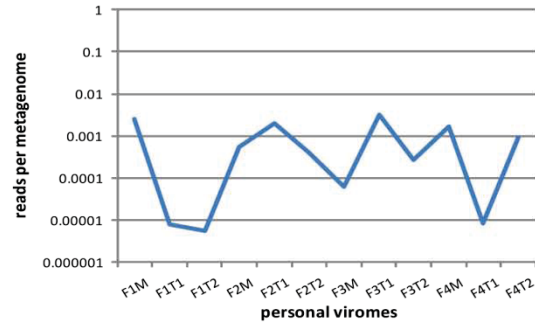
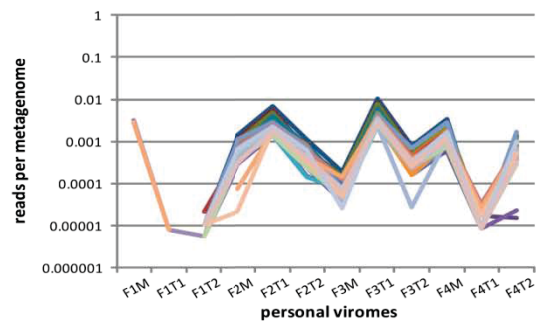


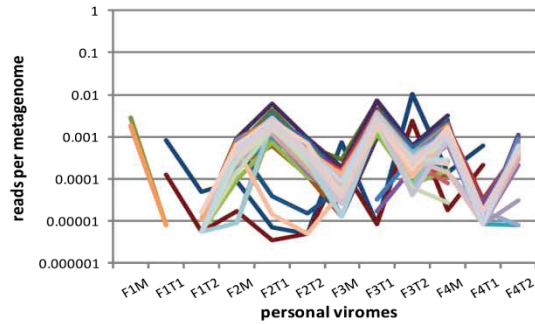
Depth profile of contig07548, present in 12 viromes



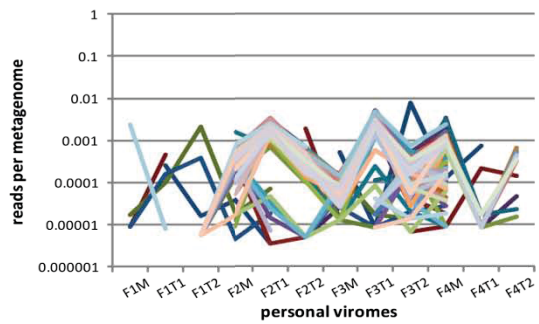
Depth profiles of 38 cross-contigs present in 11 viromes



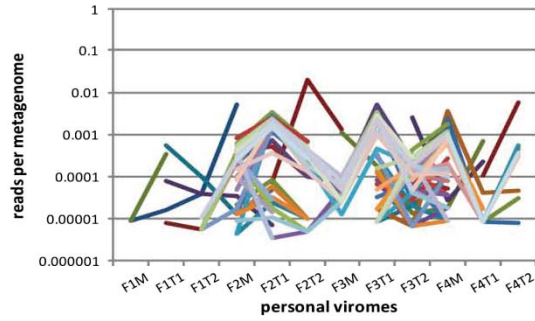
Depth profiles of 86 cross-contigs present in 10 viromes



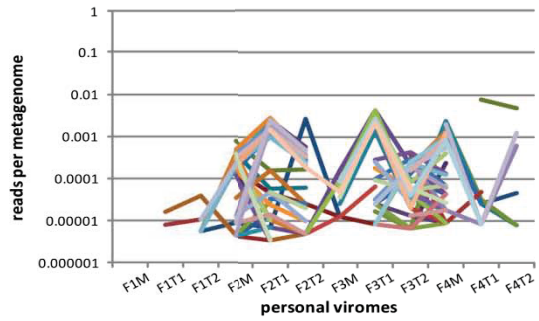
Depth profiles of 94 cross-contigs present in 9 viromes



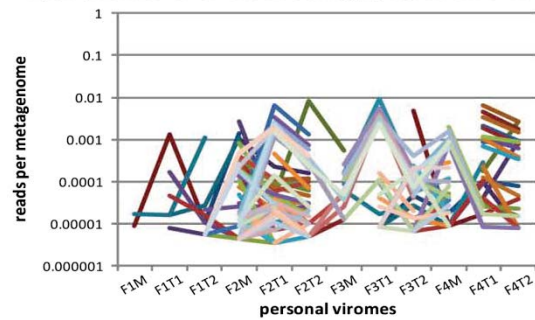
Depth profiles of 53 cross-contigs present in 8 viromes



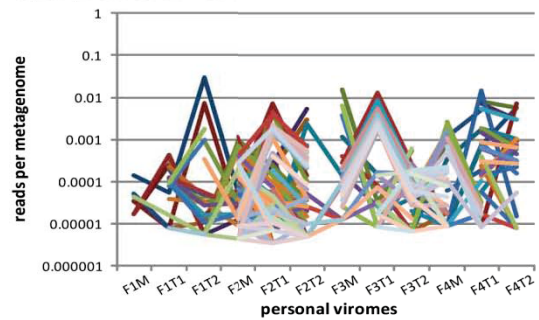
Depth profiles of 48 cross-contigs present in 7 viromes



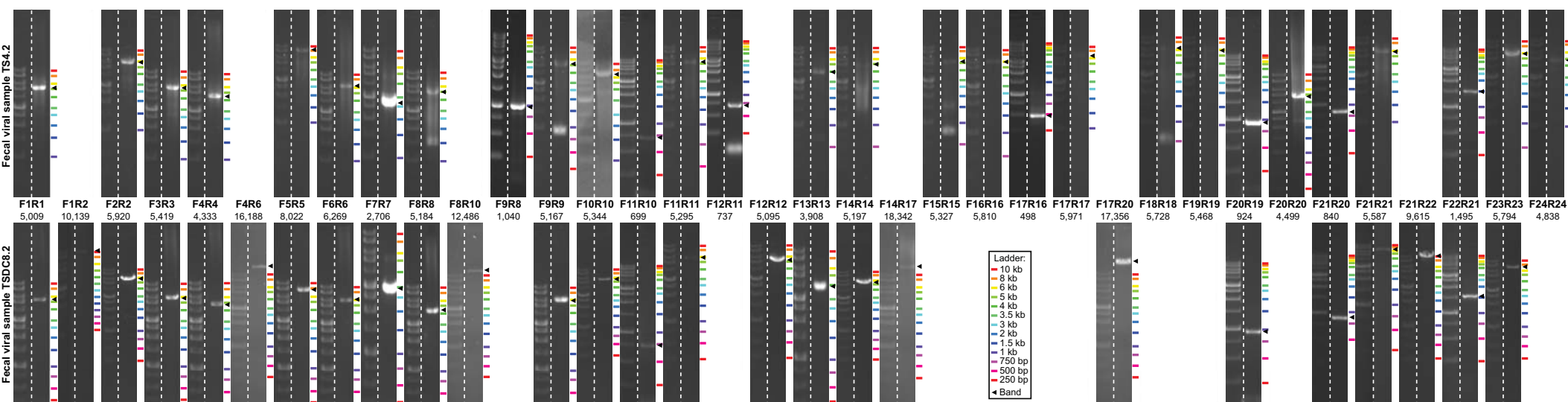
Depth profiles of 87 cross-contigs present in 6 viromes



Depth profiles of 165 cross-contigs present in 5 viromes

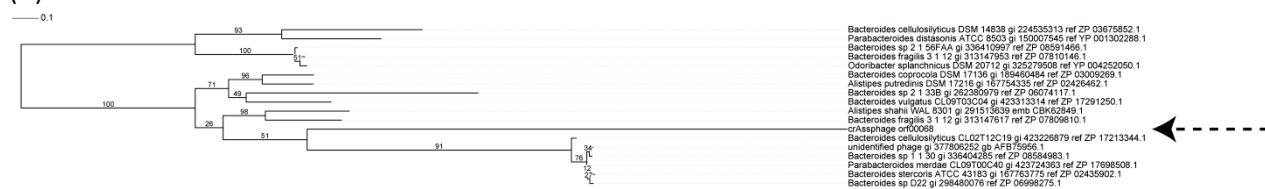


Supplementary Figure 1. Depth profiles of cross-contigs. Depth profiles of the cross-contigs containing reads from five through twelve samples, showing the highly correlating occurrence of the most ubiquitous cross-contigs. For details see Supplementary Data 1.



Supplementary Figure 2. Validation of *crAssphage* genome by long-range PCR using two unrelated fecal viral samples. All primer pairs and their expected sizes are indicated. Bands are indicated with a black arrow head showing products in the expected size range. Figure 1 and Supplementary File 13 show the genomic regions covered by these PCR products in green, Table 2 lists the primer locations. Supplementary Table 1 contains the regions of these amplicons that were sequenced by Sanger dideoxynucleotide sequencing.

