

This file contains Supplementary data for the article by R. Sieskind et al. entitled: “From Demodification to Sequencing: the Genome of Bacteriophage Carin-5 “

It contains

- 3 Supplementary Tables
- 4 Supplementary Figures

Table S1. Buffers for DNA polymerase I purification. Buffer A is used to resuspend bacterial pellets, for cells disruption and for supernatant binding on HisTrap column after high-speed centrifuge. Buffer B is the HisTrap washing buffer. The elution imidazole gradient is performed using the buffer B and C. Buffer D is a dilution buffer to decrease NaCl concentration down to 50 mM before binding on Heparin column. Buffer E is the Heparin washing buffer. The elution salt gradient is performed using the buffer E and F. The selected fractions are dialyzed against the storage Buffer G before concentration adjustment and flash freeze in liquid nitrogen.

	A	B	C	D	E	F	G
Tris-HCl (pH 8.0)	50 mM	50 mM	50 mM	50 mM	50 mM	50 mM	
NaCl	300 mM	300 mM	300 mM	-	50 mM	1 M	
MgCl ₂	10 mM	10 mM	10 mM	10 mM	10 mM	10 mM	
Imidazole	-	10 mM	500 mM	-	-	-	-
Benzonase nuclease	500 U / 100 mL	-	-	-	-	-	-
Protease inhibitor	1 tab / 100 mL	-	-	-	-	-	-

Table S2. Primers used for LongAmp PCR reaction. The ‘PCR_universal_Fwd’ primer was used as forward primer for all the PCR, the length of the amplicons depending on the reverse primers.

Name	Sequence (5' → 3')
PCR_universal_Fwd	CGAATCTGTACTATGTAGTCATCATAATCAGCATC
PCR_500bp_Rev	CAAGATGCACTAGCGTTGAATTCTACAGTAG
PCR_1000bp_Rev	GAGTCAGTACTGTATCAACTGGTATAGGTG
PCR_1500bp_Rev	TGGATGGTTCTCAGATATTACTGATATCATTCC
PCR_2000bp_Rev	CTAGGTAATATTAATACTTCAAATAATTCTTCAGTACAAGCTG
PCR_2500bp_Rev	CTAATCAAAGAGTCCTTAATGCCTCAAGG
PCR_3000bp_Rev	CAGTCCTAATAACAGTACTTCATTTGATGATCTTC

Table S3. List of genes that could be annotated directly by Blastp searches. Ninety-five of the 226 predicted gene products were homologous to proteins in the RefSeq protein database when an E-value cutoff of 1e-04 was used. Among these 95, only 50 had an annotated function.

