

GQ478431_ORF5	-----MKRIITLF--FIFIFAISCSKTQTQNTQTEVS-----SADNS	33
GQ478432_ORF6	-----MKRIITLF--FIFIFAISCSKTQNASQASS-----AADNS	33
fig 6666666.171271.peg.763	MTYNLKTKNFFIKTLFQTFIFL---IFLLIISCTKETK-KIKDELTVNLGYELQSIDPA	56
fig 6666666.171271.peg.2613	-----MKKTIYIIF---LILSIIISCNMNNKTTSNNEISISLGGPEPKTLDPT	43
fig 6666666.171271.peg.2276	-----MKKI-ILIVSFIFIYITISCAKQNS-IVNEAISINMGPEPKTLDPT	43
fig 6666666.171271.peg.2182	-----MFKKIITAI--LLILLIISCSKKEI-ITNNEIVINLAPEPLTMDTT	43
GQ478428_ORF2	-----MFKKIITAI--LLILLIISCSKKEI-ITNNEIVINLAPEPLTMDPT	43
fig 6666666.171271.peg.487	-----MKKLFAIL---SCIFILACSNNI--NNKEVIKVTVGAEPQSIDPS	40
GQ478429_ORF3	-----MKRIFAVL---SCIFILACSNNI--NNNEVIKVTVGAEPQSIDPS	40
fig 6666666.171271.peg.189	-----MKKYIIL---ILFFIIISCSKNN---TNNSLIVNIGPEPKTLDPT	38
GQ478427_ORF1	-----MKKYIIL---ILFFIIISCSKNN---TNNSLIVNIGPEPKTLDPT	38
fig 6666666.171271.peg.894	-----MKNIMRLLFFIFLLS-IGLIITSCSSKNK-NQGEGLFINVGPEPKTIDPT	47
fig 6666666.171271.peg.2198	-----MKHLTLTLLLSVLMIVSISCGGNK--TDEGAIYINVGPEPKTIDPA	43
fig 6666666.171271.peg.892	-----MYKRFLILFFSMIIFVVLSCGGKTNSLDDKTIVINAGPQPKTIDPG	46
GQ478430_ORF4	-----	0

GQ478431_ORF5	LQKVKENGKLVGLGLDDTLAPMGFRDDNGEIVGFDIDLAREVANRLGVEL-EAKPIDWSSA	92
GQ478432_ORF6	LQKVKESGKLVGLGLDDTFAPMGFRDEEGNIVGFDIDLAKEVANRLGVEL-VAKPVDWSSV	92
fig 6666666.171271.peg.763	INDETY---GFIYINHDFEGLLTKDINGKIVGGSSDKWEISEDKLYTFHIREDAKWSGD	113
fig 6666666.171271.peg.2613	LNSLSF---GSIYMIHFFEGLTKKDKNDEVTAASMAKSWDISEDGLTYTFYLRDDAKWSGD	100
fig 6666666.171271.peg.2276	LNSLGS---ASCYILHAFEGLTKMDSNNTVQPGIAERWDISDDGLTYTFHLRTNALWSGD	100
fig 6666666.171271.peg.2182	LNTDNL---TMIYILHAFEGLTKKDANNKIIGGAAETWDINKEGNIYTFHLRTNAKWSGD	100
GQ478428_ORF2	LNTDNL---TMIYILHAFEGLTKKDANNKIIGGVAETWDINKEGNIYTFHLRTNAKWSGD	100
fig 6666666.171271.peg.487	YLSAID---SMIYAVHVFEGLVTKDKEGNIVGGVSESWEVSPDGLNIIFHRLDNAKWSGD	97
GQ478429_ORF3	YLSAID---SMIYAVHVFEGLVTKDKEGNIVGGVAESWEVSPDGLNIIFHRLDNAKWSGD	97
fig 6666666.171271.peg.189	YTETYD---GSAYIAHTFEGLTSRDKYDKLIIPAVAESWDISEDYLTYYTFYLRDNAKWSGD	95
GQ478427_ORF1	YTETYD---GSAYIAHTFEGLTSRDKDDKLIIPAVAESWDISEDYLTYYTFYLRDNAKWSGD	95
fig 6666666.171271.peg.894	LNIASD---ASVYIIHAFEGLTTRNKEGVLVGGAAESWDISEDGLTYTFHLRTNGKWSGD	104
fig 6666666.171271.peg.2198	LNSTVD---GSIYIQHAFEGLATRDKDNKIIPGVAESWDISEDGLTYTFHIRENAKWSGD	100
fig 6666666.171271.peg.892	LNTALD---ACYVVIHAFEGLTTKSKEGKIVEGVAENWEILDDGSRYIFHLRTNAKWSGD	103
GQ478430_ORF4	-----	0

