Farnesyl cysteine-carboxyl methyltransferase"                                                                                                 |         |
| GWKQY01A475X    | 79       | 1       | 79    | 1E-32   | 1       | 1     | 259147625 | Eukaryota | Ascomycota       | Saccharomycetes    | Saccharomyces cerevisiae  | Saccharomyces cerevisiae EC1118 chromosom note="Similar to YKL117W SBA1 Co-chaperone that binds to and regulates Hsp90 family chaperones"                                                                              |         |
| GWKQY01A4B9H    | 115      | 1       | 115   | 2E-47   | 0.97    | 0.97  | 259147931 | Eukaryota | Ascomycota       | Saccharomycetes    | Saccharomyces cerevisiae  | Saccharomyces cerevisiae EC1118 chromosom note="10M37.1 35S ribosomal RNA transcript, encoded by the RDN1 locus, that is processed into the 25S, 18S and 5.8S rRNAs (represented by the RDN25, RDN18, and RDN58 loci)" |         |
| GWKQY01A1Q1TM   | 135      | 1       | 135   | 4E-51   | 0.94    | 0.94  | 259149327 | Eukaryota | Ascomycota       | Saccharomycetes    | Saccharomyces cerevisiae  | Saccharomyces cerevisiae EC1118 chromosom mol_type="genomic DNA"                                                                                                                                                       |         |
| GWKQY01B3K6Z_2  | 140      | 1       | 140   | 2E-19   | 0.78    | 0.78  | 23630288  | Eukaryota | Basidiomycota    | Tremellomycetes    | Cryptococcus neoformans   | Cryptococcus neoformans var. grubii strain H mol_type="genomic DNA"                                                                                                                                                    |         |
| GWKQY01B3K6Z_3  | 151      | 1       | 151   | 5E-15   | 0.78    | 0.78  | 315465309 | Eukaryota | Basidiomycota    | Ustilaginomycetes  | Sporisorium               | Sporisorium reilianum SR22 mitochondrial D mol_type="genomic DNA"                                                                                                                                                      |         |
| GWKQY01A1M1KA   | 513      | 1       | 513   | 0       | 0.91    | 0.91  | 71483398  | Eukaryota | Basidiomycota    | Microbotryales     | Microbotryum violaceum    | Microbotryum violaceum isolate Dreyestr., product="elongation factor Talpha"                                                                                                                                           |         |
| GWKQY01A1FQZ8   | 257      | 1       | 257   | 5E-62   | 0.85    | 0.85  | 72256214  | Eukaryota | Basidiomycota    | Ustilaginomycetes  | Ustilago maydis           | Ustilago maydis mitochondrion, complete ger product="small subunit ribosomal RNA"                                                                                                                                      |         |
| GWKQY01B5A1F    | 484      | 1       | 484   | 1E-56   | 0.76    | 0.76  | 71022004  | Eukaryota | Basidiomycota    | Ustilaginomycetes  | Ustilago maydis           | Ustilago maydis 521 hypothetical protein (UA product="hypothetical protein"                                                                                                                                            |         |
| GWKQY01B0BNNM   | 284      | 1       | 284   | 7E-116  | 0.84    | 0.84  | 171191502 | Eukaryota | Cercozia         | Imbricariae        | Paulinella chromatophora  | Paulinella chromatophora chromatophore, c product="23S ribosomal RNA"                                                                                                                                                  |         |
| GWKQY01B5D9R    | 103      | 1       | 100   | 2E-42   | 0.99    | 0.99  | 56159573  | Eukaryota | Chlorophyta      | Ulvophyceae        | Pseudodictyonium akinetum | Pseudodictyonium akinetum chloroplast, cor product="small subunit ribosomal RNA"                                                                                                                                       |         |
| GWKQY01A1H4RC   | 194      | 1       | 194   | 2E-40   | 0.81    | 0.83  | 45423887  | Eukaryota | Chordata         | Aves               | Gallus gallus             | Gallus gallus finished cDNA, clone CHEST2155 mol_type="mRNA"                                                                                                                                                           |         |
| GWKQY01A1A1FO   | 439      | 1       | 439   | 2E-149  | 0.89    | 0.89  | 118085917 | Eukaryota | Chordata         | Aves               | Gallus gallus             | PHEDICTED: Gallus gallus similar to putative t-mol_type="mRNA"                                                                                                                                                         |         |
| GWKQY01B3KXNM   | 328      | 1       | 328   | 1E-124  | 0.92    | 0.92  | 154937557 | Eukaryota | Chordata         | Aves               | Gallus gallus             | Gallus gallus BAC clone CH261-57G10 from c mol_type="genomic DNA"                                                                                                                                                      |         |