Query Seq-id	E-value	All subject titles
CLPJFFLI_00005	3.07e-07	ref WP_019687629 WP_019687629 hypothetical protein [Paenibacillus polymyxa] [Paenibacillus polymyxa]
CLPJFFLI_00013	1.48e-25	ref WP_176531376 WP_176531376 hypothetical protein [Bradyrhizobium quebecense] [Bradyrhizobium quebecense]
CLPJFFLI_00018	8.86e-15	ref WP_277595891 WP_277595891 nucleoside 2-deoxyribosyltransferase [Staphylococcus equorum] [Staphylococcus equorum]
CLPJFFLI_00043	4.18e-10	ref WP_169902409 WP_169902409 anti-CBASS protein Acb1 family protein [Pseudomonas rhodesiae] [Pseudomonas rhodesiae]
CLPJFFLI_00045	7.14e-43	ref WP_210601376 WP_210601376 hypothetical protein [Aeromonas sp. MrichA-1] [Aeromonas sp. MrichA-1]
CLPJFFLI_00048	1.11e-05	ref WP_152467710 WP_152467710 hypothetical protein [Gordonia terrae] [Gordonia terrae]
CLPJFFLI_00049	4.76e-108	ref YP_009302517 YP_009302517 thymidylate synthetase [Bacillus phage SP-15] [Bacillus phage SP-15]
CLPJFFLI_00050	1.00e-16	ref WP_339617236 WP_339617236 hypothetical protein [uncultured Gilvimarinus sp.] [uncultured Gilvimarinus sp.]
CLPJFFLI_00058	2.20e-12	ref WP_285168833 WP_285168833 SDR family oxidoreductase [Rhizobium sp. CNPSO 3464] [Rhizobium sp. CNPSO 3464]
CLPJFFLI_00063	1.38e-13	ref WP_395048160 WP_395048160 dCTP deaminase domain-containing protein, partial [Flavobacterium sp.] [Flavobacterium sp.]
CLPJFFLI_00066	8.38e-28	ref WP_284022679 WP_284022679 hypothetical protein [Pseudoalteromonas sp. B95] [Pseudoalteromonas sp. B95]
CLPJFFLI_00068	2.08e-25	ref YP_009814663 YP_009814663 ribonucleotide reductase NrdA-like [Proteus phage Stubb] [Proteus phage Stubb]
CLPJFFLI_00069	4.00e-07	ref YP_009816416 YP_009816416 hypothetical protein HOU61_gp139 [Synechococcus phage S-CBWM1] [Synechococcus phage S-CBWM1]
CLPJFFLI_00070	4.36e-18	ref WP_380253557 WP_380253557 hypothetical protein [Kiloniella antarctica] [Kiloniella antarctica]
CLPJFFLI_00072	3.47e-05	ref WP_057429220 WP_057429220 hypothetical protein [Pseudomonas syringae group genomsp. 3] [Pseudomonas syringae group genomsp. 3]
CLPJFFLI_00074	1.79e-10	ref WP_185210734 WP_185210734 hypothetical protein [Elizabethkingia meningoseptica] [Elizabethkingia meningoseptica]
CLPJFFLI_00075	6.40e-50	ref WP_192876686 WP_192876686 hypothetical protein [Paracoccus jeotgali] [Paracoccus jeotgali]
CLPJFFLI_00078	3.87e-09	ref WP_346856558 WP_346856558 hypothetical protein [uncultured Draconibacterium sp.] [uncultured Draconibacterium sp.]
CLPJFFLI_00082	3.24e-09	ref WP_074755540 WP_074755540 hypothetical protein [Pseudomonas abietaniphila] [Pseudomonas abietaniphila]
CLPJFFLI_00094	1.27e-13	ref WP_022611685 WP_022611685 hypothetical protein [Vibrio nigripulchritudo] [Vibrio nigripulchritudo]
CLPJFFLI_00095	1.39e-06	ref YP_009904615 YP_009904615 hypothetical protein H1O17_gp237 [Burkholderia phage BcepSauron] [Burkholderia phage BcepSauron]
CLPJFFLI_00096	1.53e-123	ref WP_411270034 WP_411270034 polysaccharide pyruvyl transferase family protein [Alteromonas sp. CYL-A6] [Alteromonas sp. CYL-A6]
CLPJFFLI_00099	7.13e-39	ref YP_009883632 YP_009883632 DprA-like DNA recombination-mediator protein [Stenotrophomonas phage Pokken] [Stenotrophomonas phage Pokken]