GQ478431_ORF5	ILSLKKGVDVDVIWNGFAVNESRQOVNFTKPYLYNRLMIAKYSRDRDDINSKEELKGKIVG	152
GQ478432_ORF6	VLSLKKGVDVDVWNGFAINETRKEQVNFTKPYLNNRLMIVKNADRDDIKTKDDLKGIKLG	152
fig 6666666.171271.peg.763	K-KLTADDFVYSYRRVVDPK-----TASPIAYLM-YYIKNAK-----D	149
fig 6666666.171271.peg.2613	E-KVKAQDFEYALKRAADPK-----TAATYSHML-NVVKNKS-----L	136
fig 6666666.171271.peg.2276	T-NVKAKDFEYSWKRSVNPS-----VAAEYAYMM-EIVQNAK-----E	136
fig 6666666.171271.peg.2182	K-EVKAEDFVYTWRRRAVDPK-----TANKYSYYF-EVIKNAK-----E	136
GQ478428_ORF2	K-EVKAEDFVYTWRRRAVDPK-----TANKYSYYF-EVIKNAK-----E	136
fig 6666666.171271.peg.487	K-KITADEFVYSFRRLVDPN-----TASSYGFLA-SPIKNAD-----K	133

GQ478429_ORF3	K-KITADEFVYSFRRLVDPN-----TASSYGFLA-SPIKNAD-----K	133
fig 6666666.171271.peg.189	K-TVTANDFVYAWQRAVDPK-----VASSVSYQL-NPIKNAS-----K	131
GQ478427_ORF1	K-TVTANDFVYAWQRAVDPK-----VASSVSYQL-NPIKNAS-----K	131
fig 6666666.171271.peg.894	T-PLTSKDFVYSWKRAVDPK-----TASEYSYQF-EPVKNAM-----A	140
fig 6666666.171271.peg.2198	K-KITAEDFVYSWQRAVDPI-----TASEYAYQF-EPVLNAM-----D	136
fig 6666666.171271.peg.892	K-AVVADDFVYAWRRVVDPA-----VGSQYSFQH-EPVKNAK-----A	139
GQ478430_ORF4	-----	0

