University of Nebraska - Lincoln
DigitalCommons@University of Nebraska - Lincoln

Virology Papers | Virology, Nebraska Center for

2-20-2007

Supplementary Data for "Sequence and annotation of the 314-kb MT325 and the 321-kb FR483 viruses that infect *Chlorella* Pbi": Appendix A: Gene Names m002R through m843L

Lisa A. Fitzgerald
University of Nebraska-Lincoln, fitz918@msn.com

Michael V. Graves
University of Massachusetts–Lowell, Michael_Graves@uml.edu

Xiao Li
University of Massachusetts–Lowell

Tamara Feldblyum
The Institute for Genomic Research, Rockville, MD

James Hartigan
Agencourt Bioscience Corporation, Beverly, MA

See next page for additional authors

Follow this and additional works at: http://digitalcommons.unl.edu/virologypub

Part of the Virology Commons

Fitzgerald, Lisa A.; Graves, Michael V.; Li, Xiao; Feldblyum, Tamara; Hartigan, James; and Van Etten, James L., "Supplementary Data for "Sequence and annotation of the 314-kb MT325 and the 321-kb FR483 viruses that infect *Chlorella* Pbi": Appendix A: Gene Names m002R through m843L" (2007). *Virology Papers*. Paper 2.
http://digitalcommons.unl.edu/virologypub/2

This Article is brought to you for free and open access by the Virology, Nebraska Center for at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Virology Papers by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

Authors
Lisa A. Fitzgerald, Michael V. Graves, Xiao Li, Tamara Feldblyum, James Hartigan, and James L. Van Etten

This article is available at DigitalCommons@University of Nebraska - Lincoln: http://digitalcommons.unl.edu/virologypub/2

Published in *Virology* 358:2 (February 20, 2007), pp. 459–471. doi:10.1016/j.virol.2006.08.034 Copyright © 2006 Elsevier Inc. Used by permission. Submitted July 27, 2006; revised August 18, 2006; accepted August 23, 2006; published online October 4, 2006. Supplementary data originally published online at doi:10.1016/j.virol.2006.08.034.

SUPPLEMENTARY DATA FOR

Sequence and annotation of the 314-kb MT325 and the 321-kb FR483 viruses that infect *Chlorella* Pbi

Lisa A. Fitzgerald[a], Michael V. Graves[b], Xiao Li[b], Tamara Feldblyum[c], James Hartigan[d], and James L. Van Etten[e, f, *]

[a]Department of Chemistry, University of Nebraska–Lincoln, Lincoln, NE 68588-0304
[b]Department of Biological Sciences, University of Massachusetts–Lowell, Lowell, MA 01854
[c]The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850
[d]Agencourt Bioscience Corporation, 500 Cummings Center, Suite 2450, Beverly, MA 01915
[e]Deparment of Plant Pathology, University of Nebraska–Lincoln, Lincoln, NE 68583-0722
[f]Nebraska Center for Virology, University of Nebraska, Lincoln, NE 68588-0666
*Corresponding author. Email: jvanetten@unlnotes.unl.edu

Abstract: Viruses MT325 and FR483, members of the family Phycodnaviridae, genus *Chlorovirus*, infect the fresh water, unicellular, eukaryotic, chlorella-like green alga, *Chlorella* Pbi. The 314,335-bp genome of MT325 and the 321,240-bp genome of FR483 are the first viruses that infect *Chlorella* Pbi to have their genomes sequenced and annotated. Furthermore, these genomes are the two smallest chlorella virus genomes sequenced to date, MT325 has 331 putative protein-encoding and 10 tRNA-encoding genes and FR483 has 335 putative protein-encoding and 9 tRNA-encoding genes. The protein-encoding genes are almost evenly distributed on both strands, and intergenic space is minimal. Approximately 40% of the viral gene products resemble entries in public databases, including some that are the first of their kind to be detected in a virus. For example, these unique gene products include an aquaglyceroporin in MT325, a potassium ion transporter protein and an alkyl sulfatase in FR483, and a dTDP–glucose pyrophosphorylase in both viruses. Comparison of MT325 and FR483 protein-encoding genes with the prototype chlorella virus PBCV-1 indicates that approximately 82% of the genes are present in all three viruses.

Supplementary data associated with this article is archived in this repository as 4 separate files: Appendices A–D. Each document, in spreadsheet format, shows Gene Name, Genome Position, A.A. length, Peptid e Mw, pI, CDD Hit Number, COGs, COG Definition, Bit Score, E-value, % Identity, % Positive, Query from-to, Hit from-to, BLASTp Hit Number, Hit Accession, BLASTp Definition, Bit Score, E-value, % Identity, % Positive, Query from-to, and Hit from-to.

Appendix A: Gene Names m002R through m843L
Appendix B: Gene Names M001L through M807R
Appendix C: Gene Names n001L through n849R
Appendix D: Gene Names N003L through N847R

Appendix A: Gene Names m002R through m843L

Gene Name	Genome Position	A.A. length	Peptide Mw	pI	CDD Hit Number	COGs	COG Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from to	BLASTp Hit Number	Hit Accession	BLASTp Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from-to
m002R	406--1008	201	21,987	11.12		No Hit Found									No Hit Found							
m004R	1345--1566	74	8,032	7.02		No Hit Found									No Hit Found							
m006L	1984--1775	70	8,219	11.40		No Hit Found									No Hit Found							
m008R	2751--3119	123	14,086	9.94		No Hit Found									No Hit Found							
m013L	5323--5072	84	9,564	6.53		No Hit Found								1	NP_048541	a194R	74.33	1.17E-12	49%	58%	1--79	1--79
m016R	8002--8271	90	9,632	10.81		No Hit Found									No Hit Found							
m017L	9445--9146	100	11,875	12.25	1	pfam05743	Tsg101, Tumour susceptibility gene 101 protein (TSG101). This family consists of the eukaryotic tumour susceptibility gene 101 protein (TSG101). Altered transcripts of this gene have been detected in sporadic breast cancers and many other human malignanc	37.01	1.80E-03	22%	43%	9--78	244--313		No Hit Found							
m018R	9473--9682	70	7,726	7.15		No Hit Found									No Hit Found							
m020R	10550--10780	77	8,549	10.27		No Hit Found									No Hit Found							
m021L	10985--10761	75	8,826	12.85		No Hit Found									No Hit Found							
m022R	12278--12502	75	8,225	10.51		No Hit Found									No Hit Found							
m027L	13919--13683	79	8,883	12.51		No Hit Found									No Hit Found							
m028R	13883--14077	65	7,805	11.07		No Hit Found								1	NP_048404	contains type 1 hydrophobic transmembrane region	59.31	3.85E-08	59%	66%	1--62	43--104
m029R	13984--14631	216	23,949	7.66		No Hit Found								1	NP_048403	a55L	72.79	7.93E-12	24%	58%	1--122	10--131
m031L	15378--14911	156	17,530	12.18		No Hit Found									No Hit Found							
m033R	16234--16551	106	11,946	6.24		No Hit Found									No Hit Found							
m035R	16704--16919	72	7,878	9.71		No Hit Found									No Hit Found							
m036R	17212--17415	68	8,219	11.66		No Hit Found									No Hit Found							
m038L	17670--17449	74	8,357	4.99		No Hit Found									No Hit Found							
m039L	18114--17800	105	11,262	8.68		No Hit Found									No Hit Found							
m040L	18421--18068	118	13,664	9.90		No Hit Found									No Hit Found							
m041L	18699--18385	105	11,221	6.49		No Hit Found									No Hit Found							
m043L	19522--19256	89	9,827	4.45		No Hit Found									No Hit Found							
m044L	19975--19673	101	11,052	4.36		No Hit Found									No Hit Found							
m045L	21577--21281	99	10,863	4.53	1	pfam01698	FLO_LFY, Floricaula / Leafy protein. This family consists of various plant development proteins which are homologues of floricaula (FLO) and Leafy (LFY) proteins which are floral meristem identity proteins. Mutations in the sequences of these proteins af	35.77	4.63E-03	24%	33%	18--72	6--60		No Hit Found							
m046R	21776--21973	66	7,552	10.84		No Hit Found									No Hit Found							
m048L	22622--22185	146	16,111	4.33		No Hit Found									No Hit Found							
m049L	22928--22632	99	10,983	8.40		No Hit Found									No Hit Found							
m050L	23345--23046	100	10,830	4.71		No Hit Found									No Hit Found							
m051L	24366--24076	97	10,644	4.31		No Hit Found									No Hit Found							
m052L	24653--24366	96	10,453	3.85		No Hit Found									No Hit Found							
m053L	24923--24687	79	8,475	8.43		No Hit Found									No Hit Found							
m054L	25580--25374	69	7,721	9.24		No Hit Found									No Hit Found							
m056L	27371--26628	248	26,897	4.65		No Hit Found									No Hit Found							
m057R	26894--27187	98	11,613	6.89		No Hit Found									No Hit Found							
m058L	27707--27489	73	7,952	4.16		No Hit Found									No Hit Found							
m059L	28397--28056	114	12,391	4.58		No Hit Found									No Hit Found							
m060L	30984--30757	76	8,621	7.81		No Hit Found									No Hit Found							
m062L	31584--31147	146	15,831	3.99		No Hit Found									No Hit Found							
m063L	32307--31633	225	23,962	4.66		No Hit Found									No Hit Found							
m064L	32916--32575	114	12,445	4.37		No Hit Found									No Hit Found							
m065L	34177--33806	124	14,508	12.15		No Hit Found									No Hit Found							
m066L	34977--34780	66	6,989	9.51		No Hit Found									No Hit Found							
m067L	35511--35083	143	15,512	10.10		No Hit Found									No Hit Found							
m068L	35587--35117	157	18,213	8.39		No Hit Found									No Hit Found							
m069R	35403--36698	432	49,571	9.77		No Hit Found								1	NP_048711	A354R	198.36	3.93E-49	42%	58%	187--432	2--235
m072L	37356--36865	164	17,831	10.78		No Hit Found									No Hit Found							
m073L	38289--38026	88	9,774	10.22		No Hit Found									No Hit Found							
m074L	38931--38665	89	9,737	4.40		No Hit Found									No Hit Found							
m075L	39174--38968	69	7,939	4.88		No Hit Found									No Hit Found							
m076R	39316--39564	83	9,924	8.93		No Hit Found									No Hit Found							
m077L	40155--39955	67	7,006	9.42		No Hit Found									No Hit Found							
m079R	41369--41662	98	11,476	12.04		No Hit Found									No Hit Found							
m080L	42909--42484	142	16,161	9.58		No Hit Found									No Hit Found							

