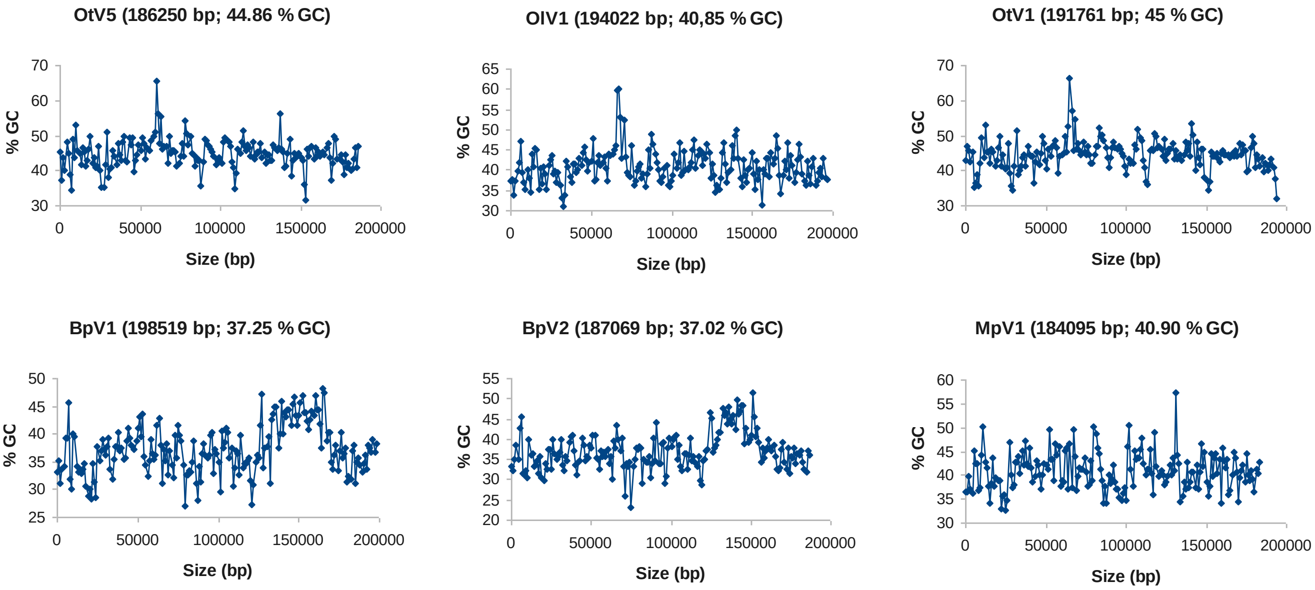
Figure S1. G+C content of the six Prasinovirus genomes.



OtV5 (186250 bp; 44.86 % GC)

OlV1 (194022 bp; 40,85 % GC)

OtV1 (191761 bp; 45 % GC)

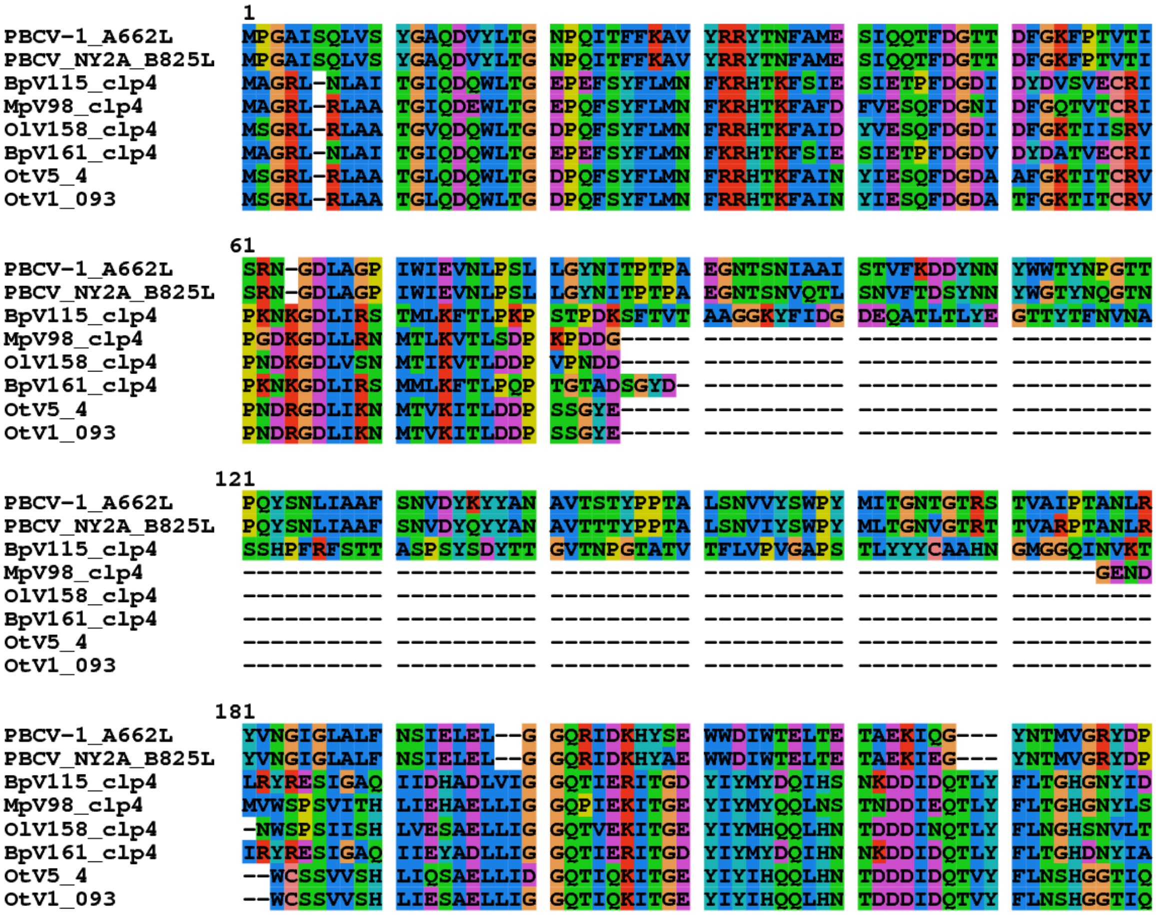
BpV1 (198519 bp; 37.25 % GC)