GQ478431_ORF5	VQSGSSNYETLVNDPVSKEIKEIRQYDSYVNAFLDLE-----	189
GQ478432_ORF6	VQTGSSNYESLEADPI SKETKEIRQYDIYANAFDLE-----	189
fig 6666666.171271.peg.763	INIGKKQIDTLGVTATDKNTLTIEL-ENPTLYFEDILASGGCYVPVREDIINKYGDDWTW	208
fig 6666666.171271.peg.2613	VISGKTNIDALGVKAIDDSTLEIVL-ENPTPYFLEYLSVSSAYFPVRKDI VEKYGDDWSR	195
fig 6666666.171271.peg.2276	INEGTMDYNELAVKAIDDYTFEVTI-ANSPYFLEFIASGIFMPLREDIINTYGDEWTL	195
fig 6666666.171271.peg.2182	VISGEKTVEELGVKAIDDYTFEVEL-NSPTAYFLELAA-YPPFYVPVREDIINEYGDKWTL	194
GQ478428_ORF2	VISGEKTVEELGVKAIDDYTFEVEL-NSPTAYFLELAA-YPPFYVPVREDIINKYGDKWTL	194
fig 6666666.171271.peg.487	IMAGKLRKEQLGIEAIDEKTLLIK-ETPTAYFLELFC-IPFSPLRADYIEDN-EKWTF	190
GQ478429_ORF3	IMAGKLRKEELGIEAIDEKTLLIK-ETPTAYFLELFC-IPFSPLRAYYIEDN-EKWTF	190
fig 6666666.171271.peg.189	I IKGELSLTELGVKAIDDYTLEVNL-ESPTAYFLDLLS-ITIYSPLRKDIIEKYGDAWTQ	189
GQ478427_ORF1	I IKGELPLTELGVKAIDDYTLEVNL-ESPTAYFLDLLS-ITIYSPLRKDIIEKYGDAWTQ	189
fig 6666666.171271.peg.894	INASKMELDTLGIRAIDDYTLEVTL-EKPTAYFLELTA-FPTFYPIREDIIEKYGDTWTI	198
fig 6666666.171271.peg.2198	INSGKKPVTDLGIKAIDENTLEVKL-NAPTAYFLELAA-FPTFYVPVRKDIIEENKDNWTL	194
fig 6666666.171271.peg.892	ITAGEMPVDSLGIKAIDDYTLEVTL-EAPTAYFLDLVA-FTFYFPVRKDIIEKYGDSWSL	197
GQ478430_ORF4	-----	0

GQ478431_ORF5	-----AKRIDAVIVDEIVARYYISK----ENADFTLLEDKPITSQYLSVGL	231
GQ478432_ORF6	-----AKRIDAVIVDEIVARYYISK----ENADFELLEYPITAKSFGVGV	231
fig 6666666.171271.peg.763	NSESYIGNGAYKMTERKPDDELIAFELNTNYWDYKNQVAKKINFVLI ADEYI SLNAVRTGD	268
fig 6666666.171271.peg.2613	NPETYIVNGAYVMTERKTDEKIVMEVNTNYDKE SIVAKKINVI IMSDSNTSLAAIKRGD	255
fig 6666666.171271.peg.2276	KPETYIVNGAYTMTERLADEKIVFSANKNYNPNEQVAKKINFVLMSEPNTAMSGIKSSS	255
fig 6666666.171271.peg.2182	KPATYIGNGAFKMTERNFDKSIILERNNTNYWNNENTKPNKLTFFLLMEEPNTSLAGVINNS	254
GQ478428_ORF2	KPATYIGNGAFKMTERNFDKSIILERNNTNYWNNENIKPNKLTFFLLMEEPNTSLAGVLNNS	254
fig 6666666.171271.peg.487	YPNTYIGNGPYKMIERKVDESISLELNTNYWKNKDNMVANRIDFIMLSDVSTAYPALKEGS	250
GQ478429_ORF3	YPNTYIGNGPYKMIERKVDESISLELNTNYWKNENMVANRIDFIMLSDVSTAYPALKEGS	250
fig 6666666.171271.peg.189	--EEYIGNGPPFFVKEHINNDKIVMEKNTNYWKNKETIIPQTITFSLLDNPNTIVASIQDGT	247
GQ478427_ORF1	--EEYIGNGPPFFVKEHINNDKIMEKNTNYWKNKETIIPQTITFSLLDNPNTIVASIQDGT	247
fig 6666666.171271.peg.894	NPESYIGNGSYVMIERNIDNNIIMVKNTNYWDYENLVPEKITFVLMNPTASVAGIKEGS	258
fig 6666666.171271.peg.2198	SPDITYIGNGPFPTLVERRTDDRLVMIKNTNYWNSENI IPEKLVF ILMQNGTAAVAGIKEGS	254
fig 6666666.171271.peg.892	NVESYIGNGPPFVTEINQDESIMVKNTNYWAFNDVVPEKLFVLMQNETASVAGIKEGS	257
GQ478430_ORF4	-----MVKNTNYWAFNDVVPEKLFVLMQNETASVAGIKEGS	37