Gene Name	Genome Position	A.A. length	Peptide Mw	pI	CDD Hit Number	COGs	COG Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from to	BLASTp Hit Number	Hit Accession	BLASTp Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from-to
m082L	43103--42711	131	14,511	8.09		No Hit Found									No Hit Found							
m084L	43646--43356	97	10,346	11.01		No Hit Found								1	NP_048548	A201L	119.40	3.15E-26	78%	91%	25--93	24--92
m087L	45838--45353	162	18,201	10.11		No Hit Found									No Hit Found							
m089L	46601--45978	208	22,820	11.93	1	COG3119	AslA, Arylsulfatase A and related enzymes [Inorganic ion transport and metabolism].	35.87	3.82E-03	27%	43%	68--143	8--71		No Hit Found							
m092L	48155--47943	71	8,214	10.98		No Hit Found								1	NP_048647	a293R	57.77	1.10E-07	56%	65%	17--71	5--59
m094L	49089--48862	76	8,968	9.95		No Hit Found									No Hit Found							
m095R	49042--49245	68	7,667	7.51		No Hit Found									No Hit Found							
m096R	49139--49405	89	9,821	8.27		No Hit Found									No Hit Found							
m098R	49636--50043	136	15,838	9.47		No Hit Found									No Hit Found							
m102L	52463--52128	112	11,772	8.35		No Hit Found									No Hit Found							
m104L	53598--53269	110	12,338	12.81		No Hit Found									No Hit Found							
m105R	53501--53740	80	8,771	10.60		No Hit Found									No Hit Found							
m106R	53592--53804	71	8,282	11.67		No Hit Found									No Hit Found							
m108L	54210--53923	96	10,934	8.46		No Hit Found									No Hit Found							
m110L	55378--55175	68	7,587	11.46		No Hit Found									No Hit Found							
m112L	56110--55808	101	11,820	8.08		No Hit Found									No Hit Found							
m113L	56421--56131	97	11,314	9.26		No Hit Found									No Hit Found							
m115R	56850--57080	77	9,170	10.32		No Hit Found									No Hit Found							
m117R	57070--57549	160	18,272	8.74		No Hit Found									No Hit Found							
m118R	57197--57397	67	7,617	11.31		No Hit Found									No Hit Found							
m119R	57772--58005	78	8,655	9.54		No Hit Found									No Hit Found							
m123L	59143--58925	73	8,218	8.43		No Hit Found									No Hit Found							
m125R	59183--59386	68	7,610	9.23		No Hit Found									No Hit Found							
m126L	59546--59349	66	7,634	10.39		No Hit Found									No Hit Found							
m127R	59969--60166	66	6,979	11.55		No Hit Found									No Hit Found							
m129L	61436--60564	291	32,015	7.60		No Hit Found									No Hit Found							
m130R	60982--61284	101	10,868	11.97		No Hit Found									No Hit Found							
m131L	61450--61181	90	10,385	10.73		No Hit Found									No Hit Found							
m132R	61348--61668	107	11,636	10.71		No Hit Found									No Hit Found							
m134L	63080--62343	246	27,343	10.28		No Hit Found									No Hit Found							
m135L	63466--63209	86	9,385	11.43		No Hit Found									No Hit Found							
m138R	63993--64250	86	9,861	8.21		No Hit Found									No Hit Found							
m141R	65554--65787	78	8,665	11.88		No Hit Found									No Hit Found							
m142R	65562--65795	78	8,803	7.15		No Hit Found									No Hit Found							
m144R	66095--66400	102	12,245	7.97		No Hit Found									No Hit Found							
m145R	66456--66686	77	8,012	4.24		No Hit Found									No Hit Found							
m146L	67062--66835	76	8,599	10.65		No Hit Found									No Hit Found							
m147R	67584--67784	67	7,447	11.07		No Hit Found									No Hit Found							
m151L	68990--68790	67	7,942	10.38		No Hit Found									No Hit Found							
m153L	70066--69116	317	35,140	6.06		No Hit Found								1	XP_499823	hypothetical protein	77.80	4.92E-13	29%	43%	9--270	149--404
														2	AAC39772	hepatitis A virus cellular receptor 1 long form	58.92	2.37E-07	28%	36%	9--234	152--356
														3	AAC39771	hepatitis A virus cellular receptor 1 short form	58.92	2.37E-07	28%	36%	9--234	147--351
														4	CAA66906	HAVcr-1 protein	58.15	4.03E-07	33%	41%	108--242	147--305
														5	CAE74002	Hypothetical protein CBG21638	57.77	5.27E-07	22%	38%	3--253	202--452
														6	BAB03062	unnamed protein product	56.61	1.17E-06	27%	38%	14--293	868--1133
														7	AAC39774	hepatitis A virus cellular receptor 1 long form	56.23	1.53E-06	27%	35%	9--234	152--360
														8	AAC39773	hepatitis A virus cellular receptor 1 short form	56.23	1.53E-06	27%	35%	9--234	147--355
														9	NP_001018272	hypothetical protein SPAPB18E9.04c	53.53	9.94E-06	26%	40%	22--254	80--335
														10	XP_502320	hypothetical protein	52.76	1.70E-05	26%	39%	28--254	311--532
m154R	69889--70083	65	7,314	6.51		No Hit Found									No Hit Found							
m155L	70945--70082	288	31,696	4.48		No Hit Found								1	XP_499823	hypothetical protein	82.80	1.31E-14	28%	46%	2--288	153--435
														2	AAC39772	hepatitis A virus cellular receptor 1 long form	65.08	2.82E-09	28%	37%	31--231	154--328
														3	AAC39771	hepatitis A virus cellular receptor 1 short form	65.08	2.82E-09	28%	37%	31--231	149--323
														4	AAC39774	hepatitis A virus cellular receptor 1 long form	63.54	8.21E-09	29%	36%	31--231	154--332
														5	AAC39773	hepatitis A virus cellular receptor 1 short form	63.54	8.21E-09	29%	36%	31--231	149--327
														6	BAA21556	hepatitis A virus receptor	59.69	1.19E-07	29%	36%	31--231	149--314
														7	CAA66906	HAVcr-1 protein	58.15	3.45E-07	29%	38%	27--231	137--305
														8	CAE74002	Hypothetical protein CBG21638	55.84	1.71E-06	22%	39%	21--285	286--535
														9	BAB03062	unnamed protein product	55.07	2.92E-06	24%	34%	4--285	198--499
														10	XP_667732	hypothetical protein Chro.10070	51.60	3.23E-05	26%	39%	24--269	2374--2578
m156L	71659--70991	223	24,748	4.54		No Hit Found								1	XP_499823	hypothetical protein	65.08	1.78E-09	25%	44%	2--211	153--409
m157L	72478--72140	113	12,142	5.81		No Hit Found									No Hit Found							
m158L	73897--73517	127	13,996	7.91		No Hit Found									No Hit Found							
m159L	74736--73984	251	28,065	9.84		No Hit Found								1	NP_048635	a281R	117.86	2.89E-25	50%	54%	99--231	14--164
m161L	75058--74843	72	8,381	9.45		No Hit Found									No Hit Found							
m162L	75287--75090	66	7,833	8.34		No Hit Found									No Hit Found							