BpV2 (187069 bp; 37.02 % GC)

MpV1 (184095 bp; 40.90 % GC)

The G + C profile represents a slipping window of 1000 bp

Figure S2. Alignment of the first 240 amino acids of predicted clp4 proteins with their putative orthologs in Chlorella viruses



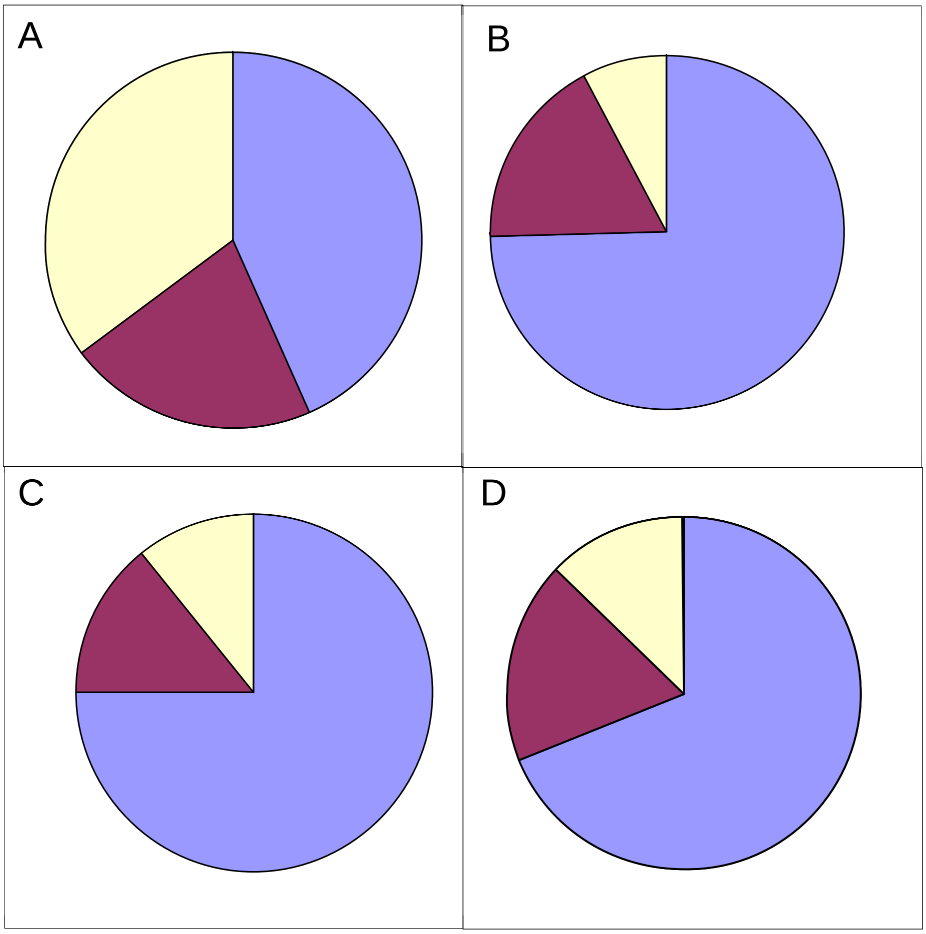
```
                   1
PBCV-1_A662L       MPGAISQLVS YGAQDVYLTG NPQITFFKAV YRRYTNFAME SIQQTFDGTT DFGKFPTVTI
PBCV_NY2A_B825L    MPGAISQLVS YGAQDVYLTG NPQITFFKAV YRRYTNFAME SIQQTFDGTT DFGKFPTVTI
BpV115_clp4        MAGRL-NLAI TGIQDQWLTG EPEFSYFLMN FKRHTKFSIE SIETPFDGDI DYDVSVECRI
MpV98_clp4         MAGRL-RLAA TGIQDEWLTG EPQFSYFLMN FKRHTKFAFD FVESQFDGNI DFGQTVTCRI
OlV158_clp4        MSGRL-RLAA TGVQDQWLTG DPQFSYFLMN FRRHTKFAID YVESQFDGDI DFGKTIISRV
BpV161_clp4        MAGRL-NLAI TGIQDQWLTG EPEFSYFLMN FKRHTKFSIE SIETPFDGDV DYDATVECRI
OtV5_4             MSGRL-RLAA TGLQDQWLTG DPQFSYFLMN FRRHTKFAIN YIESQFDGDA AFGKTITCRV
OtV1_093           MSGRL-RLAA TGLQDQWLTG DPQFSYFLMN FRRHTKFAIN YIESQFDGDA TFGKTITCRV

                  61
PBCV-1_A662L       SRN-GDLAGP IWIEVNLPSL LGYNITPTPA EGNTSNIAAI STVFKDDYNN YWWTYNPGTT
PBCV_NY2A_B825L    SRN-GDLAGP IWIEVNLPSL LGYNITPTPA EGNTSNVQTL SNVFTDSYNN YWGTYNQGTN
BpV115_clp4        PKNKGDLIRS TMLKFTLPKP STPDKSFTVT AAGGKYFIDG DEQATLTLYE GTTYTFNVNA
MpV98_clp4         PGDKGDLLRN MTLKVTLSDP KPDDG----- ---------- ---------- ----------
OlV158_clp4        PNDKGDLVSN MTIKVTLDDP VPNDD----- ---------- ---------- ----------
BpV161_clp4        PKNKGDLIRS MMLKFTLPQP TGTADSGYD- ---------- ---------- ----------
OtV5_4             PNDRGDLIKN MTVKITLDDP SSGYE----- ---------- ---------- ----------
OtV1_093           PNDRGDLIKN MTVKITLDDP SSGYE----- ---------- ---------- ----------

                 121
PBCV-1_A662L       PQYSNLIAAF SNVDYKYYAN AVTSTYPPTA LSNVVYSWPY MITGNTGTRS TVAIPTANLR
PBCV_NY2A_B825L    PQYSNLIAAF SNVDYQYYAN AVTTTYPPTA LSNVIYSWPY MLTGNVGTRT TVARPTANLR
BpV115_clp4        SSHPFRFSTT ASPSYSDYTT GVTNPGTATV TFLVPVGAPS TLYYYCAAHN GMGGQINVKT
MpV98_clp4         ---------- ---------- ---------- ---------- ---------- ------GEND
OlV158_clp4        ---------- ---------- ---------- ---------- ---------- ----------
BpV161_clp4        ---------- ---------- ---------- ---------- ---------- ----------
OtV5_4             ---------- ---------- ---------- ---------- ---------- ----------
OtV1_093           ---------- ---------- ---------- ---------- ---------- ----------

                 181
PBCV-1_A662L       YVNGIGLALF NSIELEL--G GQRIDKHYSE WWDIWTELTE TAEKIQG--- YNTMVGRYDP
PBCV_NY2A_B825L    YVNGIGLALF NSIELEL--G GQRIDKHYAE WWDIWTELTE TAEKIEG--- YNTMVGRYDP
BpV115_clp4        LRYRESIGAQ IIDHADLVIG GQTIERITGD YIYMYDQIHS NKDDIDQTLY FLTGHGNYID
MpV98_clp4         MVWSPSVITH LIEHAELLIG GQPIEKITGE YIYMYQQLNS TNDDIEQTLY FLTGHGNYLS
OlV158_clp4        -NWSPSIISH LVESAELLIG GQTVEKITGE YIYMHQQLHN TDDDINQTLY FLNGHSNVLT
BpV161_clp4        IRYRESIGAQ IIEYADLLIG GQTIERITGD YIYMYDQIHN NKDDIDQTLY FLTGHDNYIA
OtV5_4             --WCSSVVSH LIQSAELLID GQTIQKITGE YIYIHQQLHN TDDDIDQTVY FLNSHGGTIQ
OtV1_093           --WCSSVVSH LIESAELLIG GQTIQKITGE YIYIHQQLHN TDDDIDQTVY FLNSHGGTIQ
```