.*: . : . :: :

GQ478431_ORF5	-----RKTDT-----LLNAIDKALDDMRADGK-AAEISTKWF	263
GQ478432_ORF6	-----RKTDT-----LLNAIDKALDDMRDDGK-AAEISTKWF	263
fig 6666666.171271.peg.763	VDFSINAPPIGEIESLIKENLMAVSDIIGVYYLDLNNKDKTSLDKRVRKALSIAIDRNYI	328
fig 6666666.171271.peg.2613	IQFSVIEAPLGEISSLIKENYILQEPAYGIYFLEINSKKGVLTKNIRKALALAFDRNYI	315
fig 6666666.171271.peg.2276	IHFSALEPPSSEIASLAKADNLLVENNAIGTYFIELNITNNALRDKRVRQALSIAIDRNYL	315
fig 6666666.171271.peg.2182	IHFSALEPPSSEIASLAKADNLLVENNAIGTYFIELNITNNALRDKRVRQALSIAIDRNYL	313
GQ478428_ORF2	IHFSALEPPSSEIASLAKADNLLVENNAIGTYFIELNITNNALRDKRVRQALSIAIDRNYL	313
fig 6666666.171271.peg.487	LHYSS-RIQNNDIELLRKEGYLVITPSLGTAYYAVNNTNEVLKDKRVRKALALSIDRNYI	309
GQ478429_ORF3	LHYSS-RIQNNDIELLRKEGYLVITPSLGTAYYAVNNTNEVLKDKRVRKALALSIDRNYI	309
fig 6666666.171271.peg.189	LLFGS-NPPLQDIPKLTTEEGYIDYVPLLGIYFLSINTTNEVLKKNVRKALSIAIDRNYI	306
GQ478427_ORF1	LLFGS-NPPLQDIPKLTTEEGYIDYVPLLGIYFLSINTTNEVLKKNVRKALSIAIDRNYI	306
fig 6666666.171271.peg.894	LHFSD-RVPAQDIDTLKAEGLLQMVKRLGTYYYAVNHTNEVLKDKRVRKALSIAIDRNYI	317
fig 6666666.171271.peg.2198	IHFSGN-NPPLQDIPKLTTEEGYIDYVPLLGIYFLSINTTNEVLKDKRVRKALSIAIDRNYL	313
fig 6666666.171271.peg.892	IDFSR-ILPTQDIPKLTTEEGYIDYVPLLGIYFLSINTTNEVLKDKRVRKALSIAIDRNYI	316
GQ478430_ORF4	IDFSR-ILPTQDIPKLTTEEGYIDYVPLLGIYFLSINTTNEVLKDKRVRKALSIAIDRNYI	96
	: . * . * : . : . . :	