CLPJFFLI_00100	1.10e-20	ref WP_345192874 WP_345192874 hypothetical protein [Kistimonas scapharcae] [Kistimonas scapharcae]
CLPJFFLI_00101	3.19e-07	ref YP_009846604 YP_009846604 hypothetical protein HWC27_gp056 [Aeromonas phage 4L372D] [Aeromonas phage 4L372D]
CLPJFFLI_00102	2.06e-29	ref WP_013327390 WP_013327390 exonuclease domain-containing protein [Sulfurimonas autotrophica] [Sulfurimonas autotrophica]
CLPJFFLI_00104	8.60e-19	ref WP_023791953 WP_023791953 thioredoxin family protein [Candidatus Babela massiliensis] [Candidatus Babela massiliensis]
CLPJFFLI_00105	1.70e-29	ref YP_009854318 YP_009854318 hypothetical protein HWC84_gp080 [Erwinia phage Hena1] [Erwinia phage Hena1]
CLPJFFLI_00106	7.72e-06	ref WP_262430813 WP_262430813 hypothetical protein [Paratissierella segnis] [Paratissierella segnis]
CLPJFFLI_00108	7.02e-47	ref WP_126232210 WP_126232210 helicase C-terminal domain-containing protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00114	6.90e-103	ref WP_126232528 WP_126232528 Hint domain-containing protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00115	9.05e-41	ref WP_126232162 WP_126232162 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00117	4.04e-26	ref WP_161969383 WP_161969383 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00119	7.23e-87	ref WP_126262813 WP_126262813 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00121	8.63e-05	ref WP_311338752 WP_311338752 DUF2339 domain-containing protein [Capnocytophaga leadbetteri] [Capnocytophaga leadbetteri]
CLPJFFLI_00123	3.58e-04	ref WP_126232154 WP_126232154 sigma-70 domain-containing protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00124	2.14e-11	ref WP_126232123 WP_126232123 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00126	7.51e-10	ref WP_126232152 WP_126232152 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00128	1.30e-119	ref WP_126232149 WP_126232149 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00129	6.59e-22	ref WP_126262808 WP_126262808 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00131	5.82e-14	ref WP_126232164 WP_126232164 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00132	1.60e-55	ref WP_198350055 WP_198350055 hypothetical protein [Vreelandella alkaliphila] [Vreelandella alkaliphila]
CLPJFFLI_00133	3.19e-06	ref WP_138594720 WP_138594720 hypothetical protein [Pseudoalteromonas citrea] [Pseudoalteromonas citrea]
CLPJFFLI_00135	6.27e-10	ref YP_009794299 YP_009794299 hypothetical protein HOS33_gp125 [Erwinia phage vB_EamM_Y3] [Erwinia phage vB_EamM_Y3]
CLPJFFLI_00136	6.38e-07	ref WP_126262946 WP_126262946 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00138	6.62e-09	ref WP_126232526 WP_126232526 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00139	6.34e-32	ref WP_320153048 WP_320153048 hypothetical protein [uncultured Tolumonas sp.] [uncultured Tolumonas sp.]
CLPJFFLI_00140	2.55e-41	ref WP_192329588 WP_192329588 PAAR domain-containing protein [Pseudomonas sp. PDM14] [Pseudomonas sp. PDM14]
CLPJFFLI_00141	6.82e-11	ref WP_287009673 WP_287009673 GPW/gp25 family protein [Alteromonas sp.] [Alteromonas sp.]

CLPJFFLI_00142	2.66e-46	ref WP_126232165 WP_126232165 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00144	2.05e-10	ref WP_058615529 WP_058615529 hypothetical protein [Tepidimonas taiwanensis] [Tepidimonas taiwanensis]
CLPJFFLI_00147	9.07e-24	ref WP_344910486 WP_344910486 hypothetical protein [Paenibacillus hodogayensis] [Paenibacillus hodogayensis]
CLPJFFLI_00148	9.32e-12	ref WP_063545414 WP_063545414 hypothetical protein [Pseudomonas putida] [Pseudomonas putida]
CLPJFFLI_00149	4.48e-136	ref WP_247754128 WP_247754128 hypothetical protein [Citrobacter freundii] [Citrobacter freundii]
CLPJFFLI_00150	1.36e-24	ref WP_110600825 WP_110600825 hypothetical protein [Salinicola lusitanus] [Salinicola lusitanus]
CLPJFFLI_00151	1.05e-138	ref WP_069112255 WP_069112255 peptidase G2 autoproteolytic cleavage domain-containing protein [Jiangella alba] [Jiangella alba]
CLPJFFLI_00152	5.57e-117	ref WP_303334623 WP_303334623 glycosyl hydrolase family 28-related protein [Sutterella sp.] [Sutterella sp.]
CLPJFFLI_00153	6.60e-17	ref WP_059364820 WP_059364820 phage tail protein [Enterobacter roggenkampii] [Enterobacter roggenkampii]
CLPJFFLI_00154	3.65e-30	ref WP_028694884 WP_028694884 tail fiber domain-containing protein [Pseudomonas cremoricolorata] [Pseudomonas cremoricolorata]
CLPJFFLI_00155	6.04e-16	ref YP_009037300 YP_009037300 hypothetical protein GJ22_gp042 [Nitrincola phage 1M3-16] [Nitrincola phage 1M3-16]
CLPJFFLI_00156	1.66e-45	ref WP_369105948 WP_369105948 hypothetical protein [Providencia hangzhouensis] [Providencia hangzhouensis]
CLPJFFLI_00159	5.86e-80	ref WP_282476630 WP_282476630 glycosyl hydrolase 108 family protein [Vibrio sp. ABG19] [Vibrio sp. ABG19]
CLPJFFLI_00163	2.27e-84	ref YP_009822671 YP_009822671 PhoH-like phosphate starvation-inducible [Rheinheimera phage Barba5S] [Rheinheimera phage Barba5S]
CLPJFFLI_00164	7.51e-20	ref WP_340436202 WP_340436202 nucleotidyltransferase domain-containing protein [Lacibacter sp. H375] [Lacibacter sp. H375]
CLPJFFLI_00169	1.84e-60	ref XP_018026159 XP_018026159 thymidylate synthase-like isoform X2 [Hyaella azteca] [Hyaella azteca]
CLPJFFLI_00171	0.0	ref WP_335767917 WP_335767917 ribonucleoside-diphosphate reductase subunit alpha [Brachyspira pilosicoli] [Brachyspira pilosicoli]
CLPJFFLI_00172	6.56e-126	ref WP_288680296 WP_288680296 ribonucleotide-diphosphate reductase subunit beta [uncultured Brachyspira sp.] [uncultured Brachyspira sp.]
CLPJFFLI_00173	1.41e-26	ref WP_372710969 WP_372710969 glycerophosphodiester phosphodiesterase family protein [Hyphomicrobium sp.] [Hyphomicrobium sp.]
CLPJFFLI_00175	2.32e-52	ref WP_290167984 WP_290167984 phosphocholine cytidyltransferase family protein [Arthrobacter sp. APC 3897] [Arthrobacter sp. APC 3897]
CLPJFFLI_00176	9.47e-33	ref WP_149274885 WP_149274885 glycerol-3-phosphate cytidyltransferase [Pareuzebyella sediminis] [Pareuzebyella sediminis]
CLPJFFLI_00181	1.39e-23	ref WP_126262921 WP_126262921 DNA polymerase [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00182	9.05e-15	ref WP_273072499 WP_273072499 YfbR-like 5'-deoxynucleotidase [Marinobacter sp.] [Marinobacter sp.]
CLPJFFLI_00186	3.87e-54	ref WP_126232193 WP_126232193 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00187	1.78e-95	ref WP_126232176 WP_126232176 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]