Gene Name	Genome Position	A.A. length	Peptide Mw	pI	CDD Hit Number	COGs	COG Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from to	BLASTp Hit Number	Hit Accession	BLASTp Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from-to
m164L	75555--75361	65	7,300	11.52		No Hit Found									No Hit Found							
m165R	75447--75647	67	7,297	11.06		No Hit Found									No Hit Found							
m168R	77002--77304	101	12,235	8.08		No Hit Found									No Hit Found							
m169R	77057--77332	92	10,280	5.66		No Hit Found									No Hit Found							
m172R	77947--78162	72	8,084	7.32		No Hit Found									No Hit Found							
m175R	78679--78897	73	8,288	10.14		No Hit Found									No Hit Found							
m176R	79213--79434	74	8,519	5.86		No Hit Found									No Hit Found							
m181R	83089--83283	65	7,736	9.38		No Hit Found									No Hit Found							
m182R	83120--83344	75	8,985	9.41		No Hit Found									No Hit Found							
m185R	84298--84501	68	7,612	9.01		No Hit Found									No Hit Found							
m187L	84637--84401	79	9,091	10.42		No Hit Found									No Hit Found							
m188R	85442--85699	86	9,636	7.59		No Hit Found									No Hit Found							
m189L	86056--85853	68	7,770	10.47		No Hit Found									No Hit Found							
m191R	86062--86256	65	7,379	12.50		No Hit Found									No Hit Found							
m192L	86518--86288	77	8,531	11.89		No Hit Found									No Hit Found							
m193L	86978--86649	110	12,210	8.47		No Hit Found									No Hit Found							
m194R	86944--87165	74	8,176	10.23		No Hit Found									No Hit Found							
m195L	87611--87258	118	13,100	6.88		No Hit Found									No Hit Found							
m196L	87908--87678	77	8,518	7.62		No Hit Found									No Hit Found							
m197L	88675--88475	67	6,606	4.11		No Hit Found									No Hit Found							
m198L	89035--88763	91	7,908	3.58		No Hit Found									No Hit Found							
m199R	88777--89040	88	9,997	12.36		No Hit Found									No Hit Found							
m202L	90305--90036	90	10,585	10.88		No Hit Found									No Hit Found							
m203L	90361--90113	83	9,179	9.43		No Hit Found									No Hit Found							
m205R	91243--91446	68	7,275	5.66		No Hit Found									No Hit Found							
m207L	92549--91956	198	21,901	9.47		No Hit Found									No Hit Found							
m209R	92076--92375	100	10,930	11.09		No Hit Found									No Hit Found							
m211R	92723--92977	85	9,935	5.93		No Hit Found									No Hit Found							
m212R	92799--93008	70	8,300	8.17		No Hit Found									No Hit Found							
m214L	93214--93002	71	8,387	11.22		No Hit Found									No Hit Found							
m216L	94192--93989	68	7,519	11.32		No Hit Found									No Hit Found							
m217L	94725--94324	134	14,250	11.13		No Hit Found									No Hit Found							
m220R	95282--95596	105	11,594	10.62		No Hit Found									No Hit Found							
m222R	95790--96113	108	11,755	11.02		No Hit Found								1	NP_048598	a249L	53.91	1.62E-06	44%	50%	18--102	17--102
m223R	95852--96346	165	19,121	10.66		No Hit Found									No Hit Found							
m224L	96651--96373	93	10,283	10.36		No Hit Found									No Hit Found							
m226R	96669--97106	146	16,118	11.60		No Hit Found									No Hit Found							
m227R	97242--97442	67	7,658	10.52		No Hit Found									No Hit Found							
m228L	98010--97807	68	7,836	10.79		No Hit Found									No Hit Found							
m229R	98004--98408	135	15,086	10.34		No Hit Found									No Hit Found							
m230R	98541--98750	70	7,904	5.63		No Hit Found									No Hit Found							
m231R	98648--98857	70	8,372	9.09		No Hit Found									No Hit Found							
m234R	99504--99701	66	7,387	11.07		No Hit Found								1	NP 048586	a238L	83.19	2.48E-15	66%	72%	2--66	48--112
m235R	99846--100055	70	7,784	9.64		No Hit Found									No Hit Found							
m236R	100367--100780	138	15,833	8.18		No Hit Found									No Hit Found							
m239L	101646--101323	108	12,058	8.36		No Hit Found									No Hit Found							
m244R	103086--103475	130	14,802	9.99		No Hit Found								1	NP_048575	A227L	162.93	2.43E-39	60%	78%	3--129	10--136
m245L	103420--103190	77	8,876	10.69		No Hit Found								1	NP_048576	a228R	72.79	3.36E-12	51%	62%	1--77	1--76
m248R	103875--104153	93	10,394	11.82		No Hit Found									No Hit Found							
m250L	104786--104517	90	9,825	11.81		No Hit Found									No Hit Found							
m251R	104705--105112	136	15,210	8.24		No Hit Found									No Hit Found							
m252L	104988--104791	66	7,566	11.93		No Hit Found									No Hit Found							
m254R	105328--105609	94	10,949	8.36		No Hit Found									No Hit Found							
m255L	105987--105775	71	7,609	4.99		No Hit Found									No Hit Found							
m256R	106227--106427	67	7,671	9.01		No Hit Found									No Hit Found							
m257R	106575--106811	79	8,744	10.82		No Hit Found									No Hit Found							
m259R	107614--107814	67	7,648	12.22		No Hit Found									No Hit Found							
m260R	107861--108202	114	13,241	10.49		No Hit Found									No Hit Found							