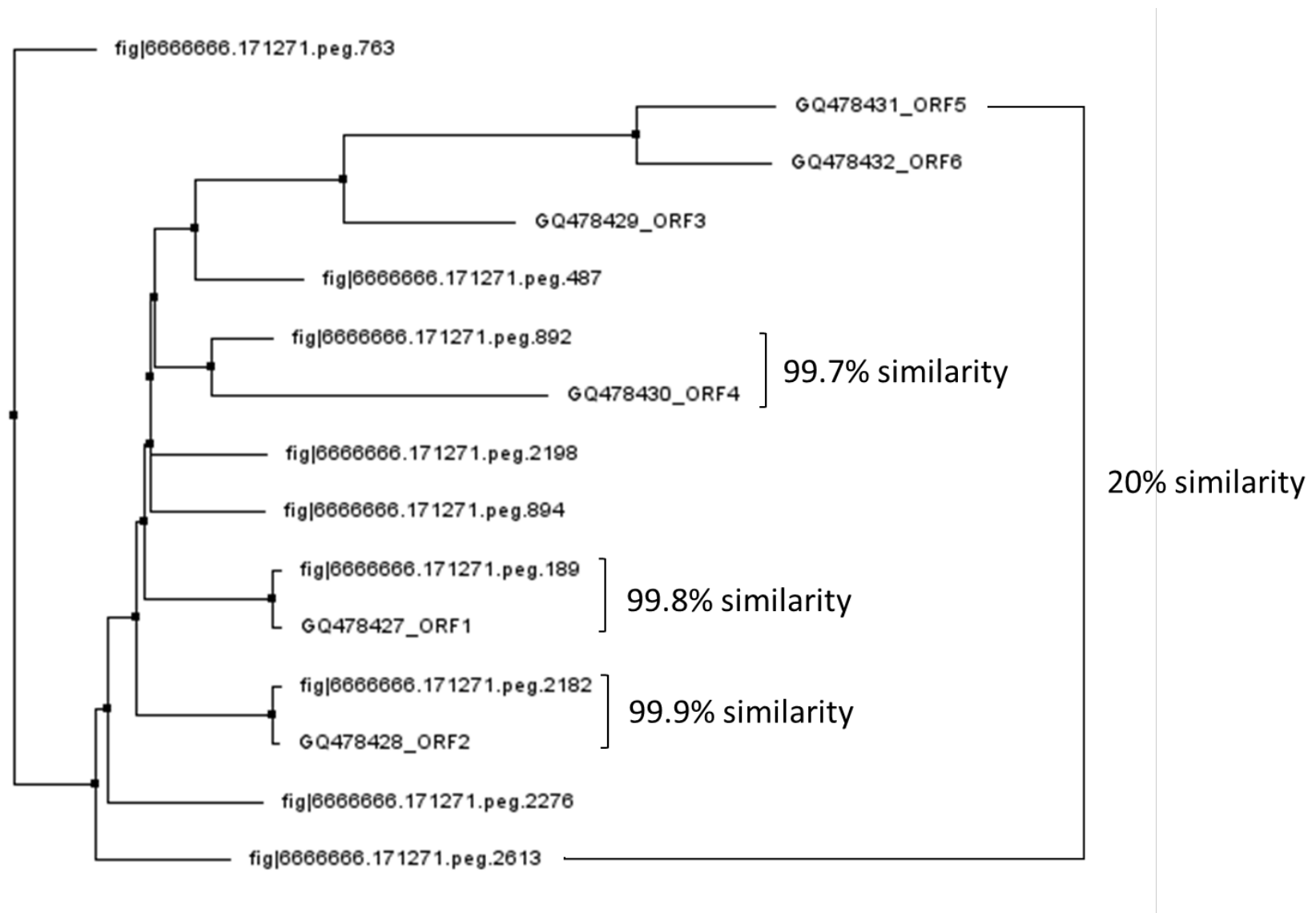
GQ478431_ORF5	GKDVFIK-----	271
GQ478432_ORF6	GKDVFVRG-----	271
fig 6666666.171271.peg.763	VSNIGHGKLIAAESFVPPVV-KGLEKSFREESSNFII--ANNYSNNIEAKKLLAEAGYP	385
fig 6666666.171271.peg.2613	ISNITKMNQTPAYAFVVPYGMKDGDKDFRENGSNYLD--LGSYDANIKEAKRLMELSGYT	373
fig 6666666.171271.peg.2276	VLNVTKGGQTPAGAFIPPTV-KGSNSTFRVENQEYIN--NKTYLANVEKAKRLMAEAGYP	372
fig 6666666.171271.peg.2182	VKSITRCGERPAGSLVPYGI-NDIDGDFRTKGGEYII--SSNYQKNIEEAKKLLSQAGYE	370
GQ478428_ORF2	VKSITRCGERPAGSLVPYGI-NDIDGDFRTKGGEYII--SSNYQKNIEEAKKLLSQAGYE	370
fig 6666666.171271.peg.487	VENITKGGVVPAGAFIPFGL-KDVSGDFRENGGNYFSVSKEDYESNIKEAKRLLAEAGYS	368
GQ478429_ORF3	VENITKGGVVPAGAFIPFGL-KDVSGDFRENGGNYFSVSKRL-----	350
fig 6666666.171271.peg.189	VEKVTKGNEIPAAALVPYSV-FDIDGSFREKGGNYFSVNENDYQKNIEEARRLLAEAGYS	365
GQ478427_ORF1	VEKVTKGNEIPAAALVPYSV-FDIDGSFREKGGNYFSVNENDYQKNIEEARRLLAEAGYS	365
fig 6666666.171271.peg.894	VDNIVKAG-VPAGAFVPPFGA-PDVNGDFREVGGDYISVKPEDYQKNIEEAKKLMEEAGYT	375
fig 6666666.171271.peg.2198	VEQVTRGGQLPASAWVPSGV-NDVNGDFREVGGDYIGIKPEDYKKNVVEAKRLLAEAGYP	372
fig 6666666.171271.peg.892	VEQVTRGGGETPANAFVVPYGI-NDAEGSFRENGGGYFDISQEGYKKNVVEAKKLMMAEAGYP	375
GQ478430_ORF4	VEQVTRGGGETPADAFVVPYGI-NDAEGSFRENGGGYFDISQEGYKKNVVEAKKLMMAEAGYP	155
	. :	