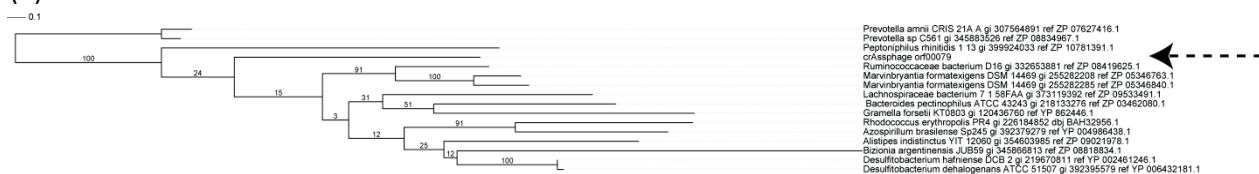
(A) orf00068



(B) orf00077



(C) orf00079

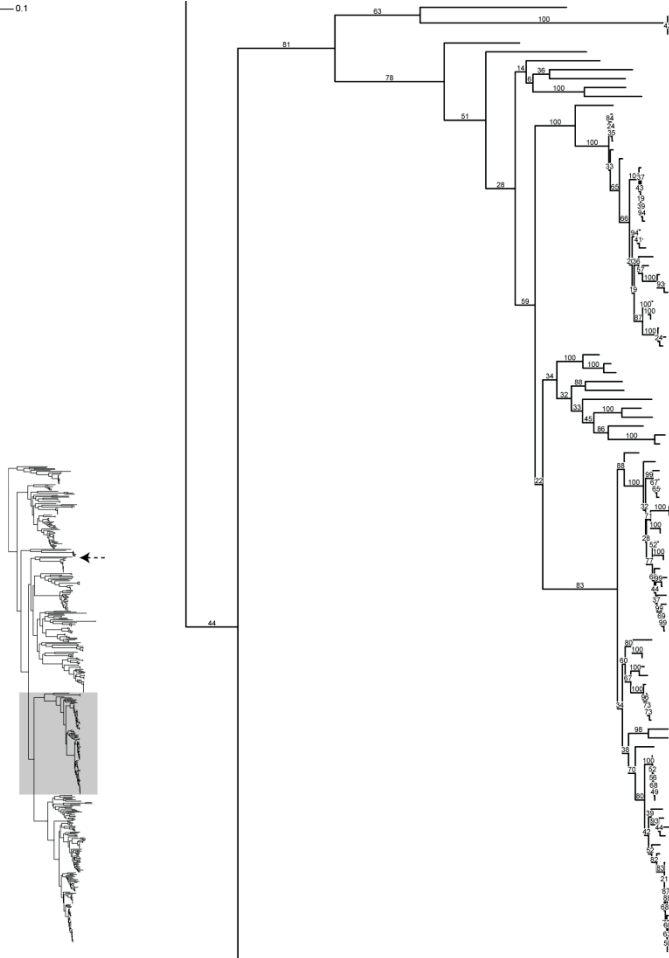


(D) orf00019



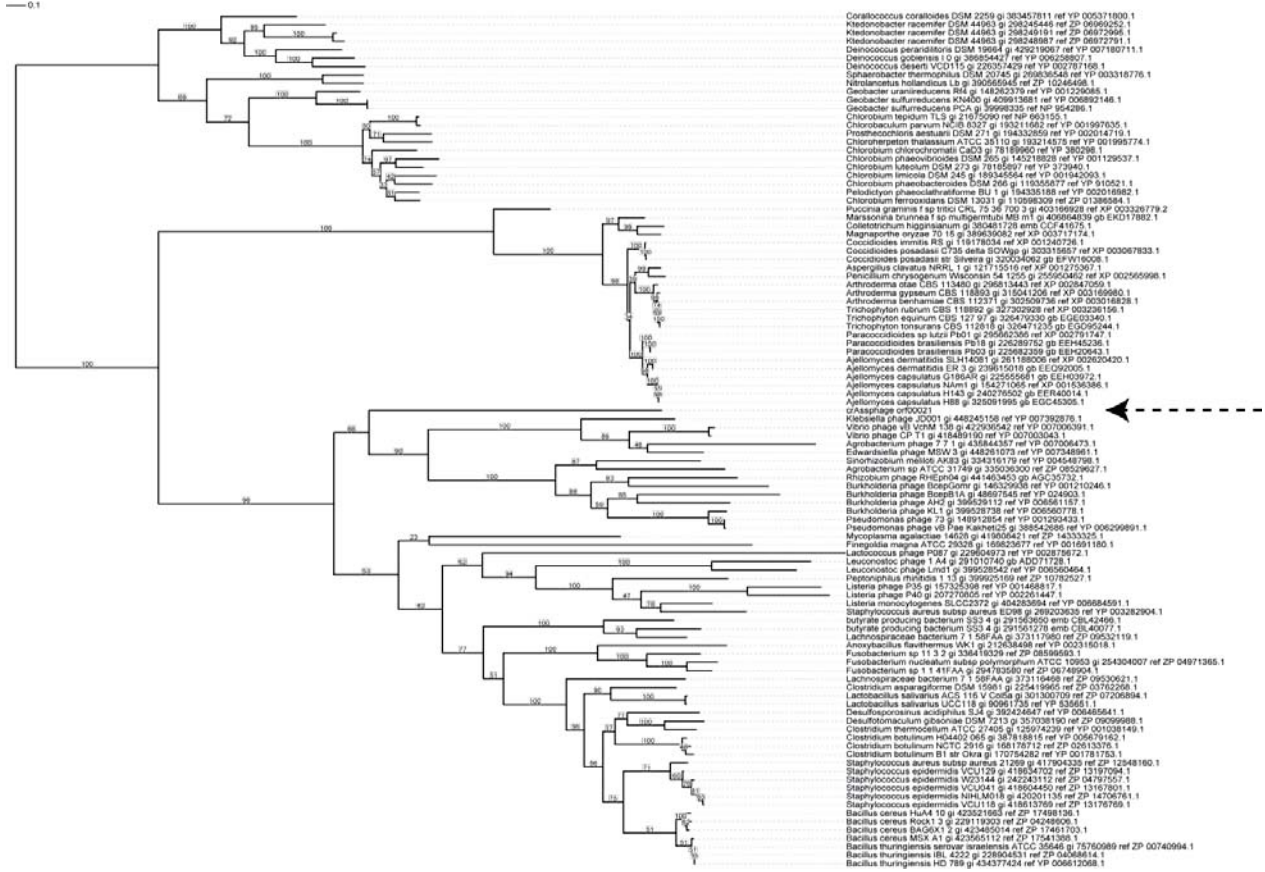
orf00019 (continued)