CLPJFFLI_00189	5.21e-46	ref YP_006987239 YP_006987239 dCMP deaminase [Cronobacter phage vB_CsaM_GAP32] [Cronobacter phage vB_CsaM_GAP32]
CLPJFFLI_00190	1.07e-60	ref WP_006261870 WP_006261870 type IA DNA topoisomerase [Myroides odoratimimus] [Myroides odoratimimus]
CLPJFFLI_00193	4.12e-04	ref WP_242221582 WP_242221582 hypothetical protein [Shinella zoogloeoides] [Shinella zoogloeoides]
CLPJFFLI_00196	6.58e-42	ref WP_126232174 WP_126232174 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00198	1.11e-16	ref YP_007237818 YP_007237818 hypothetical protein G173_gp168 [Erwinia phage phiEaH2] [Erwinia phage phiEaH2]
CLPJFFLI_00200	5.79e-27	ref WP_036852753 WP_036852753 viroplasm family protein [Porphyromonas cangingivalis] [Porphyromonas cangingivalis]
CLPJFFLI_00201	6.61e-08	ref WP_282994502 WP_282994502 hypothetical protein [Serratia sp. Se-RSBMAAmG] [Serratia sp. Se-RSBMAAmG]
CLPJFFLI_00202	8.10e-04	ref WP_308644563 WP_308644563 ribonuclease H family protein [Oceanotoga sp.] [Oceanotoga sp.]
CLPJFFLI_00205	2.64e-09	ref WP_126232216 WP_126232216 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00207	4.73e-113	ref WP_126232178 WP_126232178 AAA family ATPase [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00208	1.19e-15	ref YP_006908159 YP_006908159 exonuclease [Campylobacter phage CP30A] [Campylobacter phage CP30A]
CLPJFFLI_00209	1.92e-25	ref WP_126232456 WP_126232456 hypothetical protein [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00210	2.84e-20	ref WP_301202003 WP_301202003 DNA polymerase I [Plebeibacterium marinum] [Plebeibacterium marinum]
CLPJFFLI_00214	5.37e-33	ref YP_010671370 YP_010671370 DnaB-like replicative helicase [Rhizobium phage RHph_Y65] [Rhizobium phage RHph_Y65]
CLPJFFLI_00215	5.98e-44	ref WP_126262881 WP_126262881 DEAD/DEAH box helicase [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00217	1.00e-23	ref WP_238526954 WP_238526954 dUTP diphosphatase [Glaciecola nitratireducens] [Glaciecola nitratireducens]
CLPJFFLI_00218	7.50e-38	ref WP_187648677 WP_187648677 ParM/StbA family protein [Nitrosophilus labii] [Nitrosophilus labii]
CLPJFFLI_00221	2.73e-19	ref WP_119445719 WP_119445719 exonuclease domain-containing protein [Hymenobacter oligotrophus] [Hymenobacter oligotrophus]
CLPJFFLI_00222	3.25e-07	ref WP_052954872 WP_052954872 hypothetical protein [Microvirga vignae] [Microvirga vignae]
CLPJFFLI_00223	1.92e-53	ref WP_126232480 WP_126232480 DNA gyrase subunit A [Campylobacter jejuni] [Campylobacter jejuni]
CLPJFFLI_00224	4.27e-95	ref WP_161969393 WP_161969393 toprim domain-containing protein [Campylobacter jejuni] [Campylobacter jejuni]