GQ478431_ORF5	-----	271
GQ478432_ORF6	-----	271
fig 6666666.171271.peg.763	NGENFPILVVKVS-SGFYTTVLEAIQDMWKNLNIDVVRTEESKITLPPFRQ-SGKYQIA	443
fig 6666666.171271.peg.2613	NGDNFPVLEIRTT-PGYFTLICEAMQEMYKENLIGIDVTIKSEEYNETFQAMV-EKNYDLA	431
fig 6666666.171271.peg.2276	NGKDFPVLELKVS-PGIYVLLGETIQQMWRDNLNVYIKLLQEEFPVTLQTLV-DKNYEMA	430
fig 6666666.171271.peg.2182	NGRNFPVIDLLIATREFDINIADAVQSMLKENLNIDVRIVKHEWASYLQNMV-DRNFDLA	429
GQ478428_ORF2	NCRNFPVIDLLIATREFDINIADAVQSMLKENLNIDVRIVKHEWASYLQNMV-DRNFDLA	429

fig 6666666.171271.peg.487	NGVNFVPVLEFKTN-PGAGVTIAEAVQQMWKENLNVDMSITQEEWSVFQKNRQ-TRNYTVC	426
GQ478429_ORF3	-----	350
fig 6666666.171271.peg.189	NGSNFPVLEYSVE-SQSSLNIFEAIQQMWKENLNIDTRVNQLEWGVFMQTTRGDKNFEIA	424
GQ478427_ORF1	NGSNFPVLEYSVE-SQSSLNIFEAIQQMWKENLNIDTRVNQLEWGVFMQTTRGDKNFEIA	424
fig 6666666.171271.peg.894	NGEGFPVLQFLVD-SNREIPTFEAVQQMWKEHLGIDTSVSQQEWAVFLQTLTYTDKNYVMG	434
fig 6666666.171271.peg.2198	NGDGFVPVIEFKSN-SGEHIQIFEAVQQMWKENLNIDSTIAQEEWATFQDTRQ-NKNYMIA	430
fig 6666666.171271.peg.892	NGEGFPVFEFKAD-PGFHISIFEAVQQMWKENLGIDTKIVQEEWAVFLQTRY-DKNLTMC	433
GQ478430_ORF4	NGEGFPVFEFKAD-PGFHISIFEAVQQMWKENLGIDTKIVQEEWAVFLQTRY-DKNLTMC	213