Gene Name	Genome Position	A.A. length	Peptide Mw	pI	CDD Hit Number	COGs	COG Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from to	BLASTp Hit Number	Hit Accession	BLASTp Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from-to
m261R	108103–108387	95	10,771	8.52		No Hit Found									No Hit Found							
m262L	108508–108278	77	8,582	10.64		No Hit Found									No Hit Found							
m265L	110071–109826	82	9,249	10.32		No Hit Found									No Hit Found							
m267L	110331–110119	71	7,953	4.56		No Hit Found									No Hit Found							
m268R	110517–110729	71	8,512	11.89		No Hit Found									No Hit Found							
m270R	111648–111842	65	7,318	7.29		No Hit Found									No Hit Found							
m271L	111986–111708	93	11,395	10.66		No Hit Found									No Hit Found							
m273R	111941–112186	82	9,539	9.91		No Hit Found									No Hit Found							
m274R	112244–112468	75	8,244	6.40		No Hit Found									No Hit Found							
m275R	112616–112897	94	10,896	11.93		No Hit Found									No Hit Found							
m276L	113591–113304	96	10,835	11.12		No Hit Found									No Hit Found							
m277L	113988–113377	204	22,758	8.55		No Hit Found									No Hit Found							
m279R	113597–113959	121	13,681	12.56		No Hit Found									No Hit Found							
m280L	114363–114136	76	8,538	7.32		No Hit Found									No Hit Found							
m281L	114563–114369	65	7,038	12.23		No Hit Found									No Hit Found							
m284L	115387–115025	121	14,401	9.48		No Hit Found									No Hit Found							
m285L	115368–115105	88	9,895	11.95		No Hit Found									No Hit Found							
m286L	115960–115682	93	10,741	10.24		No Hit Found									No Hit Found							
m287L	116160–115906	85	9,533	10.91		No Hit Found									No Hit Found							
m290R	117315–117527	71	8,140	8.46		No Hit Found									No Hit Found							
m291L	117744–117439	102	11,319	4.95		No Hit Found									No Hit Found							
m292L	117991–117611	127	14,385	8.73		No Hit Found									No Hit Found							
m293L	118314–118069	82	9,088	10.16		No Hit Found									No Hit Found							
m295L	119073–118456	206	23,413	10.50		No Hit Found									No Hit Found							
m296L	119251–118901	117	12,947	8.81		No Hit Found									No Hit Found							
m297L	119390–119163	76	8,274	4.83		No Hit Found									No Hit Found							
m300L	120203–119898	102	11,861	9.48		No Hit Found									No Hit Found							
m301L	120447–120220	76	9,522	13.28		No Hit Found									No Hit Found							
m303R	120462–121241	260	26,935	12.50		No Hit Found									No Hit Found							
m304L	120771–120532	80	9,697	13.03		No Hit Found									No Hit Found							
m305R	120878–121216	113	12,349	6.46		No Hit Found									No Hit Found							
m306R	121304–121783	160	17,675	5.42		No Hit Found									No Hit Found							
m309R	122771–123256	162	17,764	11.83		No Hit Found									No Hit Found							
m311R	123590–123814	75	7,705	10.62		No Hit Found								1	NP_048551	a204L	84.34	1.12E-15	65%	67%	2--74	6--78
m316R	125397–125696	100	11,765	9.99		No Hit Found									No Hit Found							
m318R	126300–126710	137	14,710	11.07		No Hit Found									No Hit Found							
m320R	126788–127180	131	13,962	4.67		No Hit Found									No Hit Found							
m321L	127242–127021	74	8,170	11.30		No Hit Found									No Hit Found							
m325L	129162–128968	65	7,331	10.55		No Hit Found									No Hit Found							
m327L	129461–129204	86	10,035	10.29		No Hit Found									No Hit Found							
m330R	129967–130338	124	13,546	6.23		No Hit Found									No Hit Found							
m331R	130163–130543	127	14,296	10.96		No Hit Found									No Hit Found							
m333L	131493–131143	117	13,038	11.75		No Hit Found									No Hit Found							
m335L	132326–131694	211	24,704	9.01		No Hit Found								1	NP_048682	A326L	241.89	9.33E-63	58%	77%	14--208	1--197
m336R	132317–132763	149	17,791	11.09		No Hit Found									No Hit Found							
m338R	132869–133270	134	15,825	9.07		No Hit Found									No Hit Found							
m339R	133035–133403	123	14,050	8.82		No Hit Found									No Hit Found							
m340L	134014–133568	149	15,783	4.36		No Hit Found									No Hit Found							
m342L	134328–134092	79	8,598	9.54		No Hit Found									No Hit Found							
m343L	134957–134694	88	10,005	11.68		No Hit Found									No Hit Found							
m345L	136343–136149	65	7,482	9.44		No Hit Found									No Hit Found							
m347R	136321–136632	104	11,830	8.39		No Hit Found									No Hit Found							
m349R	136790–137071	94	9,874	11.44		No Hit Found									No Hit Found							
m350L	137303–136971	111	12,566	12.11		No Hit Found									No Hit Found							
m351R	137080–137358	93	10,302	7.18		No Hit Found									No Hit Found							
m352L	137918–137451	156	16,878	10.73		No Hit Found									No Hit Found							
m353L	138776–138432	115	13,428	9.08		No Hit Found									No Hit Found							

Gene Name	Genome Position	A.A. length	Peptide Mw	pI	CDD Hit Number	COGs	COG Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from to	BLASTp Hit Number	Hit Accession	BLASTp Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from-to
m355R	138625--138909	95	10,690	11.28		No Hit Found									No Hit Found							
m356R	138988--139191	68	7,483	12.04		No Hit Found									No Hit Found							
m360R	141751--140726	342	13,254	12.13		No Hit Found									No Hit Found							
m364R	142845--143192	116	12,809	11.72		No Hit Found									No Hit Found							
m365L	143505--143173	111	14,534	13.51		No Hit Found									No Hit Found							
m366R	143202--143642	147	14,331	6.43	1	COG3889	COG3889, Predicted solute binding protein [General function prediction onlv].	45.42	5.98E-06	30%	59%	12--103	758--850		No Hit Found							
m368L	144406--144146	87	10,088	9.57		No Hit Found									No Hit Found							
m369L	144817--144425	131	14,771	7.46		No Hit Found									No Hit Found							
m371L	145139--144936	68	8,057	8.09		No Hit Found									No Hit Found							
m373L	147588--147247	114	12,771	11.00		No Hit Found									No Hit Found							
m374L	147919--147722	66	7,231	4.16		No Hit Found									No Hit Found							
m378L	150262--149987	92	9,916	11.55		No Hit Found									No Hit Found							
m380L	150985--150779	69	7,991	8.67		No Hit Found									No Hit Found							
m382L	151760--151497	88	10,264	11.31		No Hit Found									No Hit Found							
m383L	152087--151890	66	7,013	8.17		No Hit Found									No Hit Found							
m384L	152546--152349	66	7,641	8.65		No Hit Found									No Hit Found							
m386R	152797--153084	96	10,846	8.04		No Hit Found									No Hit Found							
m387R	153385--154275	297	33,023	6.53		No Hit Found								1	NP_048635	a281R	114.01	5.60E-24	50%	55%	124--277	14--164
m390L	155619--155380	80	9,266	5.13		No Hit Found									No Hit Found							
m391L	156030--155731	100	11,267	6.48		No Hit Found									No Hit Found							
m392L	156173--155943	77	8,749	9.47		No Hit Found									No Hit Found							
m393R	155957--156175	73	8,628	10.76		No Hit Found									No Hit Found							
m394L	156620--156312	103	11,566	7.34		No Hit Found									No Hit Found							
m396R	156675--156941	89	10,042	11.44		No Hit Found									No Hit Found							
m398R	157206--157541	112	12,842	11.36		No Hit Found									No Hit Found							
m400L	158861--158496	122	13,544	10.38		No Hit Found									No Hit Found							
m402L	159108--158845	88	10,297	10.48		No Hit Found									No Hit Found							
m405L	160671--160357	105	11,781	11.79		No Hit Found								1	NP_048803	a446R	56.61	2.52E-07	60%	66%	55--105	2--52
m406L	161079--160885	65	7,660	11.13		No Hit Found									No Hit Found							
m408L	162340--161600	247	23,928	4.24	1	pfam05642	Sporozoite_P67, Sporozoite P67 surface antigen. This family consists of several Theileria P67 surface antigens. A stage specific surface antigen of Theileria parva, p67, is the basis for the development of an anti-soorozoite vaccine for the control of Ea	35.54	4.72E-03	39%	44%	106--194	185--275		No Hit Found							
m410L	163253--163038	72	8,079	9.69		No Hit Found									No Hit Found							
m412R	163120--163332	71	8,024	6.34		No Hit Found									No Hit Found							
m414R	163833--164213	127	13,143	9.29		No Hit Found								1	ZP_00372892	hypothetical protein WwSim0361	64.31	1.19E-09	42%	50%	19--127	15--122
														2	BAE63414	unnamed protein product	59.31	3.82E-08	40%	49%	3--127	38--166
														3	ZP_00372930	hypothetical protein WwSim0771	59.31	3.82E-08	39%	43%	11--127	7--122
m415R	163984--164427	148	16,959	11.95		No Hit Found									No Hit Found							
m418R	165741--166007	89	10,163	4.43		No Hit Found									No Hit Found							
m419R	166383--166703	107	9,403	8.68		No Hit Found									No Hit Found							
m420R	166897--167292	132	15,375	12.08		No Hit Found									No Hit Found							
m422L	168066--167671	132	15,709	10.56		No Hit Found									No Hit Found							
m424L	168014--167781	78	8,948	11.94		No Hit Found									No Hit Found							
m426L	168393--168199	65	7,194	4.91		No Hit Found									No Hit Found							
m428L	168933--168712	74	8,259	7.65		No Hit Found									No Hit Found							
m429L	169269--169072	66	7,309	7.97		No Hit Found									No Hit Found							
m431R	169803--170045	81	9,412	11.40		No Hit Found									No Hit Found							
m432R	170094--170339	82	9,029	10.25		No Hit Found									No Hit Found							
m433R	170566--170781	72	8,553	12.61		No Hit Found									No Hit Found							
m434R	170706--171065	120	13,910	5.86		No Hit Found									No Hit Found							
m436R	170959--171360	134	15,335	10.45		No Hit Found								1	NP_048778	A421R	86.66	2.25E-16	53%	68%	60--128	23--91
m439L	171801--171595	69	7,871	10.60		No Hit Found									No Hit Found							
m441L	172878--172141	246	28,123	8.64		No Hit Found									No Hit Found							
m443L	172616--172221	132	15,721	10.86		No Hit Found									No Hit Found							
m446R	172973--173236	88	9,939	10.35		No Hit Found									No Hit Found							
m447R	173255--173524	90	10,727	9.39		No Hit Found									No Hit Found							
m450R	173707--173970	88	9,471	10.11		No Hit Found									No Hit Found							
m452R	174059--174325	89	10,279	10.83		No Hit Found									No Hit Found							
m453R	175006--175218	71	8,470	8.55		No Hit Found									No Hit Found							