Figure S3. Categories of common and specific Prasinovirus genes



Distribution (%) of common and specific Prasinovirus genes. “Putative Identified Genes, PIG” 🟨, “Putative Unknown Genes, PUG” 🟥 and “Putative Orphan Genes, POG” 🟦 for the 125 CDSs common to the six Prasinoviruses (A), the *Bathycoccus* (B), the *Micromonas* (C) and the *Ostreococcus* (D) virus specific CDSs.

Figure S4. Pairwise amino acid identities of the six Prasinovirus proteomes

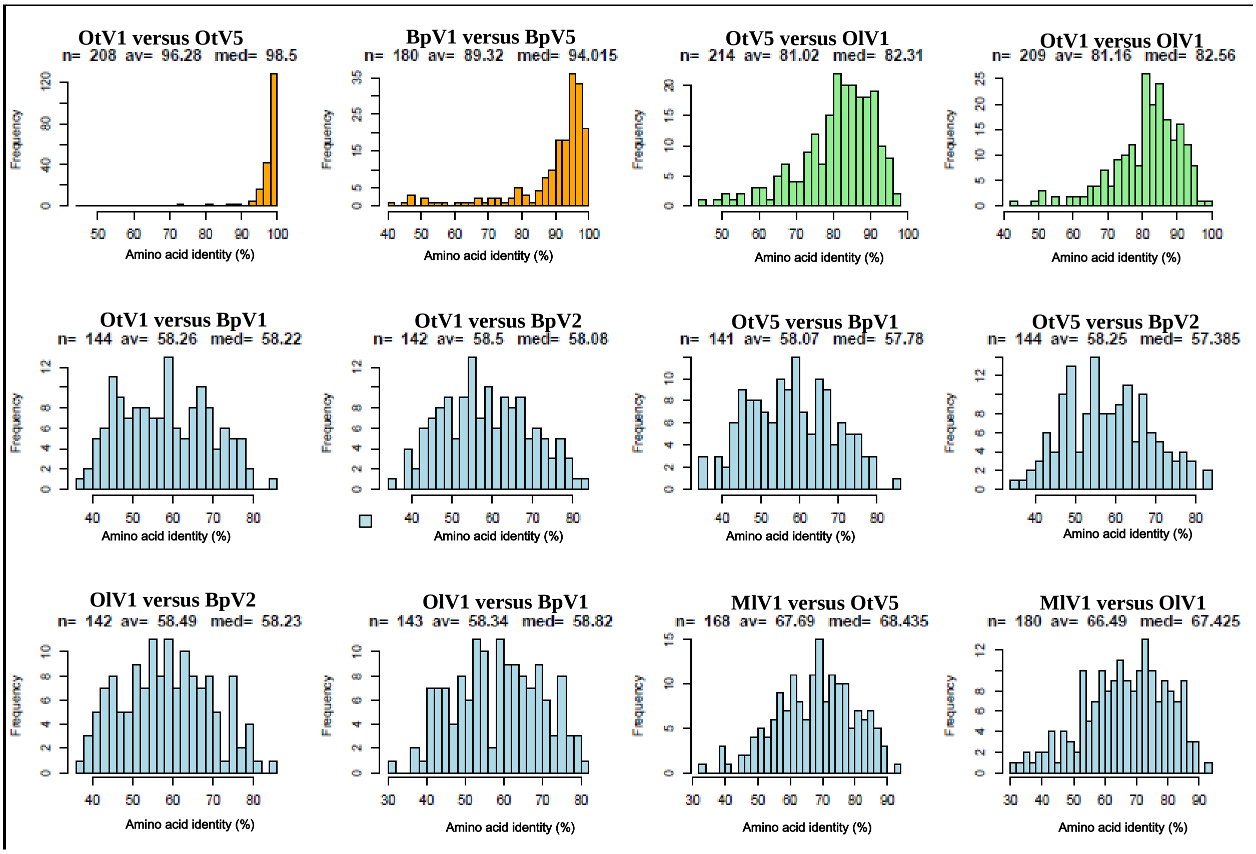


Figure S5. Phylogenies of Acetolactate synthase, 3-Dehydroquinate synthase and Asparagine synthase Prasinovirus genes with their best blast hits.