| GWKQY01A27H1Z   | 226      | 1       | 226   | 4E-18   | 0.7     | 0.7   | 69749857  | Eukaryota | Chordata         | Mammalia           | Felis                     | Felis catus clone rp86-207p24, complete seq clone_3b0="RP81-86 Male Feline BAC Library"                                                                                                                                |         |
| GWKQY01A1A1DF   | 332      | 1       | 332   | 2E-47   | 0.78    | 0.78  | 156307499 | Eukaryota | Chitridia        | Amphozoa           | Nematostella vectensis    | Nematostella vectensis predicted protein (N6 product="hypothetical protein"                                                                                                                                            |         |
| GWKQY01A1B6B7   | 51       | 1       | 51    | 3E-15   | 0.96    | 0.96  | 253981836 | Eukaryota | Cryptophyta      | Cryptophyceae      | Cryptomonas paramecium    | Cryptomonas paramecium strain CCAP97/2, product="23S large subunit ribosomal RNA"                                                                                                                                      |         |
| GWKQY01A1Y5N9   | 185      | 1       | 185   | 4E-91   | 1       | 1     | 166044520 | Eukaryota | Heterokontophyta | Oomycota           | Aphanomyces euteiches     | Aphanomyces euteiches cDNA mol_type="mRNA"                                                                                                                                                                             |         |
| GWKQY01B1B7E2   | 395      | 1       | 395   | 0       | 0.98    | 0.98  | 166044517 | Eukaryota | Heterokontophyta | Oomycota           | Aphanomyces euteiches     | Aphanomyces euteiches cDNA mol_type="mRNA"                                                                                                                                                                             |         |
| GWKQY01A1A0X5   | 28       | 1       | 28    | 5E-06   | 1       | 1     | 336170404 | Eukaryota | Kinetoplastidae  | Trypanosomatidae   | Schizotrypanum            | Trypanosoma cruzi clone He_16 SMUG 5 prot product="SMUG 5 protein"                                                                                                                                                     |         |
| GWKQY01A18PW0   | 362      | 1       | 362   | 2E-177  | 0.98    | 0.98  | 154347362 | Eukaryota | Percolozoa       | Naeheriobozoa      | Naeheria gruberi          | Naeheria gruberi extrachromosomal rDNA pl product="28S ribosomal RNA"                                                                                                                                                  |         |
| GWKQY01A3AP96   | 96       | 1       | 96    | 1E-38   | 0.98    | 0.98  | 124002151 | Eukaryota | Streptophyta     | Chlorokontophyceae | Chlorokybush aptropicus   | Chlorokybush aptropicus chloroplast, comp product="large ribosomal RNA"                                                                                                                                                |         |
| GWKQY01A2Z7H    | 105      | 1       | 105   | 2E-37   | 0.94    | 0.94  | 210148086 | Eukaryota | Streptophyta     | Cornales           | Cornus kousa              | Cornus kousa clone C08B9 microsatellite seq satellite="microsatellite"                                                                                                                                                 |         |
| GWKQY01B1H3Q0   | 218      | 1       | 218   | 7E-45   | 0.83    | 0.83  | 13445170  | Eukaryota | Streptophyta     | Liliopsida         | Festuca arundinacea       | Festuca arundinacea genomic: FaS5 repeat el.rpt_type=DISPERSED                                                                                                                                                         |         |
| GWKQY01A3B8K7   | 442      | 1       | 442   | 0       | 0.93    | 0.93  | 21779916  | Eukaryota | Streptophyta     | Liliopsida         | Aegilops tauschii         | Aegilops tauschii leucine-rich-like protein ger product="leucine-rich-like protein"                                                                                                                                    |         |
| GWKQY01B19VXZ   | 480      | 1       | 480   | 2E-109  | 0.83    | 0.83  | 32307244  | Eukaryota | Streptophyta     | Liliopsida         | Aegilops tauschii         | Aegilops tauschii clone A9E-2 tandem repeat rpt_type=tandem                                                                                                                                                            |         |
| GWKQY01B0E6K    | 501      | 1       | 501   | 0       | 0.93    | 0.93  | 32994295  | Eukaryota | Streptophyta     | Liliopsida         | Oryza sativa              | Oryza sativa Japonica Group cDNA clone.002 mol_type="mRNA"                                                                                                                                                             |         |
| GWKQY01B0B8P0   | 536      | 1       | 536   | 0       | 0.92    | 0.92  | 32995694  | Eukaryota | Streptophyta     | Liliopsida         | Oryza sativa              | Oryza sativa Japonica Group cDNA clone.002 mol_type="mRNA"                                                                                                                                                             |         |