Gene Name	Genome Position	A.A. length	Peptide Mw	pI	CDD Hit Number	COGs	COG Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from to	BLASTp Hit Number	Hit Accession	BLASTp Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from-to
m456L	176463--176071	131	15,409	11.10		No Hit Found									No Hit Found							
m458L	176799--176533	89	10,130	10.99		No Hit Found									No Hit Found							
m459L	176915--176643	91	10,412	7.81		No Hit Found									No Hit Found							
m461L	177809--177327	161	18,843	10.84		No Hit Found									No Hit Found							
m462L	177892--177671	74	8,117	4.22		No Hit Found									No Hit Found							
m464R	179187--179435	83	9,781	9.42		No Hit Found									No Hit Found							
m466R	180176--180388	71	8,158	9.73		No Hit Found									No Hit Found							
m468L	181119--180901	73	8,736	10.25		No Hit Found									No Hit Found							
m469L	181264--180977	96	10,878	11.76		No Hit Found									No Hit Found							
m470L	181909--181709	67	7,565	10.71		No Hit Found									No Hit Found							
m471L	183057--182830	76	8,083	11.66		No Hit Found									No Hit Found							
m473L	183499--183272	76	8,279	7.86		No Hit Found									No Hit Found							
m474L	183918--183592	109	12,268	10.19		No Hit Found									No Hit Found							
m477R	184604--184816	71	7,675	10.57		No Hit Found									No Hit Found							
m480L	185780--185550	77	8,473	11.31		No Hit Found									No Hit Found							
m481L	185980--185774	69	8,077	11.91		No Hit Found									No Hit Found							
m483R	186328--186606	93	10,660	11.26		No Hit Found									No Hit Found							
m486L	187726--187436	97	10,790	5.04		No Hit Found									No Hit Found							
m487R	187873--188391	173	20,718	4.10		No Hit Found								1	ZP_00433134	hypothetical protein Bmal10_01004924	59.31	5.32E-08	30%	38%	47--137	2--90
														2	ZP_00054046	COG3210: Large exoproteins involved in heme utilization or adhesion	57.77	1.55E-07	31%	42%	19--119	485--585
														3	XP_344436	PREDICTED: similar to Hypothetical protein CBG20540	56.23	4.50E-07	24%	49%	16--122	252--358
														4	AAB27498	WI-1	53.91	2.24E-06	28%	43%	28--134	1--115
														5	XP_795005	PREDICTED: similar to transmembrane protease, serine 2	51.99	8.49E-06	29%	35%	22--133	843--966
														6	CAF92931	unnamed protein product	50.45	2.47E-05	41%	45%	43--119	260--336
														7	AAA91036	WI-1 adhesin	49.29	5.51E-05	21%	35%	24--140	556--667
														8	XP_584941	PREDICTED: similar to Epsin 1 isoform 1	48.52	9.39E-05	34%	37%	29--131	278--388
														9	XP_875111	PREDICTED: similar to Epsin 1 isoform 2	48.52	9.39E-05	34%	37%	29--131	283--393
														10	XP_800603	PREDICTED: hypothetical protein XP_795510	48.52	9.39E-05	33%	38%	16--128	21--125
m489R	187941--188480	180	19,136	8.04		No Hit Found								1	AAH42633	Extraembryonic, spermatogenesis, homeobox 1-like	75.49	7.95E-13	45%	52%	7--102	294--377
														2	NP_703149	extraembryonic, spermatogenesis, homeobox 1-like	70.86	1.96E-11	45%	53%	7--102	303--377
														3	AAB16964	collagen-like protein	68.94	7.44E-11	29%	54%	4--102	140--238
														4	XP_421396	PREDICTED: similar to adenylosuccinate synthetase:ISOTYPE=muscle	68.94	7.44E-11	40%	53%	7--96	504--594
														5	ZP_00283849	COG4991: Uncharacterized protein with a bacterial SH3 domain homologue	68.94	7.44E-11	37%	53%	2--100	155--253
														6	XP_780848	PREDICTED: similar to Aryl hydrocarbon receptor precursor (Ah receptor) (AhR)	68.17	1.27E-10	45%	55%	5--94	381--481
														7	ZP_00675930	hypothetical protein TeryDRAFT_0036	65.08	1.07E-09	35%	42%	2--102	329--429
														8	BAE24926	unnamed protein product	64.31	1.83E-09	36%	53%	7--102	277--372
														9	AAF34245	zinc finger protein ZAC1	64.31	1.83E-09	36%	53%	7--102	277--372
														10	NP_033564	pleiomorphic adenoma gene-like 1	64.31	1.83E-09	36%	53%	7--102	277--372
m490L	188607--188398	70	8,412	11.87		No Hit Found									No Hit Found							
m492L	189190--188975	72	8,399	12.17		No Hit Found									No Hit Found							
m493L	189269--189006	88	9,350	3.65		No Hit Found									No Hit Found							
m494L	189635--189432	68	7,425	8.43		No Hit Found									No Hit Found							
m498L	191788--191483	102	12,357	9.89		No Hit Found									No Hit Found							
m499R	191600--191800	67	7,016	7.26		No Hit Found									No Hit Found							
m500L	191928--191734	65	6,824	9.68		No Hit Found									No Hit Found							
m503R	192663--192923	87	9,505	7.85		No Hit Found									No Hit Found							
m505R	193511--193777	89	10,512	6.76		No Hit Found									No Hit Found							
m509L	195125--194832	98	10,361	9.29		No Hit Found									No Hit Found							
m510L	195823--195557	89	10,289	8.82		No Hit Found									No Hit Found							
m513R	196352--196678	109	13,036	11.25		No Hit Found									No Hit Found							
m514L	197128--196385	248	28,609	5.32		No Hit Found								1	NP_048893	A537L	107.84	2.92E-22	30%	46%	4--248	10--263
m516L	197407--197183	75	8,403	6.78		No Hit Found									No Hit Found							
m517R	197246--197749	168	19,026	10.39		No Hit Found								1	NP_048895	A539R	177.95	9.53E-44	50%	70%	1--168	1--173
														2	NP_048482	similar to bacteriophage T4 intron-associated endonuclease, corresponds to Swiss-Prot Accession Number P13299	50.45	2.29E-05	53%	68%	28--74	11--55
m519L	198025--197813	71	8,027	12.02		No Hit Found									No Hit Found							
m520R	199333--199668	112	12,587	10.19		No Hit Found									No Hit Found							
m521L	201067--200873	65	7,479	11.17		No Hit Found									No Hit Found							
m523L	202412--202161	84	9,653	8.46		No Hit Found									No Hit Found							
m524R	203196--203426	77	9,149	8.12		No Hit Found									No Hit Found							
m527R	205326--205580	85	9,179	5.59		No Hit Found								1	NP_048897	a541R	49.68	3.08E-05	40%	48%	1--79	10--86
m528R	206769--207014	82	9,367	4.03		No Hit Found									No Hit Found							
m529R	207508--207714	69	8,312	11.00		No Hit Found									No Hit Found							
m532R	208354--208560	69	7,721	7.26		No Hit Found									No Hit Found							
m533R	208901--209155	85	9,403	7.89		No Hit Found									No Hit Found							
m534R	209205--209474	90	10,036	4.17		No Hit Found									No Hit Found							