A. Acetolactate synthase **B. 3-Dehydroquinate synthase** **C. Asparagine synthase**

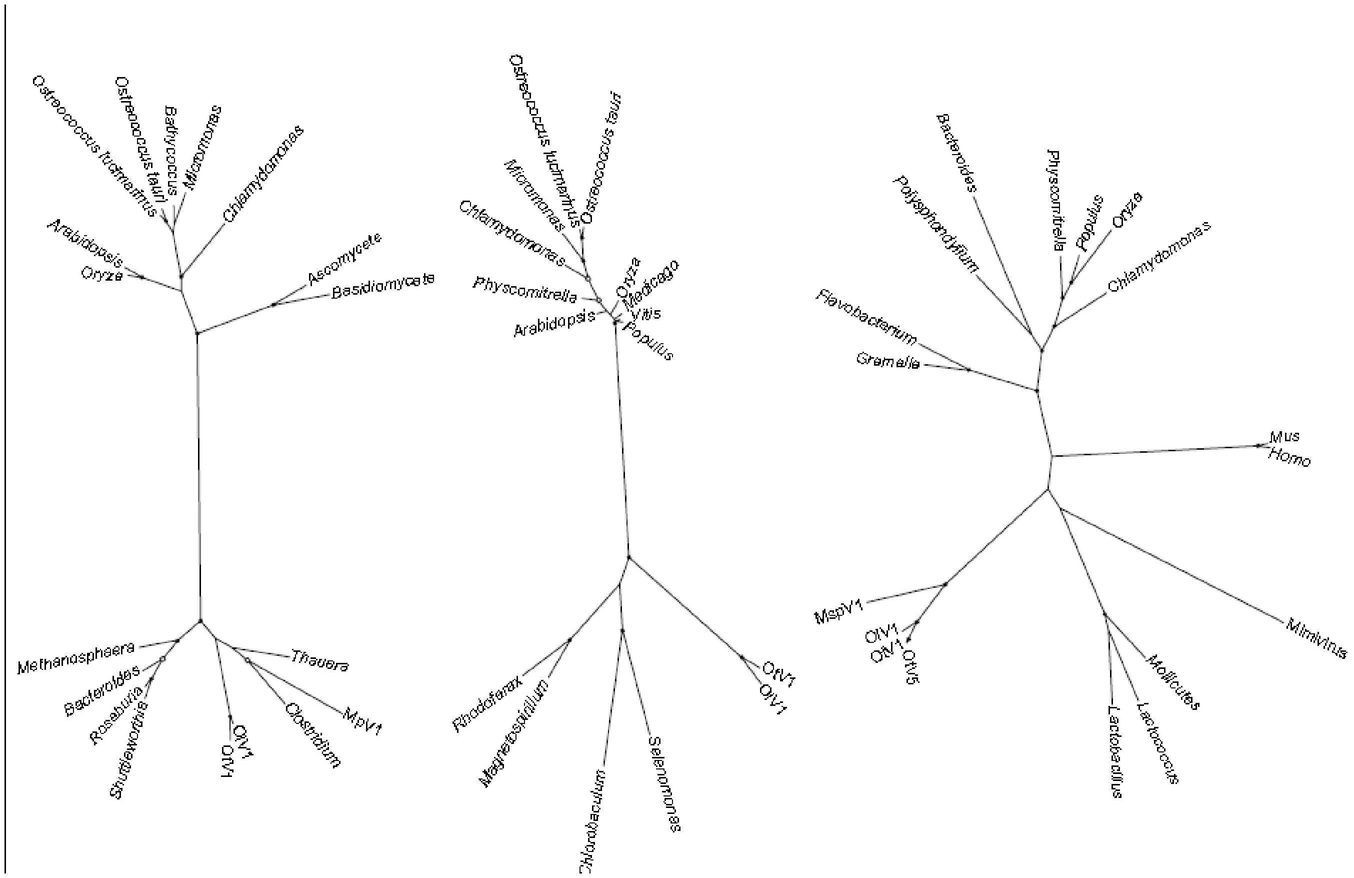
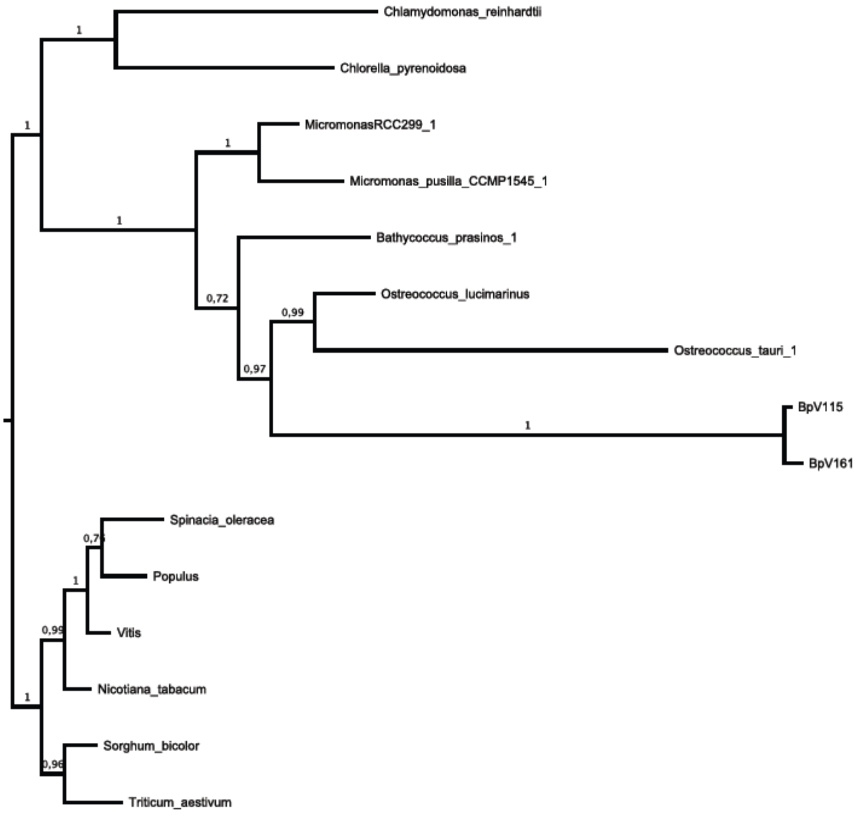
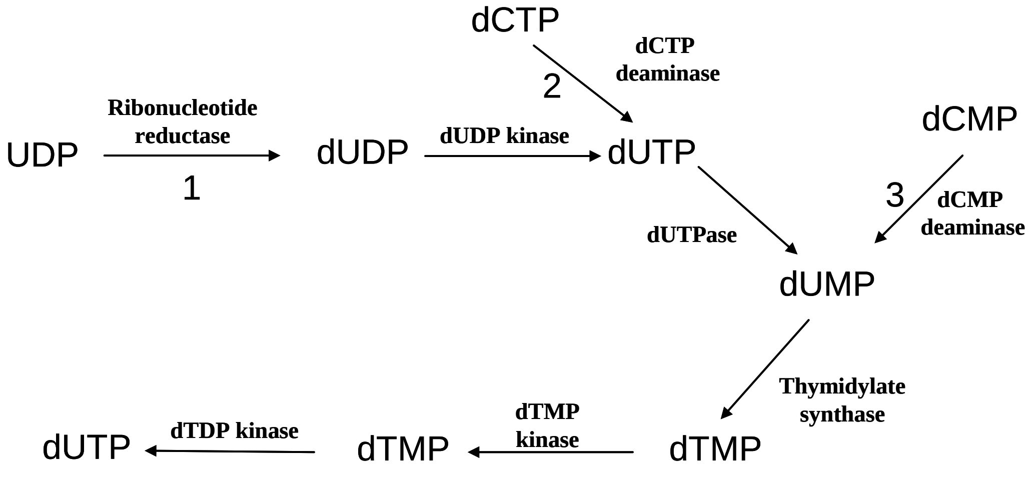


Figure S6. Phylogenies of genus-specific PIGs (putative identified genes)

A. Phylogenetic analysis of the BpV1 and BpV2 HSP70 sequences



B. The three known pathways for synthesis of dUMP, an intermediate in dTT synthesis



(A) - The ***hsp70*** gene in *Bathycoccus* viruses was probably acquired from its host (Fig. S6A). Heat-shock proteins (HSPs) participate in diverse functions including: (a) protein folding, (b) vesicular transport, (c) protein aggregation and (e) cellular signaling. HSP70 chaperones regulate RNA replication for Tombus- (1) or Closteroviruses (2). Many other phylogenetically diverse viruses have recruited an hsp70 gene, (ranging from bacteriophages such as P1 and λ, through RNA viruses such as HIV, to large DNA viruses such as herpesviruses and mimiviruses (see 3 for a review).