—0.1



- Bacillus sp BT1B CT2 g| 319645744 ref ZP 07999074.1
- Bacillus thuringiensis BL 4222 g| 22904921 ref ZP 04089975.1
- Bacillus thuringiensis serovar tenebrionis ATCC 35466 g| 57363556 ref ZP 007400420.1
- Bacillus thuringiensis HD 789 g| 434379334 ref YP 006813756.1
- Rhizobium sp CF80 g| 44204044 ref ZP 14744850.1
- Rhizobium leguminosarum bv trifolii WSM597 g| 424914240 ref ZP 18337604.1
- Phragmobacterium holotolerans S2 g| 337363368 ref YP 00489892.1
- Ochrobactrum intermedium M89 g| 444313172 ref ZP 21148726.1
- Bradyrhizobium sp STM 3843 g| 365895153 ref ZP 09433277.1
- Novosphingobium pentamethylotricans US6.1 g| 352891776 ref ZP 09190605.1
- Sphingobium yanukuyae XLDN2 5 g| 381203667 ref ZP 09910971.1
- Rhodovulum sp PH10 g| 402947653 ref ZP 10066271.1
- Bradyrhizobium sp ORS 278 g| 146342747 ref YP 001207795.1
- Bradyrhizobium sp STM 3859 g| 365895244 ref ZP 09430686.1
- Bradyrhizobium sp ORS 375 g| 365895096 ref ZP 09424112.1
- Bradyrhizobium sp ORS 269 g| 367478030 ref ZP 09478174.1
- Bradyrhizobium sp BTAI.1 g| 482623363 ref YP 001237968.1
- Bradyrhizobium sp STM 3843 g| 365895943 ref ZP 09434987.1
- Bradyrhizobium sp WSM471 g| 374457173 ref ZP 08650959.1
- Bradyrhizobium sp CCGE LA001 g| 421598305 ref ZP 16041755.1
- Bradyrhizobium sp YR581 g| 366201647 ref YP 005448610.1
- Bradyrhizobium japonicum USDA 6 g| 384216070 ref YP 005607236.1
- Bradyrhizobium sp S23321 g| 383769447 ref YP 005448610.1
- Bradyrhizobium japonicum USDA 110 g| 27382152 ref NP 773681.1
- Bradyrhizobium japonicum g| 21314952 g| 444M4137.1
- Nitrobacter hamptonensis g| 52115034 ref YP 075033.1
- Nitrobacter sp Nb311A g| 85713590 ref ZP 01044550.1
- Nitrobacter winogradskyi N2 g| 75074792 ref YP 317213.1
- Rhodospseudomonas palustris Bta53 g| 115523000 ref YP 779911.1
- Rhodospseudomonas palustris Bta18 g| 90422474 ref YP 530844.1
- Rhodospseudomonas palustris Bta5 g| 91978300 ref YP 57059.1
- Rhodospseudomonas palustris CMA05 g| 29234548 ref NP 948624.1
- Rhodospseudomonas palustris DX 1 g| 316035480 ref YP 004110462.1
- Alpina bromaeae ATCC 49717 g| 41417162 ref ZP 11426538.1
- Bradyrhizobium bacterium SG g| 3389720 ref ZP 08628577.1
- Alpina clevelandensis ATCC 46720 g| 41470289 ref ZP 11425903.1
- Alpina fella ATCC 23850 g| 414165270 ref ZP 1142191.1
- Oligotropha carboxydovorans OMS g| 209642631 ref YP 002288148.1
- Alpina sp INLS2 g| 299131987 ref ZP 07025162.1
- Methylohalobium trichoporum CB30 g| 296446633 ref YP 06888674.1
- Methylohalobium sp ATCC 45242 g| 32137325 ref ZP 08072443.1
- Methylohalobium sp SC2 g| 402771111 ref YP 006590648.1
- Methylohalobium silvestris BL2 g| 217879244 ref YP 00236244.1
- Baerleinia indica subsp indica ATCC 8029 g| 16286060 ref YP 001834551.1
- Paenibacillus baerleinii ATCC 25033 g| 304320901 ref YP 002085424.1
- Polymorphum glum S1035 204.1 g| 32842585 ref YP 004302694.1
- Pseudovibrio sp ED92 g| 254479036 ref ZP 05087300.1
- Microvibrio sp WSA0557 g| 385895963 ref ZP 1018196.1
- Methylobacterium nodulans ORS 2096 g| 220925307 ref YP 002050069.1
- Methylobacterium sp 46 g| 110738958 ref YP 001767260.1
- Bartonella tamiae Th239 g| 395768157 ref ZP 10485884.1
- Bartonella sp ref YP 007461471.1
- Bartonella clarridgeae T3 g| 318599126 ref YP 004159219.1
- Bartonella sp AR 15 3 g| 319405978 emb CB179610.1
- Bartonella sp 11c g| 319407472 emb CB81122.1
- Bartonella rochalimae ATCC BAA 1488 g| 319404479 emb CB178086.1
- Bartonella schoenbuchensis R1 g| 319406796 emb CB182653.1
- Bartonella meliopathi K 20 g| 395794036 ref ZP 10463884.1
- Bartonella bacilliformis KC53 g| 121602295 ref YP 96502.1
- Bartonella bacilliformis INS g| 421761008 ref ZP 16197813.1
- Bartonella washonensis 985 0475 g| 395770723 ref ZP 10450192.1
- Bartonella quintana RM 11 g| 403530718 ref YP 00666247.1
- Bartonella quintana str Toulouse g| 49474430 ref YP 032472.1
- Bartonella bysonii BTB3 g| 395768568 ref ZP 10468078.1
- Bartonella bysonii subsp suprensis Pm136cc g| 395793221 ref ZP 10472626.1
- Bartonella bartlesii LL WM9 g| 395788449 ref ZP 10468013.1
- Bartonella rattinasiensis 15906 g| 395787273 ref ZP 10463109.1
- Bartonella tribocorum CIP 105476 g| 163689662 ref YP 001609891.1
- Bartonella donhae NCTC 12862 g| 395767266 ref ZP 10467194.1
- Bartonella graminea steiae g| 248580961 ref YP 00297293.1
- Phyllobacterium sp YR531 g| 398828662 ref ZP 10588862.1
- Ochrobactrum intermedium LMS 3301 g| 20883292 ref ZP 04860591.1
- Ochrobactrum anthropi ATCC 49188 g| 153008136 ref YP 001707053.1
- Nitratireductor pacificus pH 38 g| 407784444 ref ZP 11127117.1
- Nitratireductor indicus C115 g| 407973615 ref ZP 11154525.1
- Mesorhizobium antracis CCMW30123 g| 35703834 ref ZP 09090373.1
- Mesorhizobium opporfunsum WSM2076 g| 337268123 ref YP 004612178.1
- Mesorhizobium austrarium WSM2073 g| 43377478 ref YP 007302954.1
- Mesorhizobium ciceri bovisiense WSM1271 g| 319782948 ref YP 004142422.1
- Mesorhizobium loti MAFF303099 g| 13471448 ref NP 103014.1
- Aurantimonas mangroveana S165 8A1 g| 90419215 ref ZP 01227125.1
- Fulvmarina pelag HTCC2506 g| 114707529 ref ZP 01440422.1
- Hoefia phototrophica DFL 431 g| 163790111 ref ZP 02196097.1
- Rhizobium lupini HPC L g| 408785755 ref ZP 11197497.1
- Rhizobium leguminosarum bv viciae USDA 2370 g| 424810769 ref ZP 18334146.1
- Agrobacterium tumefaciens CCMW350286 g| 418296309 ref ZP 12911143.1
- Agrobacterium sp H13 3 g| 325293392 ref YP 004276256.1
- Agrobacterium fabrum str. C58 g| 159180181 ref NP 324995.2
- Agrobacterium tumefaciens F2 g| 417859549 ref ZP 12504605.1
- Rhizobium sp g| 400040023 emb CCF 18614.1
- Agrobacterium alberti magni AGL 41 g| 408377830 ref ZP 11176529.1
- Agrobacterium vitis 54 g| 22149066 ref YP 00295025.1
- Rhizobium sp PDD1 076 g| 418396433 ref ZP 13480151.1
- Rhizobium sp CF80 g| 452042328 ref ZP 14746388.1
- Rhizobium etli CIAT 899 g| 440227029 ref YP 00734120.1
- Rhizobium etli CNP4512 g| 417102989 ref ZP 11960014.1
- Rhizobium etli CIAT 652 g| 156826496 ref YP 001976041.1
- Rhizobium etli CFT 42 g| 86368375 ref YP 470267.1
- Rhizobium sp CCGE 510 g| 400484875 ref ZP 1963288.1
- Rhizobium leguminosarum bv viciae 3841 g| 11625296 ref YP 768804.1
- Rhizobium leguminosarum bv trifolii WSM4125 g| 24125474 ref YP 002976570.1
- Rhizobium etli CIAT 894 g| 218682860 ref ZP 03520461.1
- Rhizobium leguminosarum bv trifolii WSM204 g| 209550387 ref YP 002282014.1
- Rhizobium leguminosarum bv trifolii WSM597 g| 424918989 ref ZP 18342353.1
- Rhizobium leguminosarum bv trifolii WSM2012 g| 424885057 ref ZP 18308566.1
- Rhizobium leguminosarum bv trifolii WSM297 g| 424896683 ref ZP 18319487.1