| GWKQY01A1YH1Q   | 243      | 1       | 243   | 3E-108  | 0.97    | 0.97  | 40849982  | Eukaryota | Streptophyta     | Liliopsida         | Triticum turgidum         | Triticum turgidum A genome HMW glutenin i mobile_element_type="transposon:LTR Retrotransposon, Erika_634M12-1p"                                                                                                        |         |
| GWKQY01A1X1MY   | 81       | 1       | 81    | 2E-36   | 0.95    | 0.95  | 7256311   | Eukaryota | Streptophyta     | Liliopsida         | Triticum aestivum         | Triticum aestivum clone wmc1621 microsat rpt_type=tandem                                                                                                                                                               |         |
| GWKQY01A1B7M7   | 526      | 1       | 526   | 0       | 0.98    | 0.98  | 99651997  | Eukaryota | Streptophyta     | Liliopsida         | Oryza sativa              | Oryza sativa Japonica Group genomic DNA, c mol_type="genomic DNA"; chromosome="9"                                                                                                                                      |         |
| GWKQY01B1H6M6   | 334      | 1       | 334   | 6E-63   | 0.81    | 0.83  | 102567891 | Eukaryota | Streptophyta     | Liliopsida         | Zea perennis              | Zea perennis mitochondrion, complete genom mol_type="genomic DNA"                                                                                                                                                      |         |
| GWKQY01A1QJW0   | 300      | 1       | 299   | 2E-113  | 0.92    | 0.92  | 111392331 | Eukaryota | Streptophyta     | Liliopsida         | Triticum monococcum       | Triticum monococcum subsp. aegilopoides c1 note="target site duplication"                                                                                                                                              |         |
| GWKQY01A1QJW6_2 | 181      | 1       | 181   | 5E-70   | 0.94    | 0.94  | 115392331 | Eukaryota | Streptophyta     | Liliopsida         | Triticum monococcum       | Triticum monococcum subsp. aegilopoides c3 mobile_element_type="transposon:truncated retrotransposon Wge1-2_TbBAC5-1p"                                                                                                 |         |
| GWKQY01B1WVC5   | 492      | 1       | 492   | 0       | 0.91    | 0.91  | 124007144 | Eukaryota | Streptophyta     | Liliopsida         | Triticum urartu           | Triticum urartu clone BAC 41C8 genomic seq mol_type="genomic DNA"; chromosome="6A"                                                                                                                                     |         |
| GWKQY01A1W7W    | 285      | 1       | 285   | 6E-131  | 0.97    | 0.97  | 148372275 | Eukaryota | Streptophyta     | Liliopsida         | Triticum monococcum       | Triticum monococcum subsp. aegilopoides c1 mol_type="genomic DNA"                                                                                                                                                      |         |
| GWKQY01B1Y5J2   | 492      | 1       | 492   | 0       | 0.99    | 0.99  | 148910867 | Eukaryota | Streptophyta     | Liliopsida         | Triticum turgidum         | Triticum turgidum overlapping clones BAC 32 mobile_element_type="retrotransposon:Gypsy"                                                                                                                                |         |
| GWKQY01A1Q4S7   | 348      | 1       | 348   | 3E-100  | 0.87    | 0.87  | 194131647 | Eukaryota | Streptophyta     | Liliopsida         | Triticum turgidum         | Triticum turgidum subsp. dicoccoides clones mobile_element_type="retrotransposon:Jel1_391M13_2                                                                                                                         |         |
| GWKQY01B1B7F    | 262      | 1       | 262   | 2E-106  | 0.94    | 0.94  | 194239308 | Eukaryota | Streptophyta     | Liliopsida         | Triticum aestivum         | Triticum aestivum 3A chromosome, clone BA mobile_element_type="transposon:RLG_Laura_TA3895C-1"                                                                                                                         |         |
| GWKQY01B1F6GK   | 432      | 1       | 432   | 0       | 1       | 1     | 209361311 | Eukaryota | Streptophyta     | Liliopsida         | Cos lacyma-jabi           | Cos lacyma-jabi chloroplast, complete geno product="16S ribosomal RNA"                                                                                                                                                 |         |
| GWKQY01A1QACWM  | 154      | 1       | 154   | 2E-13   | 0.77    | 0.77  | 212007811 | Eukaryota | Streptophyta     | Liliopsida         | Triticum aestivum         | Triticum aestivum clone BAC 502E09, comple mol_type="genomic DNA"                                                                                                                                                      |         |
| GWKQY01B1W1X8   | 383      | 1       | 383   | 0       | 0.99    | 0.99  | 255099160 | Eukaryota | Streptophyta     | Liliopsida         | Dendrocalamus latiflorus  | Dendrocalamus latiflorus chloroplast, complete product="23S ribosomal RNA"                                                                                                                                             |         |