(B) – dCMP **deaminase,** found in all of the known Chlorellavirus genomes, is encoded by MpV1 and OlV1 but is absent in the other Prasinoviruses (Fig. S6B). In contrast the other route represented by ribonucleotide reductase (coded by 432 dsDNA viruses) is present in all Prasinoviruses. Chlorella virus CMP deaminase has a double activity, i.e. the classical dCMP deaminase activity producing dUMP, but also dCTP deaminase activity producing dUTP. The most conserved domain of dCMP deaminases, which forms the catalytic site, is conserved, as well as the amino acid residues of the zinc ion-binding site (4).

References

1. Yung-Liang W-R., Stork J, Nagy P-D (2009) A key role for heat shock protein 70 in the localization and insertion of Tombusvirus replication proteins to intracellular membranes. J Virology 83: 3276-3287.

2. Dolja V-V, Kreuze JF, Valkonen J-P (2006) Comparative and functional genomics of closteroviruses.Virus Res 117:38-51.

3. Mayer M-P (2005) Recruitment of Hsp70 chaperones: a crucial part of viral survival strategies. Rev Physiol Biochem Pharmacol 153:1-46.

4. Graziani S, Bernauer J, Skouloubris S, Graille M, Zhou C-Z, Marchand C, Decottignies P, van Tilbeurgh H, Myllykallio H, Liebl U (2006) Catalytic mechanism and structure of viral flavin-dependent thymidylate synthase ThyX. J Biol Chem.281:24048-24057.

Table S1. Genomic densities of Prasinovirus core ORFs

Position (Kb)	BpV1	BpV2	MpV1	OlV1	OtV1	OtV5
10	2	2	2	2	2	4
20	11	10	9	4	8	8
30	5	6	1	8	4	2
40	7	7	6	2	4	6
50	5	5	6	9	7	5
60	6	9	5	5	5	5
70	5	11	5	1	5	7
80	13	9	11	8	8	11
90	7	5	11	13	15	12
100	11	10	5	13	3	1
110	9	10	14	9	11	11
120	10	11	10	11	7	9
130	9	7	10	6	9	8
140	2	0	7	8	9	9
150	0	0	9	6	4	7
160	0	10	3	7	8	7
170	3	8	6	3	7	7
180	11	5	5	7	4	4
190	8			3	5	2
200	1					

Values in this table represent the number of ORFs determined every 10 Kb for each viral genome.

Table S2. Identified genes in Prasino- and Chlorella-viruses

Gene names	**Bpv1**	**BpV2**	**MpV1**	**OlV1**	**OtV1**	**OtV5**	**PBCV-1**	**NY-2A**	**AR 158**	**MT 325**	**FR 483**	**ATCV-1**
DNA, replication, recombination and repair												
DNA polymerase	1	1	1	1	1	1	1	1	1	1	1	1
Archaeo-eukaryotic primase	0	0	0	0	0	0	1	1	1	1	1	1
PCNA	1	1	1	1	1	1	2	2	2	2	2	2
Replication factor C	0	0	0	0	0	0	1	1	1	1	1	1
Rnase H	1	1	1	1	1	1	1	1	1	1	1	1
Helicase	1	1	1	1	1	1	1	1	1	1	1	1
DNA topoisomerase	1	1	1	1	2	2	1	1	1	1	1	1
ATPase (PP-loop)	0	0	0	0	0	0	1	1	1	1	1	1
ATPase (DNA packaging)	1	1	1	1	1	1	1	1	1	1	1	1
Pyrimidine dinner-specific glycosylase	0	0	0	0	0	0	1	1	1	1	1	1
ATP dependant DNA ligase	1	1	1	1	1	1	1	1	1	0	0	1
Exonuclease	2	2	2	2	2	2	1	1	1	1	1	1
NTPase/helicase	1	1	1	1	1	1	0	0	0	0	0	0
Transcription												
Transcription factor IIB	1	1	1	1	1	1	1	1	1	1	1	1
Transcription factor IID	0	0	0	0	0	0	1	1	1	1	1	1
Transcription factor IIS	1	1	1	1	1	1	1	1	1	1	1	1
Transcription activator /SWI/SNF helicase	1	1	1	1	1	1	1	1	1	1	1	1
Tata-box Binding Protein (TBP)	1	1	1	1	1	1	0	0	0	0	0	0
Superfamily II helicase	0	0	0	0	0	0	3	3	3	3	3	1
mRNA guanylyltransferase/mRNA capping enzyme	2	2	2	2	2	2	1	1	1	1	1	1
RNA triphosphatase	0	0	0	0	0	0	1	1	1	1	1	1
Histone H3, Lys27 methylase	0	0	0	0	0	0	1	2	2	1	1	1
RNase III	1	1	1	1	1	1	1	1	1	1	1	1
Cytosine deaminase	0	0	0	0	0	0	1	1	1	1	1	1
Sugar manipulation enzymes												
GDP-D-mannose dehydratase epimerase	1	1	1	1	1	1	1	1	1	0	0	1
6-Phospho fructokinase	0	0	1	1	1	1	0	0	0	0	0	0
α1, 2-galactosyl-transferase	0	0	1	0	1	1	0	0	0	0	0	0
glycosyltransferase	3	3	3	6	6	6	0	0	0	0	0	0
UDP-glucose 6-deshydrogenase	0	0	0	0	0	0	1	1	2	1	1	1
Mannose-6P- isomerase	0	0	0	0	0	0	0	0	0	0	0	1
Glucosamine synthase	0	0	0	0	0	0	1	1	1	1	1	0
Hyaluronan synthase	0	0	0	0	0	0	1	0	0	1	1	0
dTDP-D-glucose 4,6-dehydratase	0	0	1	2	2	2	0	0	0	0	0	1
Chitin synthase	0	0	0	0	0	0	0	2	3	0	0	0
Cellulase precursor	0	0	0	0	0	0	0	0	1	1	1	1
dTDP-glucose pyrophosphorylase	0	0	0	0	0	0	0	0	0	1	1	0
Polysaccaride deacetylase	0	0	0	0	0	0	0	1	1	0	0	0
DNA methylation and site-specific endonucleases												
Adenine DNA methyltransferase	1	1	1	2	3	3	2	11	9	0	1	0
Cytosine DNA methyltransferase	0	0	0	0	0	0	3	7	7	1	2	1
Methyltransferase FkbM	0	0	0	1	1	1	0	1	1	1	1	0
cytosine-C5 specific DNA methylase	0	0	1	0	1	0	0	0	0	0	0	0