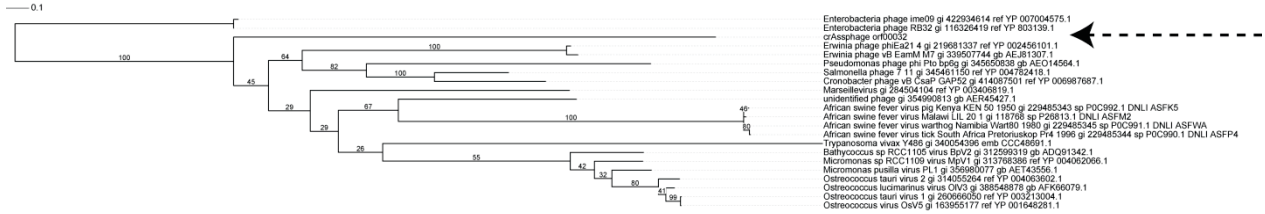
(E) orf00021



(F) orf00028



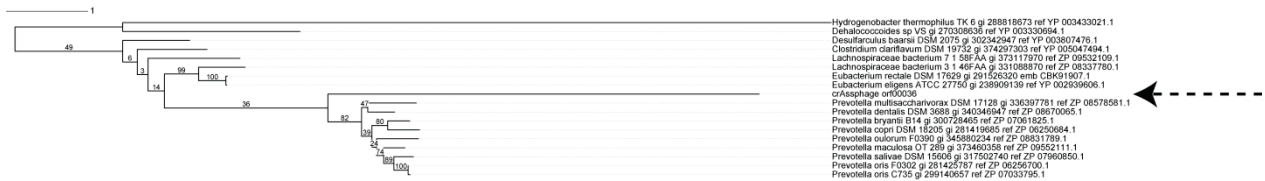
(G) orf00032



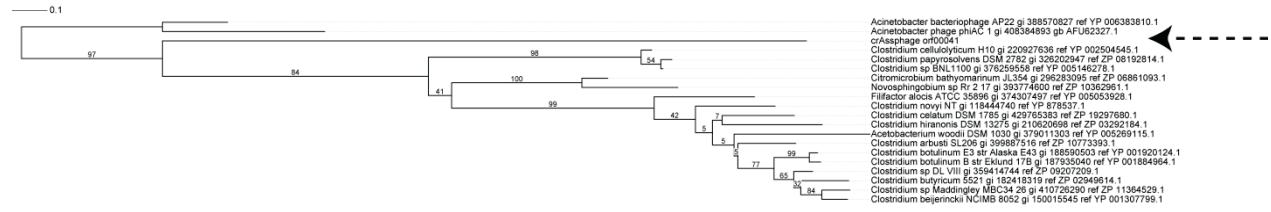
(H) orf00034



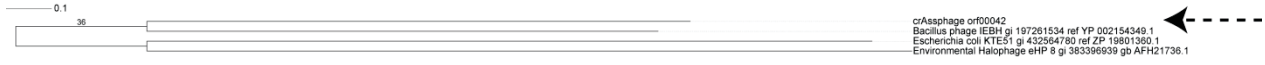
(I) orf00036



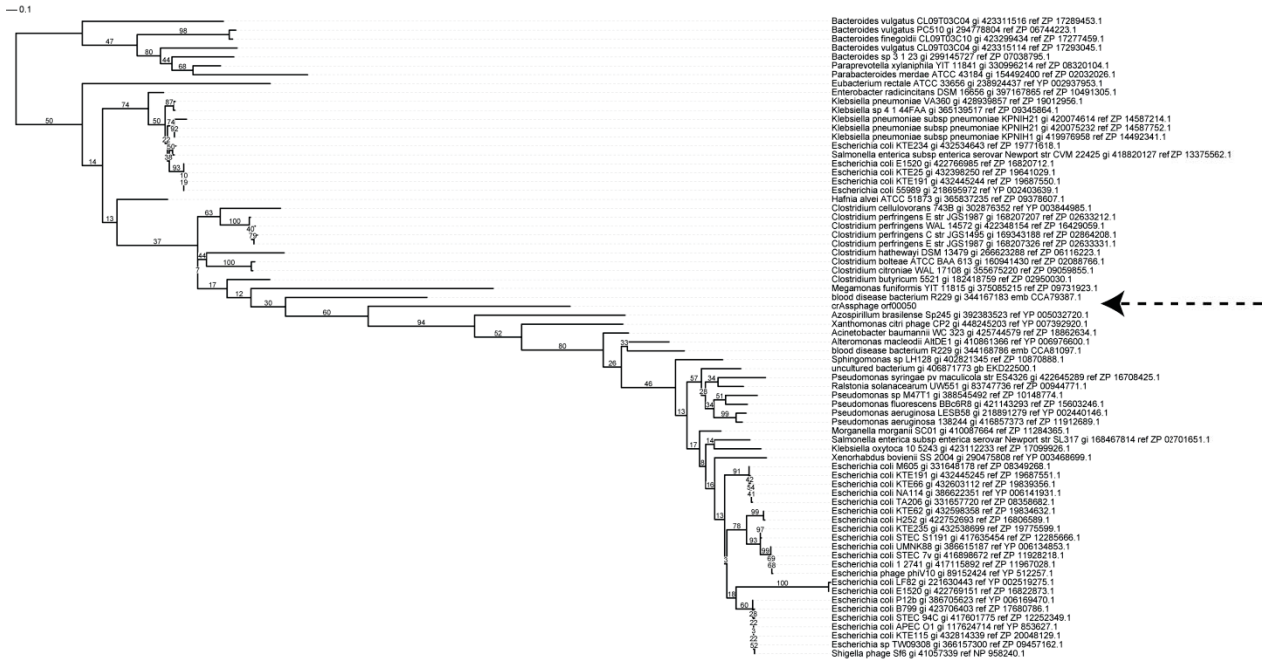
(L) orf00041



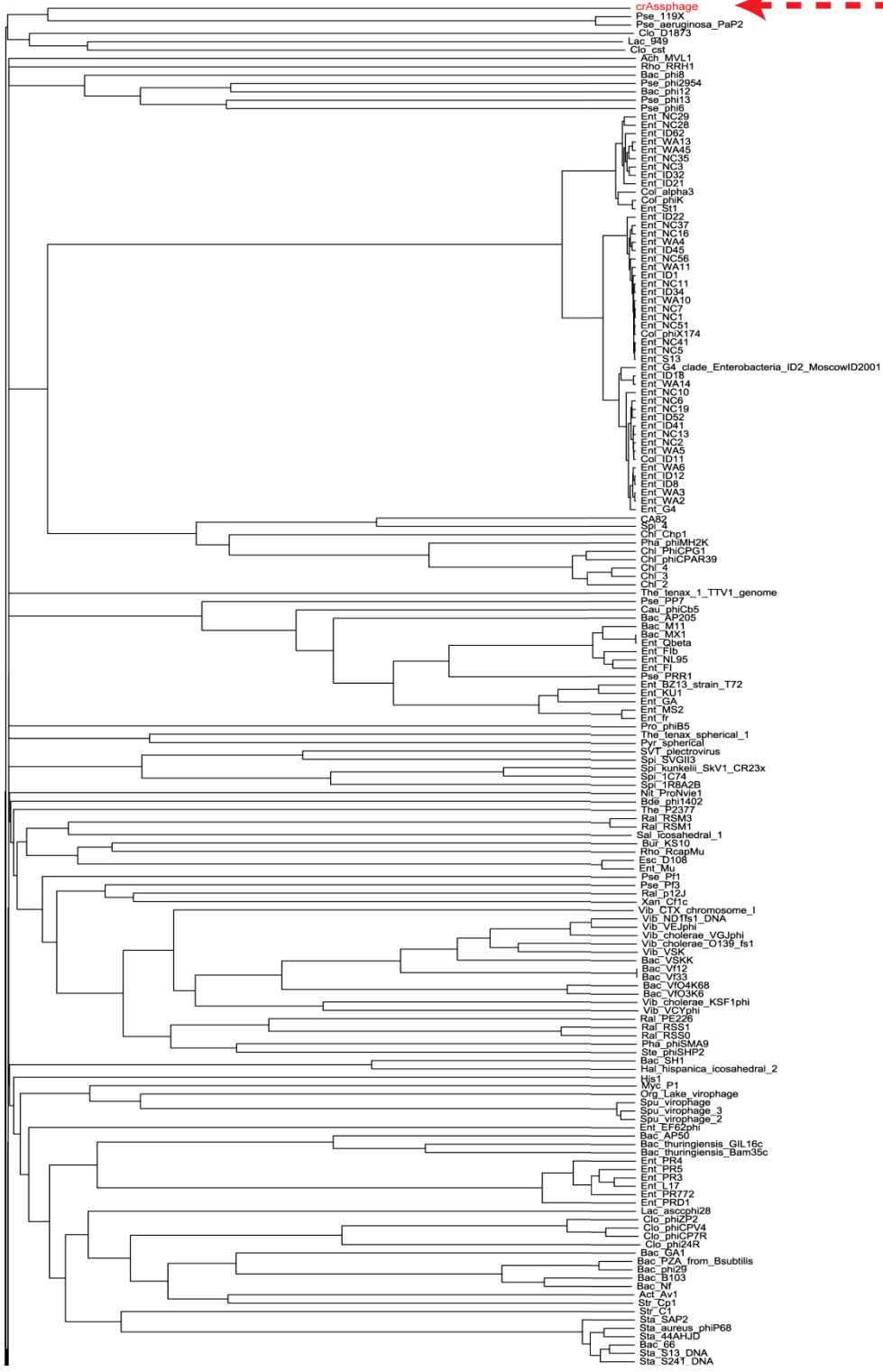
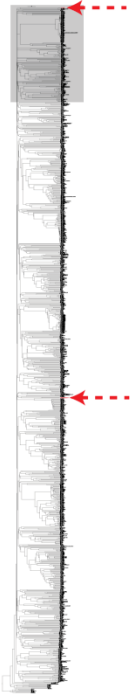
(M) orf00042

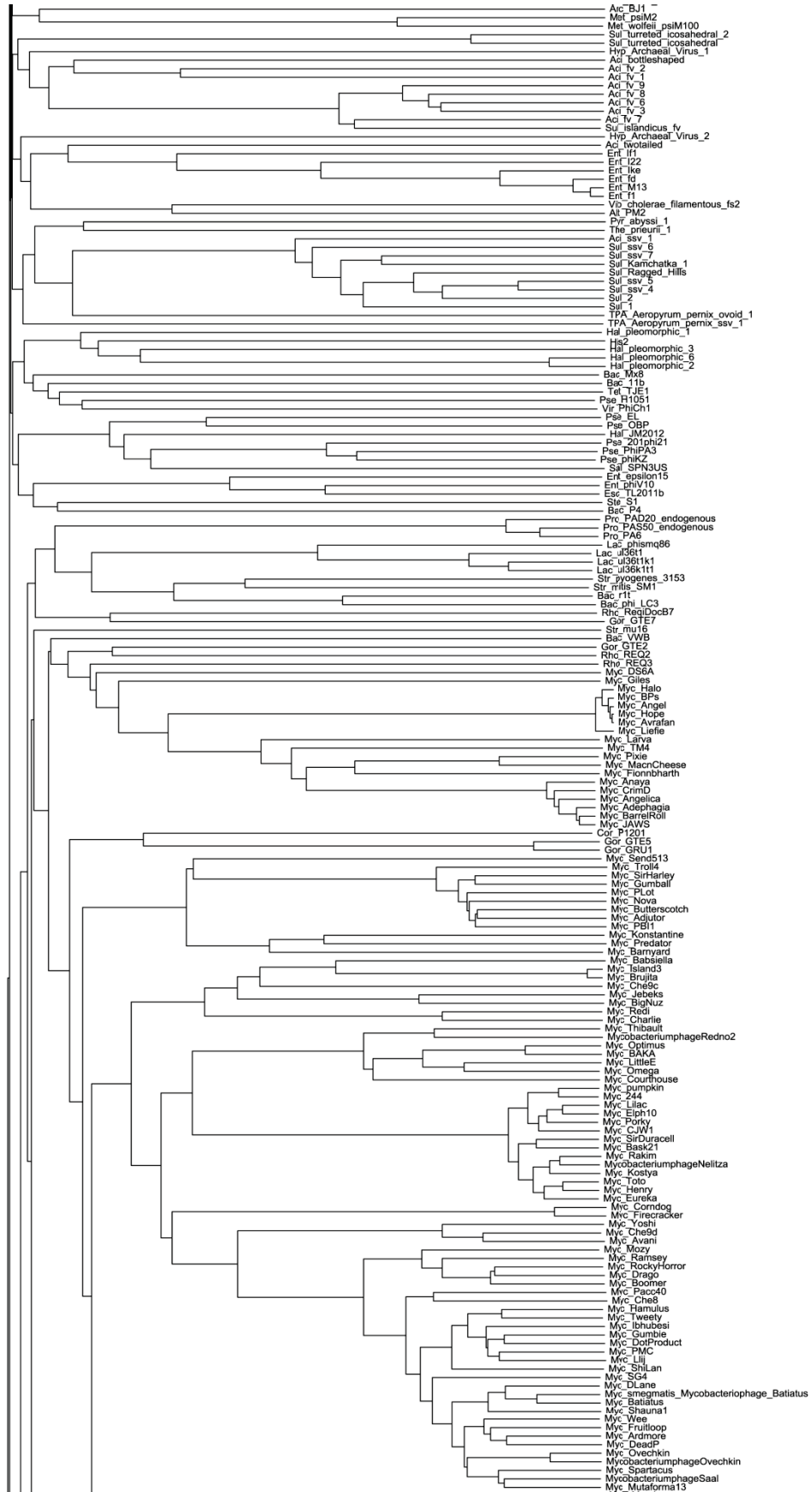
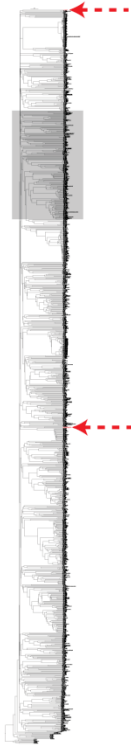