Gene Name	Genome Position	A.A. length	Peptide Mw	pI	CDD Hit Number	COGs	COG Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from to	BLASTp Hit Number	Hit Accession	BLASTp Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from-to
m536R	209245--209604	120	13,509	11.01		No Hit Found									No Hit Found							
m537R	209650--210063	138	14,807	10.13		No Hit Found								1	XP_719031	hypothetical protein CaO19_6069	51.22	1.03E-05	47%	57%	73--138	1--71
m538R	210184--210414	77	8,142	9.88		No Hit Found									No Hit Found							
m539R	210439--210873	145	16,355	7.46		No Hit Found									No Hit Found							
m540L	211104--210829	92	10,460	8.25		No Hit Found									No Hit Found							
m541R	210931--211455	175	19,176	10.91		No Hit Found								1	ZP_00886604	hypothetical protein CsacDRAFT_2497	51.22	1.50E-05	36%	45%	73--169	5--112
m542R	211603--211881	93	10,213	10.82		No Hit Found									No Hit Found							
m544L	212294--212097	66	7,409	10.36		No Hit Found									No Hit Found							
m545L	212862--212653	70	8,227	7.74		No Hit Found									No Hit Found							
m547R	213534--213743	70	8,062	8.84		No Hit Found									No Hit Found							
m548L	214314--213919	132	14,428	12.46		No Hit Found									No Hit Found							
m549L	214851--214564	96	10,868	10.35		No Hit Found								1	NP_048942	a586R	73.94	1.52E-12	60%	68%	1--64	12--75
m550L	215851--214862	330	36,801	7.75		No Hit Found									No Hit Found							
m551L	215751--215542	70	7,776	12.41		No Hit Found								1	NP_048940	contains prenyl group binding site (CAAX) box	78.95	4.62E-14	58%	70%	1--70	29--98
m552L	216399--215962	146	17,293	9.37		No Hit Found									No Hit Found							
m554L	216622--216275	116	13,759	9.68		No Hit Found									No Hit Found							
m555L	216794--216498	99	11,134	9.37		No Hit Found									No Hit Found							
m558R	218073--218720	216	28,188	11.65		No Hit Found								1	CAE56520	Hypothetical protein CBG24243	71.25	2.31E-11	41%	53%	88--207	37--166
														2	XP_509745	PREDICTED: similar to Coagulation factor VII (Serum prothrombin conversion accelerator)	63.93	3.68E-09	25%	50%	112--177	3980--4045
														3	AAX27666	unknown	62.39	1.07E-08	35%	35%	87--184	14--111
														4	AAB01097	glycine-rich protein	60.85	3.12E-08	43%	48%	91--172	20--99
														5	XP_364198	hypothetical protein MG09043.4	56.61	5.88E-07	28%	40%	96--200	370--472
														6	CAE57148	Hypothetical protein CBG25081	56.23	7.68E-07	36%	44%	88--210	206--342
														7	EAL26857	GA19788-PA	53.91	3.81E-06	46%	53%	92--172	421--503
														8	ABA47855	hypothetical protein BURPS1710b_1103	53.14	6.50E-06	40%	41%	105--181	58--139
														9	BAC86188	unnamed protein product	52.76	8.49E-06	30%	32%	139--213	438--533
														10	CAE73786	Hypothetical protein CBG21336	51.60	1.89E-05	33%	40%	87--210	90--164
m559R	218182--218796	205	16,941	3.64		No Hit Found									No Hit Found							
m561R	218995--219360	122	14,412	12.56		No Hit Found									No Hit Found							
m562R	219308--219763	152	17,062	7.97		No Hit Found									No Hit Found							
m565R	220197--220544	116	12,700	7.38		No Hit Found									No Hit Found							
m566R	220465--220707	81	8,736	6.41		No Hit Found									No Hit Found							
m568R	220749--221102	118	12,606	5.13		No Hit Found									No Hit Found							
m573R	222557--222784	76	8,550	11.01		No Hit Found									No Hit Found							
m575R	222990--223187	66	7,658	7.82		No Hit Found									No Hit Found							
m576R	223021--223305	95	10,472	8.56		No Hit Found									No Hit Found							
m579L	224482--224252	77	8,952	8.84		No Hit Found									No Hit Found							
m581R	224324--224605	94	11,130	9.55		No Hit Found									No Hit Found							
m582R	224481--224771	97	11,724	10.66		No Hit Found									No Hit Found							
m583L	225246--224911	112	12,616	4.54		No Hit Found								1	NP_048751	A394R	59.69	2.90E-08	50%	67%	17--71	63--116
m585R	225310--225513	68	7,667	7.29		No Hit Found									No Hit Found							
m587R	225640--225969	110	12,466	11.03		No Hit Found									No Hit Found							
m588R	225711--225953	81	9,833	7.60		No Hit Found									No Hit Found							
m589L	225961--225755	69	8,335	11.74		No Hit Found									No Hit Found							
m590L	226414--226163	84	10,381	8.50		No Hit Found									No Hit Found							
m592L	226406--226182	75	7,909	11.73	1	pfam03346	Lipoprotein_14, Actinobacillus constitutively-expressed outer membrane lipoprotein A..	39.74	2.82E-04	33%	48%	2--56	64--118	1	XP_536513	PREDICTED: similar to AHNAK nucleoprotein isoform 1	64.70	9.22E-10	72%	74%	21--67	237--283
														2	AAN60242	unknown	60.85	1.33E-08	60%	82%	22--67	71--116
														3	NP_849949	lipid binding	60.85	1.33E-08	60%	82%	22--67	91--136
														4	NP_046976	tail protein	55.84	4.28E-07	40%	65%	21--71	639--690
														5	XP_365380	predicted protein	53.91	1.63E-06	38%	76%	21--67	31--77
														6	XP_343560	PREDICTED: similar to LAF-4 protein (Lymphoid nuclear protein related to AF4)	50.06	2.35E-05	52%	77%	23--66	1735--1778
														7	XP_926450	PREDICTED: hypothetical protein XP_921357	49.68	3.07E-05	50%	70%	25--68	133--176
														8	EAL47534	hypothetical protein 100.t00015	48.14	8.93E-05	50%	61%	23--66	67--110
m593L	226879--226571	103	12,107	5.88		No Hit Found									No Hit Found							
m594L	227162--226689	158	16,876	7.39		No Hit Found								1	NP_048746	a389R	57.77	1.18E-07	52%	53%	55--118	2--64
m595R	226710--226958	83	9,668	10.72		No Hit Found									No Hit Found							
m596L	227504--227190	105	10,901	11.09		No Hit Found									No Hit Found							
m597R	227549--227782	78	9,371	12.11		No Hit Found								1	CAE73829	Hypothetical protein CBG21394	59.69	2.94E-08	28%	77%	3--69	55--125
														2	NP_033126	repetin	56.61	2.49E-07	54%	58%	16--72	337--417
														3	CAE62907	Hypothetical protein CBG07098	52.37	4.69E-06	25%	71%	3--69	22--92
														4	XP_584748	PREDICTED: similar to concentrative Na+-nucleoside cotransporter	51.60	8.00E-06	27%	53%	3--71	568--640
														5	XP_136142	PREDICTED: similar to periphilin 1 isoform 1	51.60	8.00E-06	35%	67%	11--69	248--306
														6	EAL47849	fimbriae-associated protein, putative	48.91	5.18E-05	28%	57%	2--70	743--815
														7	AAW31832	TraC	48.52	6.77E-05	29%	54%	2--71	551--621
														8	XP_918493	PREDICTED: similar to periphilin 1 isoform 1	48.14	8.84E-05	33%	64%	11--69	256--314
m598L	228408--228196	71	8,177	6.37		No Hit Found									No Hit Found							
m600L	229103--228675	143	13,338	4.46		No Hit Found									No Hit Found							