DNA methylase,N6_N4_Mtase	0	0	0	1	0	0	0	0	0	0	0	0
DNA restriction endonuclease	0	0	0	0	0	0	2	2	2	0	0	0
Cell wall degradation												
Chitinase	0	0	0	0	0	0	2	1	1	2	1	2
Chitosanase	0	0	0	0	0	0	1	1	1	1	1	1
Beta, alpha 1,4- glucuronic lyase	0	0	0	0	0	0	1	2	2	1	1	1
Beta-1,3-glucanase	0	0	0	0	0	0	1	1	1	1	1	1
α-L-arabino furanosidase	0	0	0	0	0	0	0	0	0	0	0	1
Nucleotide metabolism												
Ribonucleotide reductase Lsubunit	1	1	1	1	1	1	1	1	1	1	1	1
Ribonucleotide reductase Ssubunit	1	1	1	1	1	1	1	1	1	1	1	1
Thymidylate S	1	1	1	1	1	1	1	1	1	1	1	1
Thymidine Kinase	1	1	1	1	1	1	0	0	0	0	0	0
dUTP pyrophosphatase	1	1	1	1	1	1	1	1	1	1	1	1
Aspartate transcarbamylase	0	0	0	0	0	0	1	1	1	0	0	0
Deoxynucleoside kinase	0	0	0	0	0	0	1	0	0	1	1	1
Glutaredoxin	0	0	0	0	0	0	1	1	1	2	2	2
Thioredoxin	0	0	0	0	0	0	1	1	1	3	3	2
Ribonucleotide-triphosphate reductase	0	0	0	0	0	0	0	0	0	0	0	1
dCMP deaminase	0	0	0	0	0	0	1	1	1	1	1	1
CMP/dCMP deaminase zinc-binding	0	0	1	1	0	0	0	0	0	0	0	0
Protein synthesis and degradation												
Acetolactate synthase (2-2-1-6) (TPP)	0	0	1	1	1	1	0	0	0	0	0	0
4-hydroxy-2-oxovalerate aldolase	0	0	1	0	0	0	0	0		0	0	0
acetaldehyde dehydrogenase, MhpF	0	0	1	0	0	0	0	0	0	0	0	0
3-Dehydroquinate synthase	0	0	0	1	1	1	0	0	0	0	0	0
Oxoacyl (acyl-carrier protein) dehydrogenase	0	0	0	1	1	1	0	0	0	0	0	0
Aminotransferase family protein	2	0	1	1	1	1	0	0	0	0	0	0
Asparagine synthase	0	0	1	1	1	1	0	0	0	0	0	0
NAD-dependant epimerase dehydratase	0	0	1	1	1	1	0	0	0	0	0	0
N-myristoyl transferase	1	1	0	1	1	1	0	0	0	0	0	0
HSP70	1	1	0	0	0	0	0	0	0	0	0	0
Prolyl 4-hydroxylase	1	1	1	1	1	1	1	1	1	1	1	1
33 kDa in vitro translation peptide	2	2	3	2	2	2	0	0	0	0	0	0
Translation elongation factor-3	0	0	0	0	0	0	2	0	1	2	1	1
Thiol oxydoreductase	0	0	0	0	0	0	1	1	1	1	1	1
Protein disulfide isomerase	0	0	0	0	0	0	1	1	1	1	1	1
SKP-1 protein	0	0	0	0	0	0	1	1	1	1	1	1
Ubiquitin	0	0	0	0	0	0	0	1	0	0	0	1
Ring finger ubiquitin ligase	0	0	0	0	0	0	1	1	1	1	1	1
proline dehydrogenase	0	0	0	1	1	1	0	0	0	0	0	0
TRP domain protein	0	0	0	1	1	1	0	0	0	0	0	0
2OG-Fe(II) oxygenase	1	1	1	1	1	1	0	0	0	0	0	0
Aspartyl/Asparaginyl beta-hydrolase	1	1	1	2	2	1	0	0	0	0	0	0
Ubiquitin hydrolase-like cystein peptidase	1	1	1	1	1	1	1	1	1	1	1	1
FtsH (3-4-24) metalloendopeptidase	1	1	1	1	1	1	0	0	0	0	0	0
Zinc metallo protease	1	1	0	1	1	1	2	2	2	1	1	1
Initiation factor 2	0	0	0	0	0	0	0	0	0	0	1	1
ATP-dependant protease proteolytic subunit	1	1	1	1	1	1	0	0	0	0	0	0
Lipid – Fatty acid metabolism												