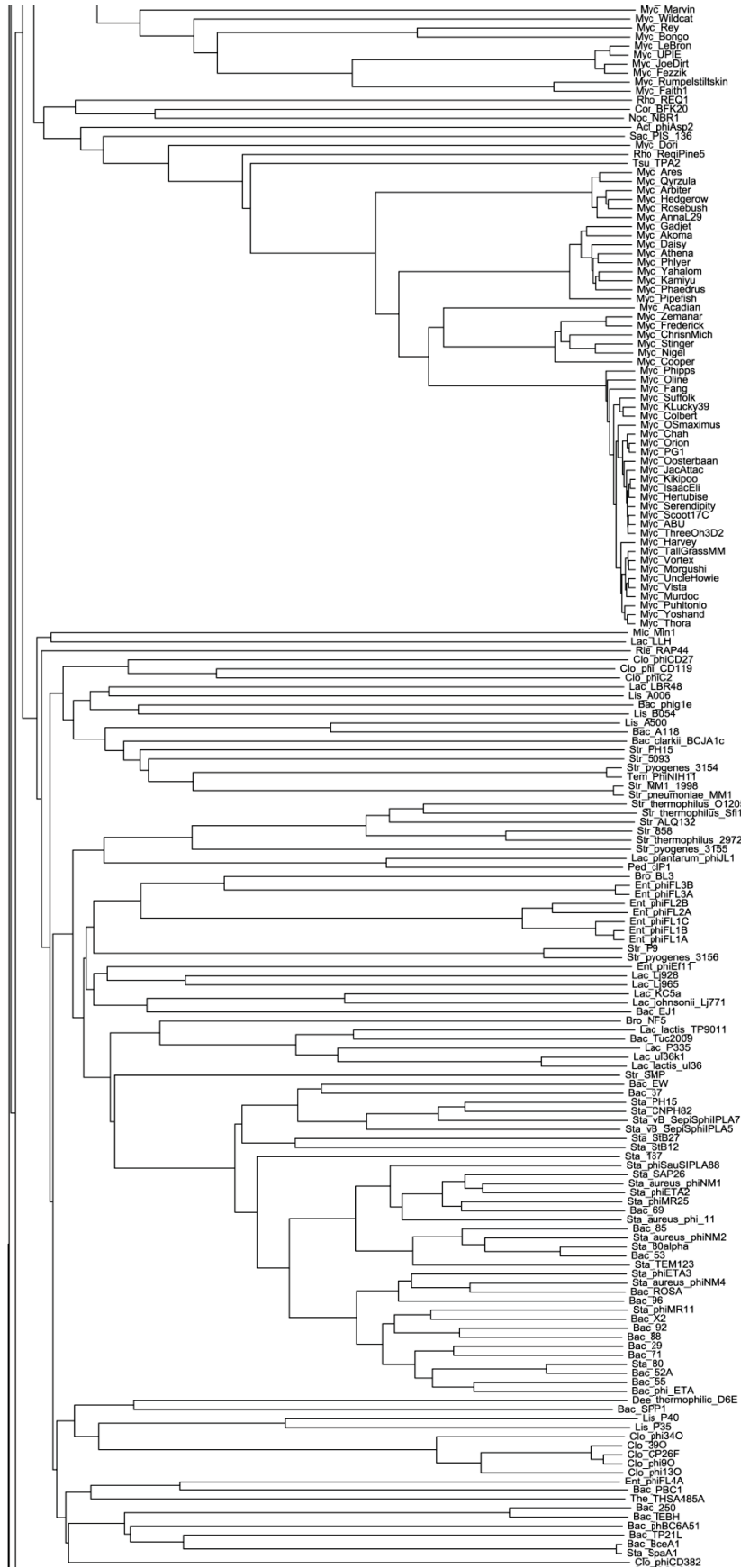
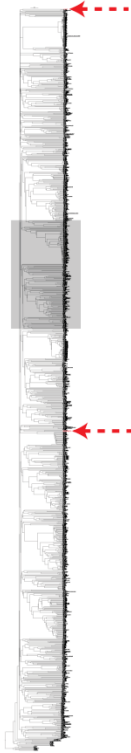
(N) orf00050

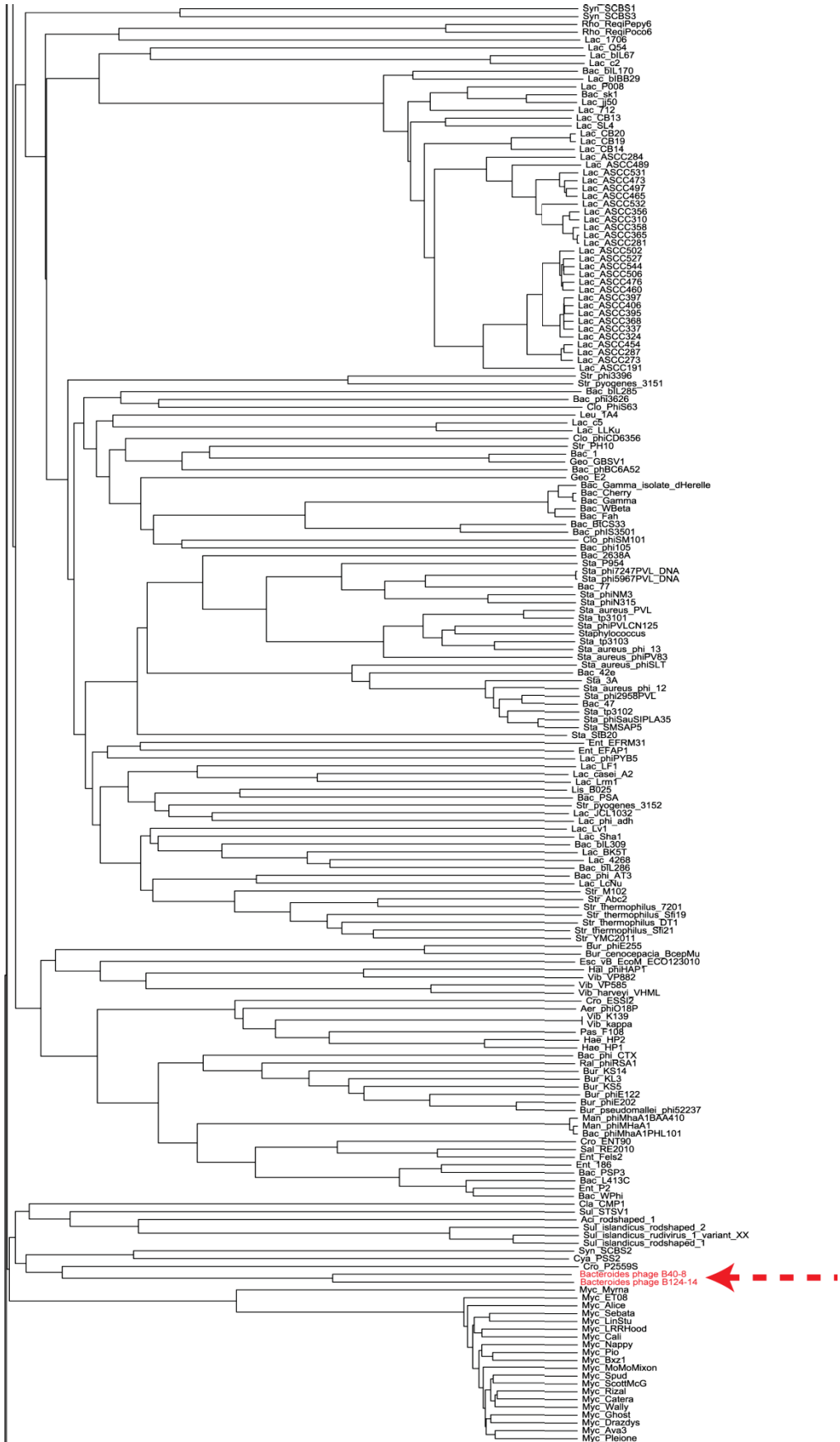
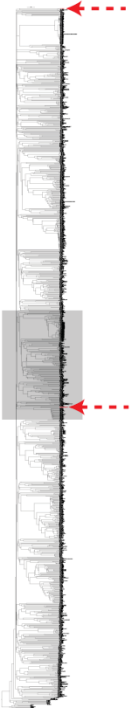