Gene Name	Genome Position	A.A. length	Peptide Mw	pI	CDD Hit Number	COGs	COG Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from to	BLASTp Hit Number	Hit Accession	BLASTp Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from-to
m602R	229665–229874	70	7,777	4.71		No Hit Found									No Hit Found							
m603R	229877–230146	90	10,142	9.02		No Hit Found									No Hit Found							
m604R	230175–230435	87	10,144	10.03		No Hit Found									No Hit Found							
m608R	231258–231482	75	8,553	11.18		No Hit Found									No Hit Found							
m609L	232244–231798	149	16,808	10.68		No Hit Found									No Hit Found							
m611L	233193–232747	149	17,374	9.49		No Hit Found									No Hit Found							
m613L	233551–233306	82	9,898	10.64		No Hit Found									No Hit Found							
m614L	234221–233913	103	10,728	9.55		No Hit Found									No Hit Found							
m617R	234588–234998	137	12,172	10.48		No Hit Found									No Hit Found							
m619L	235658–235443	72	8,060	9.59		No Hit Found									No Hit Found							
m621R	236035–236253	73	7,561	12.50		No Hit Found									No Hit Found							
m623R	236877–237314	146	16,741	10.23		No Hit Found									No Hit Found							
m624L	237337–237107	77	9,004	11.30		No Hit Found									No Hit Found							
m626L	238229–237798	144	16,890	11.00		No Hit Found								1	NP_048841	A485R	176.41	2.14E-43	57%	78%	4--144	7--147
m630R	239412–239957	182	20,851	10.12		No Hit Found									No Hit Found							
m632R	240321–240938	206	24,644	10.39		No Hit Found									No Hit Found							
m633R	240856–241290	145	16,724	9.34		No Hit Found								1	NP_048635	a281R	54.30	1.22E-06	68%	73%	73--112	10--50
m636L	242219–241983	79	8,978	8.34		No Hit Found									No Hit Found							
m639R	243064–243258	65	7,680	6.87		No Hit Found									No Hit Found							
m640L	243558–243079	160	16,920	4.28		No Hit Found									No Hit Found							
m641L	243676–243473	68	8,028	8.83		No Hit Found									No Hit Found							
m642L	243786–243568	73	7,563	8.37		No Hit Found									No Hit Found							
m643L	243970–243713	86	9,840	6.71		No Hit Found									No Hit Found							
m644L	244353–244102	84	8,333	4.17		No Hit Found									No Hit Found							
m645R	244473–244868	132	15,774	11.24		No Hit Found									No Hit Found							
m646L	244759–244535	75	9,037	11.43		No Hit Found									No Hit Found							
m648L	245910–245242	223	24,608	4.73		No Hit Found								1	NP_048635	a281R	58.54	1.66E-07	44%	47%	140--217	70--164
m650L	246234–246025	70	7,650	9.58		No Hit Found									No Hit Found							
m652L	247565–247158	136	15,259	10.07		No Hit Found									No Hit Found							
m654R	247737–248300	188	21,003	10.58		No Hit Found								1	CAG98466	unnamed protein product	97.06	2.84E-19	45%	53%	4--136	1--133
														2	AAL79317	unknown	54.68	1.62E-06	41%	51%	67--136	2--71
														3	NP_851581	CPXV051A protein	50.83	2.34E-05	53%	63%	85--131	31--77
														4	NP_063682	ribonucleoside-diphosphate reductase	48.91	8.87E-05	53%	65%	85--127	31--73
														5	AAW23723	ribonucleoside-diphosphate reductase	48.91	8.87E-05	53%	65%	85--127	31--73
m655L	248307–248044	88	9,770	12.20		No Hit Found									No Hit Found							
m656R	248276–248581	102	11,361	5.51		No Hit Found									No Hit Found							
m658R	248739–249017	93	10,389	10.78		No Hit Found									No Hit Found							
m659L	249484–249281	68	7,714	12.51		No Hit Found									No Hit Found							
m661R	249468–249791	108	11,503	6.73		No Hit Found									No Hit Found							
m662R	249604–249804	67	7,526	12.17		No Hit Found									No Hit Found							
m663L	250168–249914	85	9,431	10.81		No Hit Found									No Hit Found							
m665R	250297–250515	73	8,493	8.56		No Hit Found									No Hit Found							
m666R	250307–250546	80	8,886	4.61		No Hit Found									No Hit Found							
m668L	252326–251328	333	37,720	8.88		No Hit Found									No Hit Found							
m669L	252400–252110	97	10,877	10.55		No Hit Found									No Hit Found							
m671L	252679–252404	92	10,622	12.29		No Hit Found								1	NP_048822	a466L	67.40	1.39E-10	52%	58%	24--91	19--86
m673R	252619–252858	80	8,851	8.52		No Hit Found									No Hit Found							
m675L	254829–254323	169	19,043	10.45		No Hit Found								1	NP_048817	a460R	55.07	9.28E-07	31%	60%	67--142	2--77
m676L	254668–254471	66	7,208	6.65		No Hit Found									No Hit Found							
m677L	255539–254910	210	23,333	10.44		No Hit Found								1	NP_048816	a459R	105.92	8.03E-22	72%	78%	131--205	1--75
														2	NP_048814	a457R	75.87	8.89E-13	60%	64%	14--88	2--76
m678R	255023–255283	87	9,356	9.89		No Hit Found									No Hit Found							
m679R	255756–256115	120	13,417	10.35		No Hit Found									No Hit Found							
m681L	256491–256267	75	8,246	7.95		No Hit Found									No Hit Found							
m682L	256816–256610	69	7,908	7.25		No Hit Found									No Hit Found							
m684R	256735–257034	100	11,393	7.09		No Hit Found									No Hit Found							
m685R	257437–257700	88	9,272	7.06		No Hit Found									No Hit Found							
m687L	257723–257484	80	8,779	12.52		No Hit Found									No Hit Found							
m688R	257588–257965	126	12,965	4.66		No Hit Found									No Hit Found							
m689L	259013–258789	75	8,358	8.85		No Hit Found									No Hit Found							