UDP-N-acetylglucosamine O-acyltransferase	1	0	0	0	0	0	0	0	0	0	0	0
Lipopolysaccharide biosynthesis protein rf6H	0	0	0	0	0	1	0	0	0	0	0	0
Acetyltransferase	0	0	0	0	0	0	1	0	0	0	0	0
N-acetyltransferase	0	0	0	0	0	0	1	1	1	1	1	1
Lipoprotein lipase	0	0	0	0	0	0	1	1	1	1	1	1
Lipase	0	0	0	1	0	0	0	0	0	0	0	0
Glycerophosphoryl diesterase	0	0	0	0	0	0	1	1	1	0	0	1
Lysophospho-lipase	0	0	0	0	0	0	1	1	1	0	0	1
Patatin-like phospholipase	1	1	1	1	1	1	1	1	1	1	1	1
Signaling												
Potassium channel protein, PhoH	1	1	1	1	1	1	1	1	1	1	0	1
Phosphate transporter family, Pho4	1	0	0	1	0	0	0	0	0	0	0	0
Ligand-gated channel protein	0	0	0	0	0	0	1	0	0	0	1	1
Glutamate receptor	0	0	0	0	0	0	1	0	0	0	0	0
Ca2+ transporting ATPase	0	0	0	0	0	0	0	0	1	1	0	0
Aquaglyceroporin	0	0	0	0	0	0	0	0	0	1	0	1
Rhodanese-like domain protein	0	0	1	1	1	1	1	1	1	1	1	1
Serine/Threonine protein kinase	1	1	1	1	1	1	8	6	7	6	6	5
Miscellaneous												
Ornithine:Arginine decarboxylase	0	0	0	0	0	0	1	1	1	1	1	1
Agmatine iminohydrolase	0	0	0	0	0	0	1	1	1	1	0	1
N-carbamoyl amidohydrolase	0	0	0	0	0	0	1	1	1	1	1	1
Homospermidine synthase	0	0	0	0	0	0	1	1	1	1	1	1
Histidine decarbo1ylase	0	0	0	0	0	0	1	1	1	1	1	1
Monoamine o1idase	0	0	0	0	0	0	1	1	1	1	1	1
Amidase	0	0	0	0	0	0	1	1	1	1	1	1
mannitol dehydrogenase	0	0	1	0	0	0	0	0	0	0	0	0
CuZn-superoxyde dismutase	0	0	1	0	0	0	1	0	0	1	1	1
O-methyl transferase	0	0	0	0	0	0	1	0	0	0	0	0
ABC transporter protein	0	0	0	0	0	0	1	1	1	1	1	1
ATPase (AAA+ class)	0	0	0	0	0	0	1	1	1	0	0	0
Erithrocyte binding protein	0	0	0	0	0	0	0	0	0	0	0	1
Mucin-desulfating sulfatase	0	0	0	0	0	0	0	0	0	0	0	1
Hep_Hag family protein	0	0	0	1	1	1	0	0	0	0	0	0
ABC-1 domain protein	1	1	1	1	1	1	0	0	0	0	0	0
Fibronectin binding protein	0	0	0	1	1	0	1	0	0	1	1	1
echinonectin	0	0	0	0	1	1	0	0	0	0	0	0
ATP/GTP binding site motif A AGB-1	1	1	1	1	1	1	0	0	0	0	0	0
Viral A-type inclusion protein	1	1	1	1	1	1	0	0	0	0	0	0
Capsid-like protein	7	7	8	8	8	8	5	6		5		
tRNA												
Lys-tRNA	0	0	0	0	0	0	3	1	0	1	1	1
Leu-tRNA	1	1	1	0	0	0	2	2	2	1	1	1
Arg-tRNA	0	0	0	0	0	0	1	1	1	1	1	1
Val-tRNA	0	0	0	0	0	0	1	1	1	0	0	2
Ile-tRNA	1	1	1	1	1	1	1	0	1	1	1	0
Gln-tRNA	0	0	1	1	1	1	0	0	0	1	1	0
Tyr-tRNA	0	1	1	1	0	1	1	1	0	1	1	1
Asn-tRNA	1	1	1	1	1	1	2	1	1	2	2	2
Thr-tRNA	0	0	1	1	0	1	0	0	0	1	1	0
Gly-tRNA	0	0	0	0	0	0	0	0	0	1	1	1

Phe-tRNA	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**1**	**0**
Ser-tRNA	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**
Asp-tRNA	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**1**
Integration and Transposition												
Transposase	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**6**	**4**	**0**	**0**	**0**
Resolvase	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**2**	**1**	**0**	**0**	**0**
Tlr 6Fp DNA mobile protein	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**1**	**1**	**1**	**1**	**1**
GIY-YIG endonuclease	**0**	**0**	**0**	**0**	**0**	**0**	**7**	**16**	**13**	**10**	**9**	**10**
HNH endonuclease	**0**	**0**	**0**	**0**	**0**	**0**	**6**	**18**	**4**	**4**	**6**	**7**