| GWKQY01A1X1LX   | 127      | 1       | 127   | 2E-57   | 0.90    | 0.99  | 218619090 | Eukaryota | Streptophyta     | Apiales            | Daucus carota             | Daucus carota sativa, sativus clone BAC CD34 note="clone in pinstigBAC5 vector, end sequence from reverse end"                                                                                                         |         |
| GWKQY01B1VCS0   | 578      | 1       | 578   | 2E-123  | 0.83    | 0.83  | 88656961  | Eukaryota | Streptophyta     | Asterales          | Lactuca sativa            | Lactuca sativa cultivar Salinas chloroplast, cor product="23S ribosomal RNA"                                                                                                                                           |         |
| GWKQY01B1BNTY   | 300      | 273     | 300   | 0.021   | 1       | 1     | 259526388 | Eukaryota | Streptophyta     | Asterales          | Artemisia annua           | Artemisia annua strain Artemis Contig17 don't match anything, junk sequence                                                                                                                                            |         |
| GWKQY01A2ZBH8   | 232      | 1       | 232   | 1E-107  | 0.98    | 0.98  | 194132090 | Eukaryota | Streptophyta     | Mahales            | Gomphylus                 | Gomphylus bancanus chloroplast, partial gen product="23S ribosomal RNA"                                                                                                                                                |         |
| GWKQY01A1Y7P7   | 275      | 1       | 275   | 3E-139  | 0.99    | 0.99  | 170523360 | Eukaryota | Streptophyta     | Brassicales        | Carica papaya             | Carica papaya mitochondrion, complete gene product="26S ribosomal RNA"                                                                                                                                                 |         |
| GWKQY01A1A43T   | 198      | 1       | 198   | 1E-96   | 0.99    | 0.99  | 62149314  | Eukaryota | Streptophyta     | Caryophyllales     | Silene latifolia          | Silene latifolia chloroplast rp8b, rpCL1, rpCL2 mol_type="genomic DNA"; sex="male"                                                                                                                                     |         |
| GWKQY01A1A9BN   | 428      | 1       | 428   | 0       | 0.94    | 0.94  | 193789321 | Eukaryota | Streptophyta     | Fabales            | Trifolium subterraneum    | Trifolium subterraneum chloroplast, complet product="23S ribosomal RNA"                                                                                                                                                |         |
| GWKQY01A1A746   | 335      | 1       | 335   | 2E-152  | 0.96    | 0.96  | 293338822 | Eukaryota | Streptophyta     | Fabales            | Lathyrus sativus          | Lathyrus sativus cultivar Clericha Marichlar product="16S ribosomal RNA"                                                                                                                                               |         |
| GWKQY01A1ATFQ   | 514      | 1       | 514   | 0       | 0.93    | 0.93  | 210143279 | Eukaryota | Streptophyta     | Fabales            | Glycine max               | Glycine max cDNA, clone: GML01-21-C10 mol_type="mRNA"                                                                                                                                                                  |         |
| GWKQY01B1G6G0   | 394      | 1       | 394   | 1E-159  | 0.93    | 0.93  | 218115405 | Eukaryota | Streptophyta     | Fabales            | Medicago truncatula       | Medicago truncatula chromosome 4 clone m mol_type="genomic DNA"                                                                                                                                                        |         |
| GWKQY01A1A1C4   | 345      | 1       | 345   | 4E-178  | 0.99    | 0.99  | 329124647 | Eukaryota | Streptophyta     | Solanales          | Solanum                   | Solanum tuberosum isolate DM1-3-516-R44 c product="16S ribosomal RNA"                                                                                                                                                  |         |
| GWKQY01B0B3Q3   | 467      | 1       | 467   | 0       | 0.96    | 0.96  | 147776538 | Eukaryota | Streptophyta     | Vitales            | Vitis vinifera            | Vitis vinifera contig VV780U60272.3, whole g mol_type="genomic DNA"                                                                                                                                                    |         |
| GWKQY01B1Y7ZM   | 52       | 1       | 52    | 5E-13   | 0.94    | 0.94  | 147826096 | Eukaryota | Streptophyta     | Vitales            | Vitis vinifera            | Vitis vinifera contig VV780U60272.2, whole g mol_type="genomic DNA"                                                                                                                                                    |         |
| GWKQY01B1Y7U0   | 537      | 1       | 537   | 0       | 0.99    | 0.99  | 236764707 | Eukaryota | Streptophyta     | Vitales            | Vitis vinifera            | Vitis vinifera strain PN4024 mitochondrion, product="26S ribosomal RNA"                                                                                                                                                |         |