Gene Name	Genome Position	A.A. length	Peptide Mw	pI	CDD Hit Number	COGs	COG Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from to	BLASTp Hit Number	Hit Accession	BLASTp Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from-to
m691L	259232--259017	72	7,860	9.69		No Hit Found									No Hit Found							
m692L	260188--259514	225	26,392	7.64		No Hit Found								1	NP 048924	A568L	74.71	2.29E-12	27%	53%	41--194	29--173
m696L	261340--261029	104	11,865	12.43		No Hit Found									No Hit Found							
m698R	261490--261708	73	8,055	9.38		No Hit Found									No Hit Found							
m702R	263252--263500	83	10,082	11.72		No Hit Found									No Hit Found							
m703L	263952--263533	140	15,929	7.34		No Hit Found									No Hit Found							
m705L	265052--264711	114	12,690	11.86		No Hit Found									No Hit Found							
m707R	264862--266199	446	50,164	5.21		No Hit Found								1	NP 048861	A505L	456.45	8.38E-127	56%	72%	48--443	73--478
m708L	265403--265128	92	10,272	10.78		No Hit Found									No Hit Found							
m709L	265694--265461	78	8,488	9.47		No Hit Found									No Hit Found							
m710R	265685--265891	69	7,834	12.50		No Hit Found									No Hit Found							
m711L	266105--265725	127	14,367	12.13		No Hit Found								1	NP_048862	a506R	85.11	6.51E-16	62%	70%	38--107	1--70
m712R	265925--266323	133	14,943	10.72		No Hit Found									No Hit Found							
m713R	266286--267017	244	27,642	5.49		No Hit Found								1	NP_048859	contains prokaryotic membrane lipoprotein lipid attachment site	183.34	5.29E-45	37%	62%	6--243	38--297
m716L	267588--267304	95	10,824	10.44		No Hit Found									No Hit Found							
m718R	268194--268400	69	7,789	10.21		No Hit Found									No Hit Found							
m720R	269695--270036	114	13,297	8.26		No Hit Found								1	NP_048573	a225L	48.91	5.26E-05	39%	69%	62--114	2--54
m722L	270314--270078	79	8,741	8.21		No Hit Found									No Hit Found							
m723R	270224--270475	84	9,947	11.87		No Hit Found								1	NP_048572	a224L	82.80	3.29E-15	55%	61%	1--82	6--88
m724R	270551--270775	75	8,091	11.61		No Hit Found								1	NP_048571	a223aL	59.31	3.87E-08	41%	48%	2--75	19--92
m725R	270778--271284	169	19,093	11.39		No Hit Found								1	T17711	hypothetical protein a221L - Chlorella virus PBCV-1	50.83	1.75E-05	37%	64%	66--124	1--57
m726R	271199--271459	87	10,101	6.78		No Hit Found									No Hit Found							
m728R	271756--271956	67	7,785	11.50		No Hit Found									No Hit Found							
m730R	272215--272433	73	8,098	3.99		No Hit Found									No Hit Found							
m731R	272950--273153	68	7,677	12.13		No Hit Found									No Hit Found							
m732L	273939--273745	65	7,597	6.35		No Hit Found									No Hit Found							
m733R	273860--274807	316	37,047	9.72		No Hit Found								1	NP_048973	similar to Variola virus orf E10L, corresponds to Swiss-Prot Accession Number P33801	303.52	5.48E-81	48%	69%	4--315	6--318
														2	YP_142754	S/T protein kinase, similar to Paramecium bursaria chlorella virus 1 A617R	60.85	6.19E-08	24%	42%	7--245	105--358
														3	AAF36040	Hypothetical protein Y73B6A.1	51.22	4.91E-05	23%	42%	4--219	167--375
m736L	275201--274995	69	7,983	12.91		No Hit Found									No Hit Found							
m737R	275006--275209	68	7,204	11.00		No Hit Found									No Hit Found							
m739L	276077--275793	95	11,219	10.79		No Hit Found									No Hit Found							
m743L	277075--276824	84	9,212	11.75		No Hit Found									No Hit Found							
m744L	277130--276852	93	10,647	5.56		No Hit Found									No Hit Found							
m745L	277792--277586	69	7,813	12.42		No Hit Found									No Hit Found							
m746L	278098--277805	98	10,800	9.60		No Hit Found									No Hit Found							
m747L	279611--279402	70	7,840	11.37		No Hit Found									No Hit Found							
m749R	279997--280473	159	17,780	10.18		No Hit Found									No Hit Found							
m750R	281024--281275	84	9,750	11.31		No Hit Found									No Hit Found							
m751R	281164--281469	102	11,369	6.99		No Hit Found									No Hit Found							
m754R	281720--282265	182	20,563	11.47		No Hit Found									No Hit Found							
m755L	282855--282598	86	9,347	9.33		No Hit Found									No Hit Found							
m759L	284429--284067	121	13,717	8.20		No Hit Found									No Hit Found							
m760L	284676--284332	115	13,329	7.92		No Hit Found									No Hit Found							
m761R	284444--285031	196	21,931	11.13		No Hit Found									No Hit Found							
m764R	285340--285687	116	12,055	11.37		No Hit Found									No Hit Found							
m765R	286458--286760	101	11,498	8.81		No Hit Found									No Hit Found							
m767R	286794--287156	121	13,630	10.82		No Hit Found									No Hit Found							
m768L	287392--287159	78	8,796	9.11		No Hit Found									No Hit Found							
m769R	287562--287771	70	8,318	10.27		No Hit Found									No Hit Found							
m771R	287942--288343	134	15,317	10.40		No Hit Found									No Hit Found							
m772L	288195--287998	66	7,467	7.04		No Hit Found									No Hit Found							
m774R	288510--288707	66	7,648	9.97		No Hit Found									No Hit Found							
m775L	288923--288570	118	13,206	9.29		No Hit Found								1	NP_048989	A633R	157.53	1.03E-37	61%	80%	1--114	2--115
m776L	288978--288706	91	10,642	9.47		No Hit Found									No Hit Found							
m778R	289113--289316	68	7,759	9.64		No Hit Found									No Hit Found							
m779R	289217--290179	321	35,384	7.69		No Hit Found									No Hit Found							
m780R	289629--289868	80	8,856	11.97		No Hit Found									No Hit Found							

Gene Name	Genome Position	A.A. length	Peptide Mw	pI	CDD Hit Number	COGs	COG Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from to	BLASTp Hit Number	Hit Accession	BLASTp Definition	Bit Score	E-value	% Identity	% Positive	Query from-to	Hit from-to
m781R	290315--290689	125	13,025	10.01		No Hit Found								1	NP_048987 a631L		51.22	1.05E-05	41%	62%	64--125	4--65
m782R	290843--291052	70	7,387	4.20		No Hit Found									No Hit Found							
m783R	291039--291290	84	9,703	8.84		No Hit Found									No Hit Found							
m784R	291512--291961	150	16,785	9.69		No Hit Found									No Hit Found							
m786L	291834--291589	82	9,155	12.29		No Hit Found									No Hit Found							
m787R	292539--292772	78	8,637	11.76		No Hit Found									No Hit Found							
m788L	292928--292716	71	8,080	11.04		No Hit Found									No Hit Found							
m790R	293075--293305	77	9,084	10.37		No Hit Found									No Hit Found							
m792L	293804--293382	141	15,179	11.22		No Hit Found									No Hit Found							
m793L	294470--294060	137	15,030	8.52		No Hit Found									No Hit Found							
m795L	295699--295085	205	23,601	8.18		No Hit Found									No Hit Found							
m796L	295692--295486	69	7,690	12.41		No Hit Found									No Hit Found							
m798L	296146--295814	111	12,491	8.64		No Hit Found									No Hit Found							
m799L	296174--295845	110	11,651	7.72		No Hit Found									No Hit Found							
m800L	296908--296483	142	16,811	9.76		No Hit Found									No Hit Found							
m802L	297748--297254	165	19,348	6.86		No Hit Found									No Hit Found							
m804R	297846--298169	108	12,010	7.43		No Hit Found									No Hit Found							
m805L	298566--298246	107	12,559	11.95		No Hit Found									No Hit Found							
m806R	298743--299015	91	7,840	4.11		No Hit Found									No Hit Found							
m808L	299794--299570	75	8,363	11.25		No Hit Found									No Hit Found							
m809L	300075--299881	65	7,514	12.15		No Hit Found									No Hit Found							
m812R	300587--301033	149	17,652	9.95		No Hit Found									No Hit Found							
m815R	302491--302688	66	7,546	10.94		No Hit Found									No Hit Found							
m816R	302543--302770	76	8,360	11.88		No Hit Found									No Hit Found							
m821L	304786--304520	89	10,147	9.99		No Hit Found									No Hit Found							
m823L	305535--305107	143	13,300	8.84		No Hit Found									No Hit Found							
m824L	305839--305423	139	16,292	8.64		No Hit Found									No Hit Found							
m825L	306579--306361	73	8,114	8.84		No Hit Found								1	NP_048675 a319L		64.31	1.21E-09	48%	54%	1--73	1--82
m826L	307347--306856	164	19,181	9.63		No Hit Found									No Hit Found							
m827L	307655--307446	70	7,766	6.76		No Hit Found									No Hit Found							
m830R	308525--308809	95	10,948	11.00		No Hit Found									No Hit Found							
m831R	308602--308862	87	10,140	9.67		No Hit Found									No Hit Found							
m832R	309017--309241	75	8,779	8.93		No Hit Found									No Hit Found							
m834L	309526--309287	80	9,021	9.23		No Hit Found									No Hit Found							
m835R	309618--309971	118	13,226	5.06		No Hit Found									No Hit Found							
m836R	309842--310261	140	16,555	12.54		No Hit Found									No Hit Found							
m837L	310068--309865	68	7,845	12.56		No Hit Found									No Hit Found							
m839R	310328--310765	146	16,481	5.10		No Hit Found									No Hit Found							
m840L	311207--310986	74	8,032	7.02		No Hit Found									No Hit Found							
m841R	311031--311405	125	13,988	9.87		No Hit Found									No Hit Found							
m843L	312146--311544	201	21,987	11.12		No Hit Found									No Hit Found							