

**Table S1: 5mers, 9mers, and 13mers evaluated**

Name	Sequence	Hydrophobicity	Organism†
5mer1	CHKWD	20%	<p><b><u>Eukarya</u></b>  <b>Fungi:</b>  <i>Puccinia graminis</i> (1-phosphatidylinositol-3-phosphate 5-kinase)  <i>Naumovozya dairenensis</i> (hypothetical protein)  <b>Plantae:</b>  <i>Melampsora larici-populina</i> (hypothetical protein)  <i>Brachyscome latisquamea</i> (maturase)  <i>Brachyscome ciliocarpa</i> (maturase)  <b>Animalia:</b> <i>Amphimedon queenslandica</i> (hypothetical protein)  <b>Bacteria</b>  <i>Opitutaceae bacterium</i> (hypothetical protein)  <i>Peptostreptococcus anaerobius</i> (transcriptional regulatory, LytR family)  <i>Denitrovibrio acetiphilus</i> (hypothetical protein)  <b>Archaea</b>  <i>Nitrosopumilus maritimus</i> (hypothetical protein)  <i>Ramlibacter tataouinensis</i> (hypothetical protein)  <i>Candidatus Nitrosoarchaeum limnia</i> (phospho-2-dehydro-3-deoxyheptonate aldolase)</p>
5mer2	WHKCE	20%	<p><b><u>Eukarya</u></b>  <b>Fungi:</b> <i>Pyrenophora tritici-repentis</i> (predicted protein)  <i>Metarhizium anisopliae</i> (hypothetical protein)  <b>Chromalveolata:</b> <i>Blastocystis hominis</i> (unnamed protein product)  <b>Animalia:</b> <i>Nematostella vectensis</i> (predicted protein)  <i>Ornithorhynchus anatinus</i> (Platypus, hypothetical protein)  <i>Drosophila yakuba</i> (GE10666)  <b>Plantae:</b> <i>Gluconacetobacter diazotrophicus</i> (hypothetical protein)  <i>Populus trichocarpa</i> (predicted protein)  <b>Unranked:</b> <i>Monosiga brevicollis</i> (hypothetical protein)  <b>Bacteria</b>  <i>Gluconacetobacter diazotrophicus</i> (Superfamily I DNA and RNA helicase and helicase subunits-like protein)  <i>Agrobacterium tumefaciens</i> (acetyltransferase)  <i>Capnocytophaga</i> sp (hypothetical protein)  <i>Clostridium hathewayi</i> (hypothetical protein)</p>
5mer3	CKWRC	20%	<p><b><u>Eukarya</u></b>  <i>Sporisorium reilianum</i> (related to high-affinity nickel transport protein nic1)  <i>Pyrenophora tritici-repentis</i> (hypothetical protein)  <i>Neurospora tetrasperma</i> (hypothetical protein)  <b>Animalia:</b> <i>Ciona intestinalis</i> (predicted similar to Sodium channel protein type 5 subunit alpha)  <i>Brugia malayi</i> (Ubiquitin carboxyl-terminal hydrolase family protein)  <i>Loa loa</i> (hypothetical protein)  <i>Caenorhabditis remanei</i> (hypothetical protein)  <i>Osmerus mordax</i> (Centaurin-alpha 2)  <i>Sabellaria alveolata</i> (fructose-bisphosphate aldolase)</p>

			<i>Limnodrilus</i> sp (fructose-bisphosphate aldolase)
5mer3			<b>Plantae:</b> <i>Vitis vinifera</i> (hypothetical protein) <i>Selaginella moellendorffii</i> (hypothetical protein) <i>Physcomitrella patens subsp. patens</i> (predicted protein) <i>Volvox carteri</i> f. <i>nagariensis</i> (hypothetical protein) <b>Bacteria</b> <i>Neisseria bacilliformis</i> (hypothetical protein)
5mer4	KWCEC	20%	<b>Eukarya</b> <b>Excavata:</b> <i>Leishmania mexicana</i> (hypothetical protein) <b>Chromalveolata:</b> <i>Cryptosporidium muris</i> (hypothetical protein) <b>Plantae :</b> <i>Physcomitrella patens subsp. patens</i> , <i>Populus trichocarpa</i> (conserved hypothetical protein) <b>Bacteria</b> <i>Bathycoccus</i> sp., <i>Paenibacillus curdlanolyticus</i> , <i>Clostridium nexile</i> , <i>Clostridium ljungdahlii</i> , <i>Bacteroides pectinophilus</i> , <i>Rhodococcus erythropolis</i> , <i>Veillonella</i> sp., <i>Ahrensia</i> sp. (conserved hypothetical protein) <i>Atopobium rima</i> e (methyltransferase)
5mer5	DCWMD	40%	<b>Eukarya</b> <b>Chromalveolata:</b> <i>Emiliana huxleyi</i> (ankyrin-1-like protein) <b>Fungi:</b> <i>Candida tenuis</i> (hypothetical protein) <i>Ajellomyces</i> sp. (C6 transcription factor) <i>Paracoccidioides brasiliensis</i> (hypothetical protein) <i>Trichophyton</i> sp. (C6 transcription factor) <i>Arthroderma</i> sp (C6 transcription factor) <b>Animalia:</b> <i>Pediculus humanus corporis</i> (hypothetical protein) <i>Daphnia pulex</i> (hypothetical protein) <i>Taeniopygia guttata</i> (hypothetical protein) <b>Plantae:</b> <i>Chlorella variabilis</i> (hypothetical protein) <b>Bacteria</b> <i>Asticcacaulis biprosthecum</i> (pectinesterase family protein) <i>Bacteroides</i> sp. (Glycoside hydrolase) <i>Clostridium</i> sp (hypothetical protein) <i>Bacillus megaterium</i> (allantoate amidohydrolase) <i>Parvibaculum lavamentivorans</i> (beta-lactamase-like protein) <i>Agrobacterium</i> sp. (competence protein F) <i>Campylobacter gracilis</i> (hypothetical protein) <i>Bordetella petrii</i> (glutathione S-transferase) <b>Virus</b> Duck reovirus GZ/CHN/2007 (S3 protein) Grapevine leafroll-associated virus 7 (heat shock protein 70)
5mer6	KYMCW	60%	<b>Eukarya</b> <i>Ciona intestinalis</i> (hypothetical protein) <b>Bacteria</b> <i>Desulfobacterium</i> sp. (hypothetical protein)
9mer1	CWKCWCMFE	44%	Not found in any proteome
9mer3	WNWCMHWDC	44%	Not found in any proteome
9mer4	WHWCMMCWDC	56%	Not found in any proteome
13mer1	HEHWCMWHCCMI	46%	Not found in any proteome
13mer3	HMMCHWMCWCDMH	46%	Not found in any proteome
13mer4	CHMMCHWMWCCMD	46%	Not found in any proteome