GQ478431_ORF5	-----	271
GQ478432_ORF6	-----	271
fig 6666666.171271.peg.763	RTSWTGDYNDPLTMLQIMTSESDINYSGFSNERYDYLINFATTSTNEKDRMEALKEAEAI	503
fig 6666666.171271.peg.2613	RTGWTADYSDPLAMISFFSEVSAVNNSGFSREFNDLLKFASSTTNTDERTKALHKAEDL	491
fig 6666666.171271.peg.2276	RMGWTGDYNDPMTMLEIMLSYSGINHTGFSNADFDNLIINIAKTSADNDVRMNAMKEAEAI	490
fig 6666666.171271.peg.2182	VYLWYADYNDPINFNLNIFKSDAPNNYGSYSNKTFFDDYIDIASTNKNNDIRMPALHSAENI	489
GQ478428_ORF2	VYLWYADYNDPINFNLNIFKSDAPNNYGSYSNKAFFDDYIDIASTNKNNDIRMPALHSAENI	489
fig 6666666.171271.peg.487	RADWIGDYLDPMTFQAQLFTSTTSAGNRVGYSNINYSLSIKDAQSTIDNNRRMNMHIAEDM	486
GQ478429_ORF3	-----	350
fig 6666666.171271.peg.189	RSGWMGDYNDPMTFFLDTFLSYSPQNTGSYYNKEYEQLIQSALTNGDKIYRMQTLHKAEDV	484
GQ478427_ORF1	RSGWMGDYNDPMTFFLDTFLSYSPQNTGSYYNKEYEQLIQSALTNGDKISRMQTLHKAEDV	484
fig 6666666.171271.peg.894	RSGWTADYNDAMTFLGMFLSYSPQNHAFNTNKEYDNLQIAMNTIDQNIIRMDAMHKAEKI	494
fig 6666666.171271.peg.2198	RHGWIADYNDPMSFLGVFLSYSPVQNNGGYSNKNYDDKLLAMSTIDQDVRMKAMHEAEDI	490
fig 6666666.171271.peg.892	RGGWFGDFNDPINFSLYTSYSPNNYSSYSNAQYDAYIQTALTTGDQKVRMEAMHKADEL	493
GQ478430_ORF4	RGGWFGDFNDPINFSLYTSYSPNNYSSYSNAQYDAYIQTALTTGDQKVRMEAMHKADEL	273

GQ478431_ORF5	-----	271
GQ478432_ORF6	-----	271
fig 6666666.171271.peg.763	LF-EEMPIIPFIYRTDFLVVNPVKLNKNIIDDPGRYRFNYAYIEE	546
fig 6666666.171271.peg.2613	IF-DYMPIIPIIYRMDPFMINPKLKGAI FNPLGRYRFHYAYLEK	534
fig 6666666.171271.peg.2276	LM-DEMPIIPLYRFV--YGKPNTKKCCIKPFRKT-----	522
fig 6666666.171271.peg.2182	FM-NDNAIIPYFYSEALLVSPKLGVEYDSQGLYRFFNAYLE-	531
GQ478428_ORF2	FM-NDNAIIPYFYSEALLVSPKLGVEYDSQGLYRFFNAYLE-	531
fig 6666666.171271.peg.487	LIGDDMALIPLYYYTAPSMKSPKLDVVDVDTLEIRRFFYSYLQ-	529
GQ478429_ORF3	-----	350
fig 6666666.171271.peg.189	LM-EDMPFIPLYFYNRVILVRPELKDVMISNVIPRFFFYAYVEK	527
GQ478427_ORF1	LM-EDMPFIPLYFYNRVILVRPELKDVMISNVIPRFFFYAYVEK	527
fig 6666666.171271.peg.894	FM-DNDV IIPLYYYATP TLVSPKLGVVYDNLAKHKFSYAYIEN	537
fig 6666666.171271.peg.2198	LM-NDMGLIPLYFYTDPIIMISKKLSGVI FDPLGAHKFYAKLAK	533
fig 6666666.171271.peg.892	LV-NSFAIIPMYFYTEPLLVNPKLKD VYYDALSMHKFTYASKEQ	536
GQ478430_ORF4	LV-NSFAIIPMYFYTEPLLVNPKLKD VYYDALSMHKFTYASKEQ	316



Supplementary Figure 1. Clustal alignment and dendrogram of the OppA family proteins identified by RV and ORFs 1-6 oligopeptide-binding proteins from study of Movahedi and Hampson, 2010