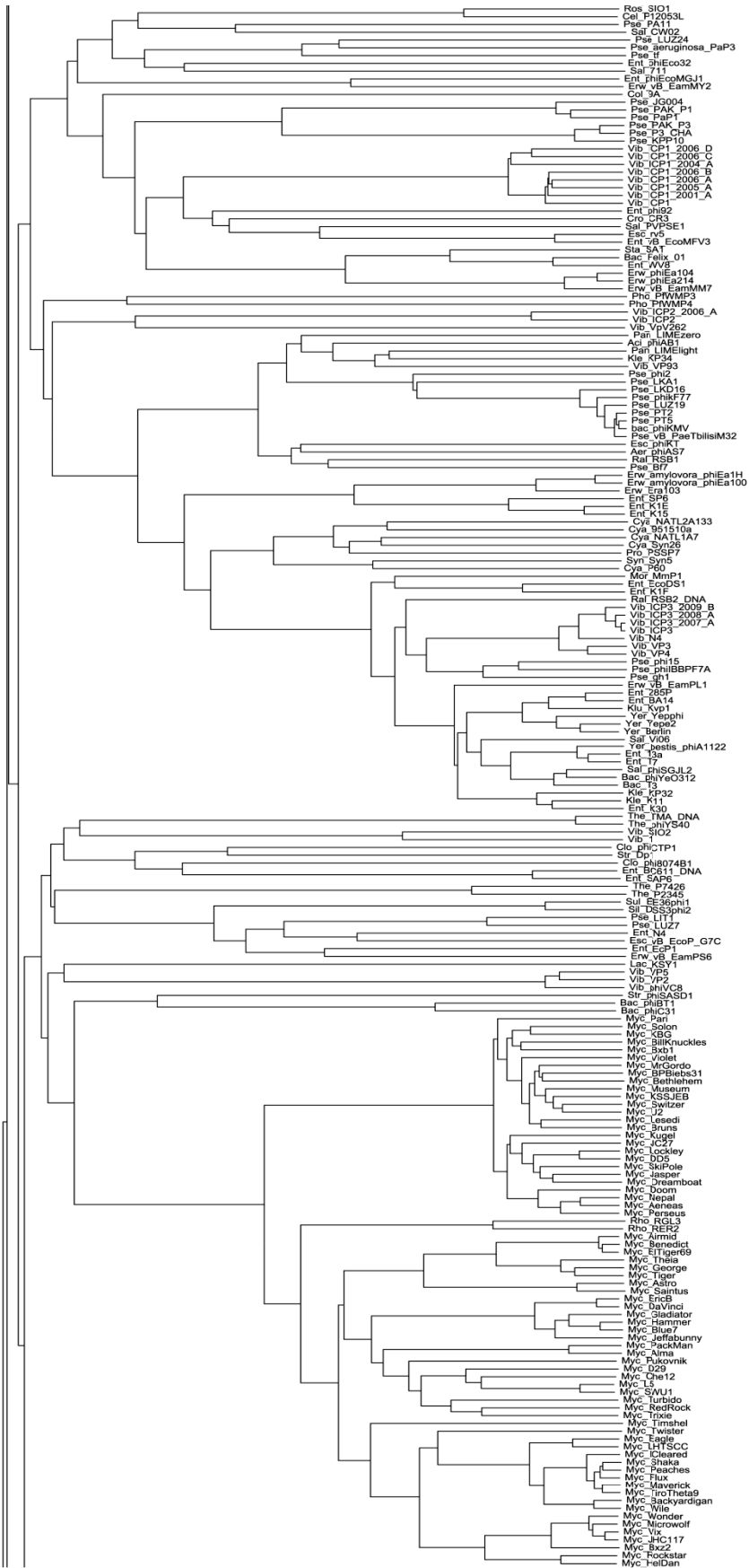
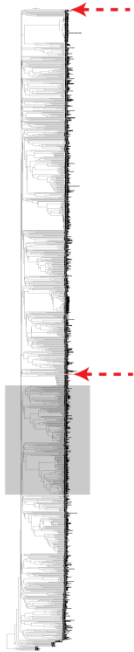
Supplementary Figure 3. Unrooted maximum likelihood phylogenetic trees of crAssphage ORFs orf00068 (A); orf00077 (B); orf00079 (C); orf00019 (D), this large tree is split over four pages, the inset on each page shows which part of the tree is displayed); orf00021 (E); orf00028 (F); orf00032 (G); orf00034 (H); orf00036 (I); orf00040 (J); orf00097 (K) orf00041 (L); orf00042 (M); and orf00050 (N) with Genbank homologs. In each tree, the crAssphage ORF is indicated with a dashed arrow. The Phylip files for these trees are available from the authors on request.

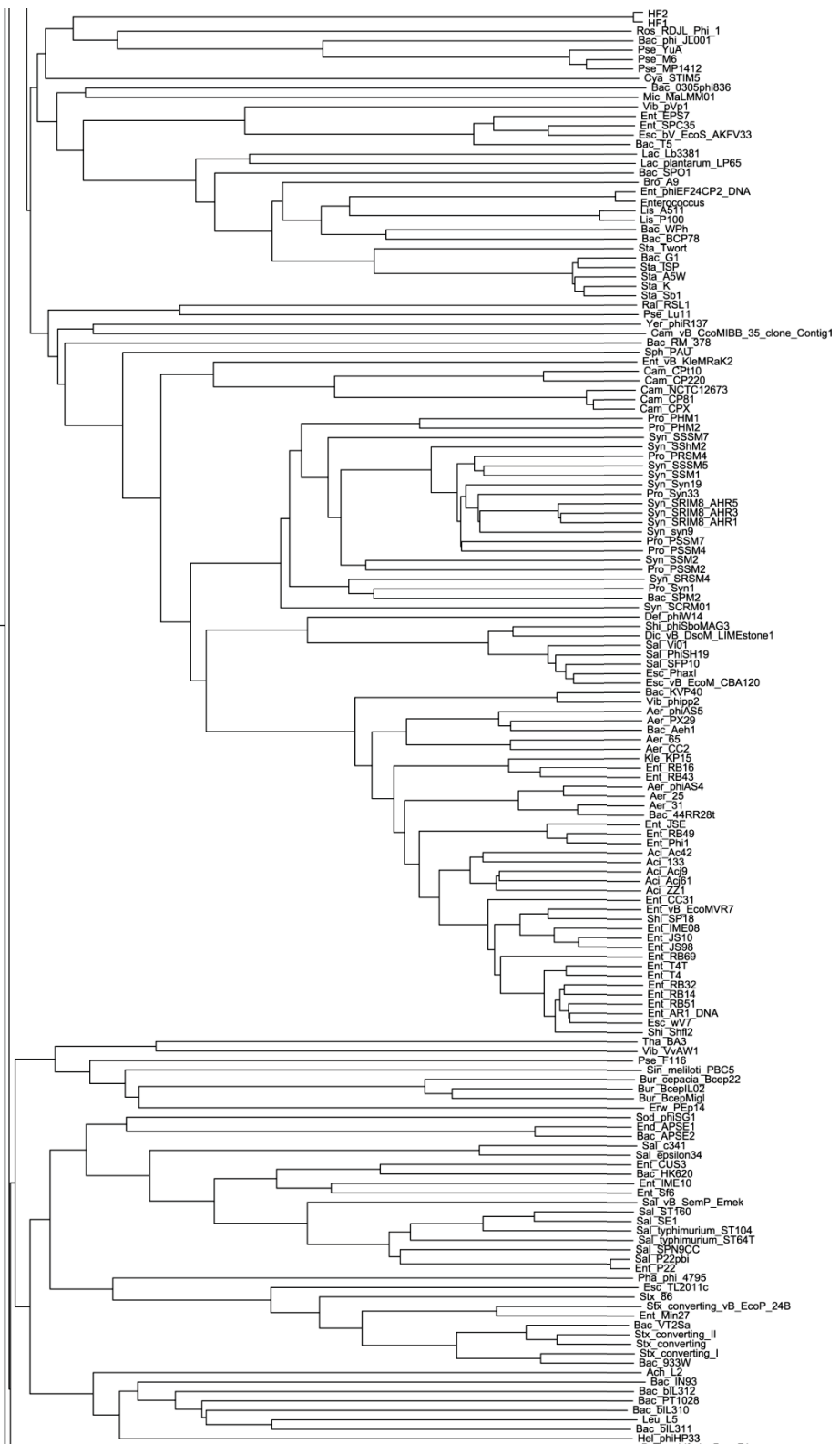
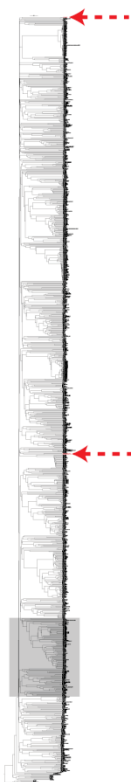


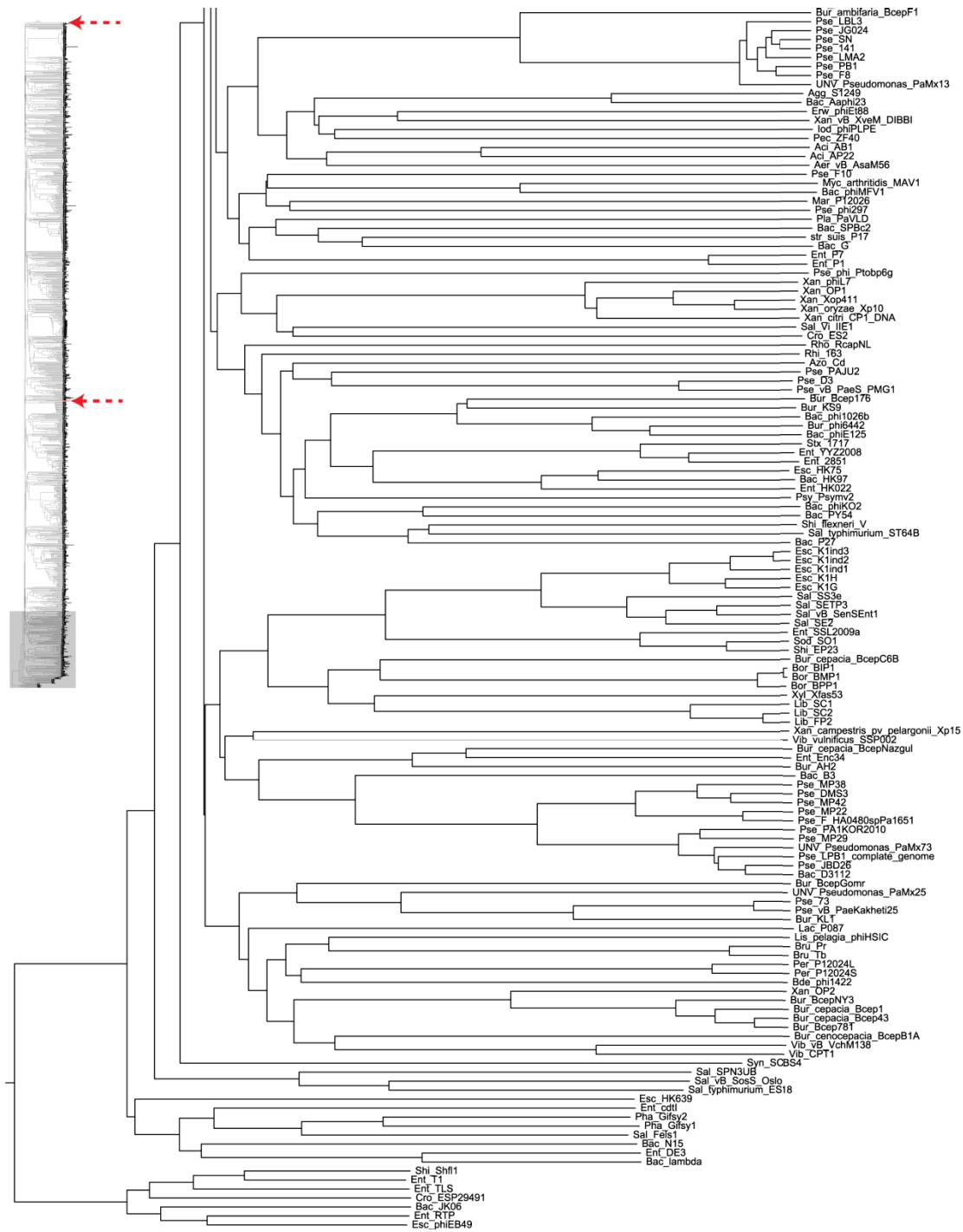




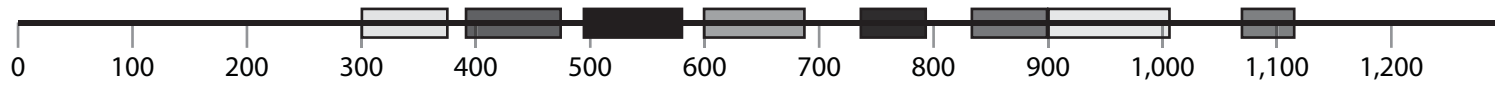




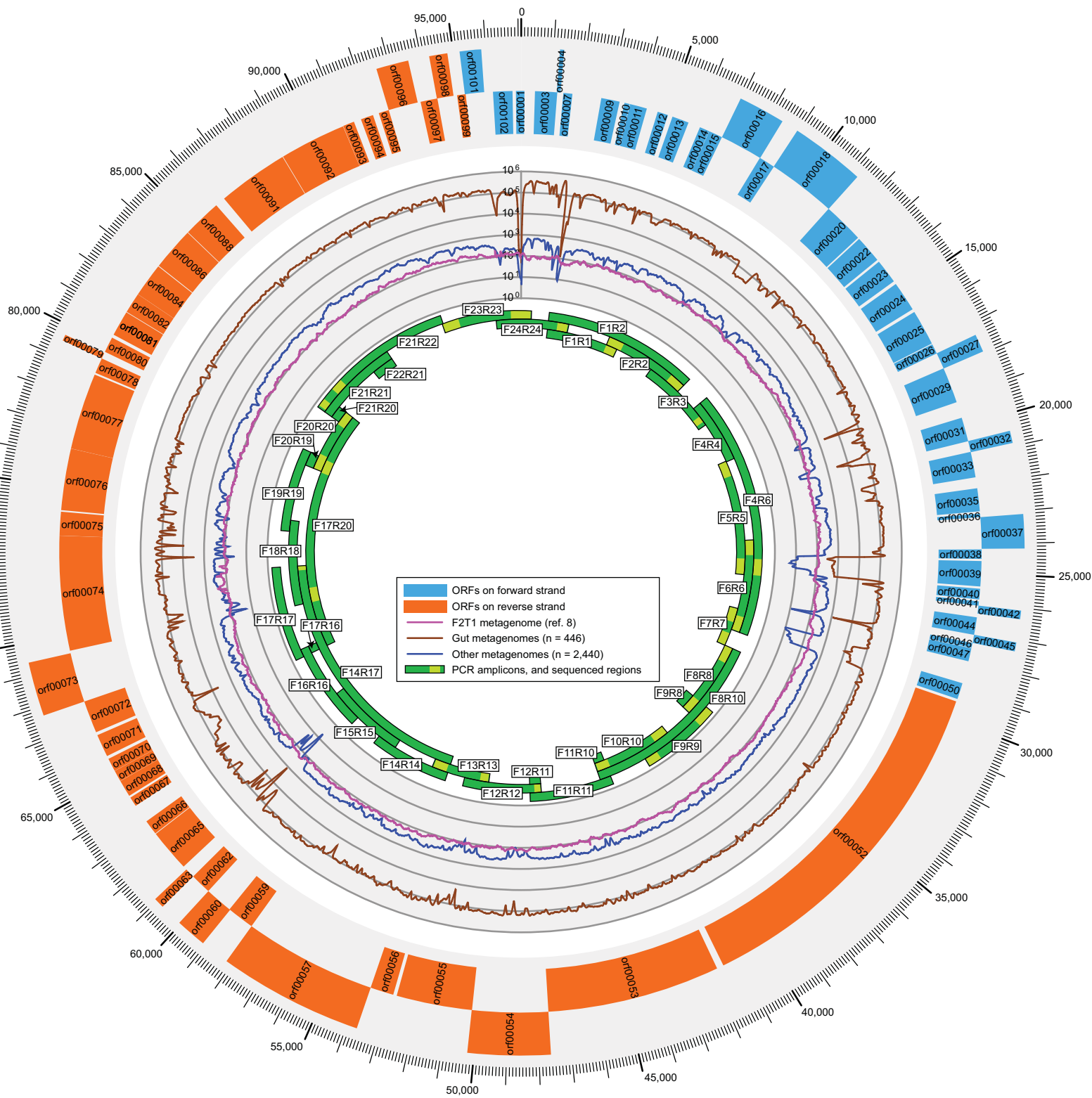




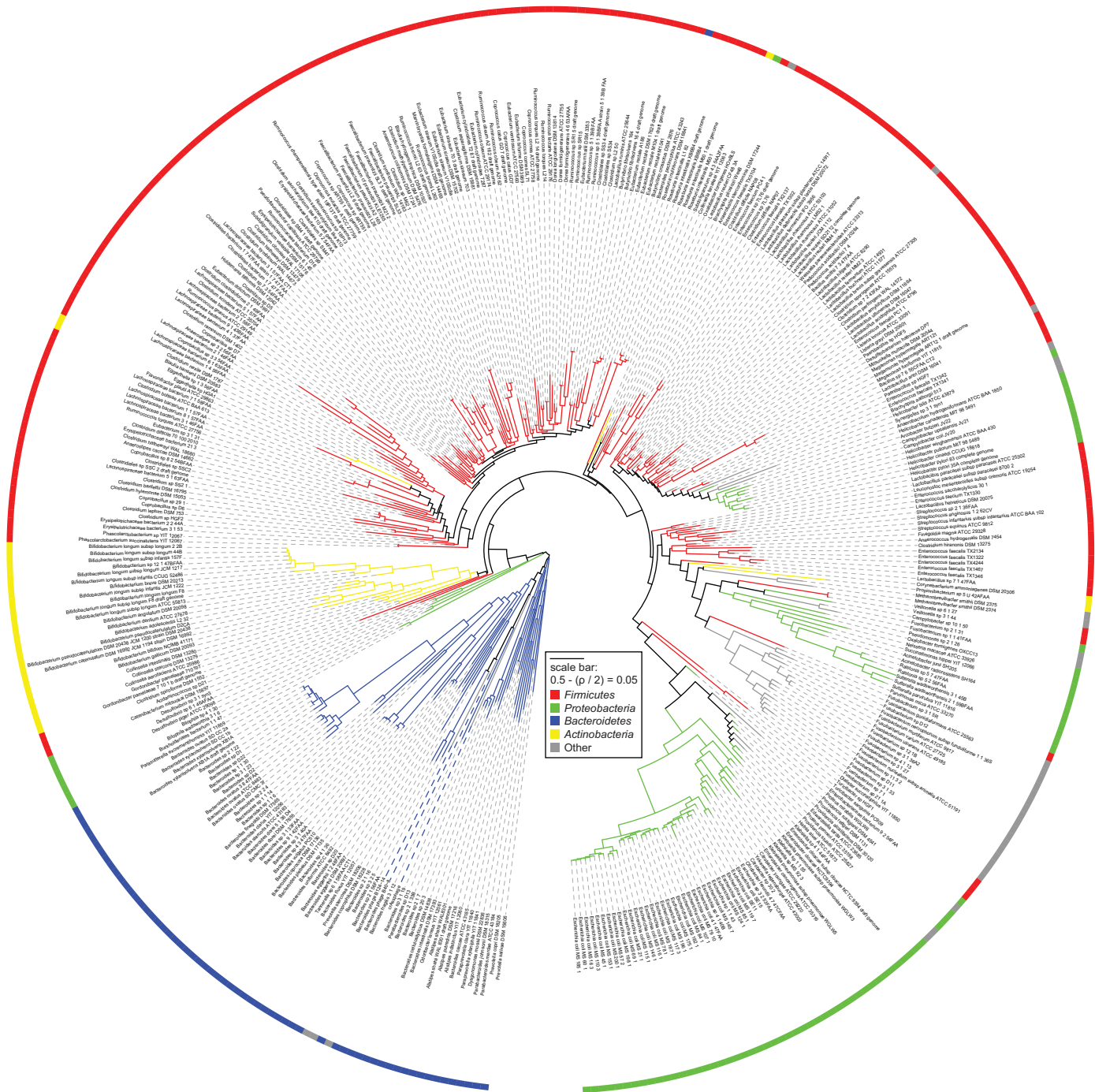
Supplementary Figure 4. Phage proteomic tree of 1,193 complete viral genomes including crAssphage. CrAssphage and the two *Bacteroides* phages B124-14 and B40-8 are indicated with a dashed arrow in red. This large tree is split over four pages, the inset on each page shows which part of the tree is displayed.



Supplementary Figure 5. Domain structure of orf00074, including eight BACON domains. The shading intensity indicates the bitscore, from left to right: 3.00, 19.06, 36.88, 10.76, 23.95, 16.13, 2.52, and 15.20.



Supplementary Figure 6. Schematic representation of the circular crAssphage genome. The genome contains 80 ORFs that were predicted with Glimmer trained on Caudovirales. The total coverage of each nucleotide in the F2T1 metagenome, and in all public metagenomes in MG-RAST is indicated (466 human fecal and 2,440 other metagenomes, as determined by blastn mapping: ≥ 75 bp aligned with $\geq 95\%$ identity, see Methods). Green bars indicate the 36 regions that were validated by long-range PCR (see Table 2 and Supplementary Table 1). Selected regions of several PCR amplicons (indicated as light green regions in the green bars) were sequenced by Sanger dideoxynucleotide sequencing to validate that the amplicons were indeed derived from the crAssphage genome (Supplementary Table 1).



Supplementary Figure 7. Phage-host prediction based on co-occurrence across metagenomes. Unrooted co-occurrence cladogram of correlated depth profiles across 151 HMP fecal metagenomes of the crAssphage, two known *Bacteroides fragilis*-infecting phages, and 404 potential hosts. Colors indicate bacterial phyla. The phages are indicated with blue dashed lines.

Supplementary Table 1. Regions sequenced by Sanger dideoxynucleotide sequencing and their origin. Sample: sample of fecal virus-like particle (VLP) derived DNA used for amplification, see Methods for details. Primers: primer pair used for amplification by long-range PCR; see Table 2 for sequences; see Supplementary Figure 6 for region; see Supplementary Figure 2 for band. Amplified: start and end positions of the amplicon. Sequenced: start and end positions of the region sequenced by Sanger sequencing. Identity: percent identity of the region sequenced by Sanger sequencing.

Identifier	Sample	Genbank ID	Primers	Amplified	Sequenced	Identity
TS4.2-F1R1-1	TS4.2	KM000086	F1-R1	1,784-6,793	6,010-6,732	97.79
TS4.2-F2R2-2	TS4.2	KM000087	F2-R2	6,003-11,923	6,040-7,068	99.90
TS4.2-F2R2-3	TS4.2	KM000088	F2-R2	6,003-11,923	10,855-11,882	99.61
TSDC8.2-F3R3-4	TSDC8.2	KM000089	F3-R3	9,511-14,930	14,214-14,651	91.55
TS4.2-F4R6-5	TS4.2	KM000090	F4-R6	13,525-29,713	24,740-25,760	99.71
TS4.2-F4R6-6	TS4.2	KM000091	F4-R6	13,525-29,713	24,748-25,785	100.00
TS4.2-F5R5-7	TS4.2	KM000092	F5-R5	17,800-25,822	17,855-18,859	99.80
TS4.2-F5R5-8	TS4.2	KM000093	F5-R5	17,800-25,822	24,779-25,785	99.80
TS4.2-F6R6-9	TS4.2	KM000094	F6-R6	23,444-29,713	23,487-24,478	99.80
TS4.2-F6R6-10	TS4.2	KM000095	F6-R6	23,444-29,713	28,632-29,679	99.24
TS4.2-F7R7-11	TS4.2	KM000096	F7-R7	28,191-30,897	28,223-29,125	99.78
TSDC8.2-F7R7-12	TS4.2	KM000097	F7-R7	28,191-30,897	28,223-29,040	94.01
TS4.2-F7R7-13	TSDC8.2	KM000098	F7-R7	28,191-30,897	30,024-30,831	99.75
TSDC8.2-F7R7-14	TSDC8.2	KM000099	F7-R7	28,191-30,897	30,009-30,838	97.47
TS4.2-F8R8-15	TS4.2	KM000100	F8-R8	30,874-36,058	30,913-31,880	99.69
TS4.2-F8R8-16	TS4.2	KM000101	F8-R8	30,874-36,058	35,064-36,006	100.00
TS4.2-F9R9-17	TS4.2	KM000102	F9-R9	35,018-40,185	35,064-36,060	99.80
TS4.2-F9R9-18	TS4.2	KM000103	F9-R9	35,018-40,185	39,256-40,134	98.75
TS4.2-F10R10-19	TS4.2	KM000104	F10-R10	38,016-43,360	38,067-39,023	100.00
TS4.2-F10R10-20	TS4.2	KM000105	F10-R10	38,016-43,360	42,451-43,309	99.77
TSDC8.2-F12R12-22	TSDC8.2	KM000106	F12-R12	47,319-52,314	47,289-47,646	93.58
TSDC8.2-F13R13-23	TSDC8.2	KM000107	F13-R13	50,677-54,585	50,746-51,290	92.46
TSDC8.2-F13R13-24	TSDC8.2	KM000108	F13-R13	50,677-54,585	53,703-54,518	97.92
TSDC8.2-F14R17-25	TSDC8.2	KM000109	F14-R17	53,476-71,818	71,512-71,805	97.28
TSDC8.2-F17R20-26	TSDC8.2	KM000110	F17-R20	65,847-83,203	69,229-70,168	95.43
TSDC8.2-F17R20-27	TSDC8.2	KM000111	F17-R20	65,847-83,203	78,679-79,572	98.77
TS4.2-F20R20-28	TS4.2	KM000112	F20-R20	78,704-83,203	78,761-79,782	99.80
TS4.2-F20R20-29	TS4.2	KM000113	F20-R20	78,704-83,203	82,218-83,152	99.89
TSDC8.2-F21R22-30	TSDC8.2	KM000114	F21-R22	82,363-91,978	82,470-83,000	91.34
TSDC8.2-F21R22-31	TSDC8.2	KM000115	F21-R22	82,363-91,978	83,756-84,563	97.28
TS4.2-F23R23-32	TS4.2	KM000116	F23-R23	91,909-628	91,940-92,951	99.51
TS4.2-F23R23-33	TS4.2	KM000117	F23-R23	91,909-628	96,503-97,064	96.80
TS4.2-F23R23-34	TS4.2	KM000118	F23-R23	91,909-628	4-601	99.50
TS4.2-F23R23-35	TS4.2	KM000119	F23-R23	91,909-628	96,513-96,998	97.12
TSDC8.2-F23R23-36	TSDC8.2	KM000120	F23-R23	91,909-628	91,957-92,893	97.24
TS4.2-F24R24-37	TS4.2	KM000121	F24-R24	95,394-3,167	2,420-3,109	99.28