Supplementary Figures

Figure S1. MS/MS spectra of precursor ion 1074.3 in positive mode, fragmented with a normalized collision energy (NCE) of 10. A hexose residue was detected as well as a fragment corresponding to a thymidine moiety with a linker of MW 88 Da.

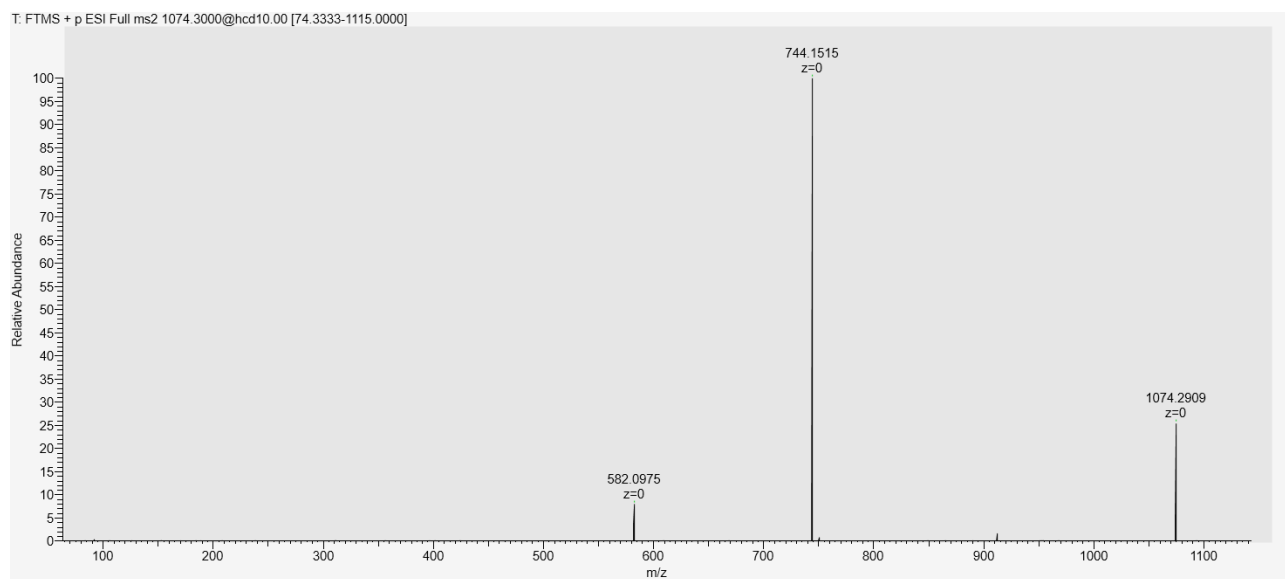


Figure S2. MS/MS spectra of precursor ion 1074.3 in positive mode, fragmented with a normalized collision energy (NCE) of 25. A hexose residue and an *N*-acetyl glycosamine molecule were detected.

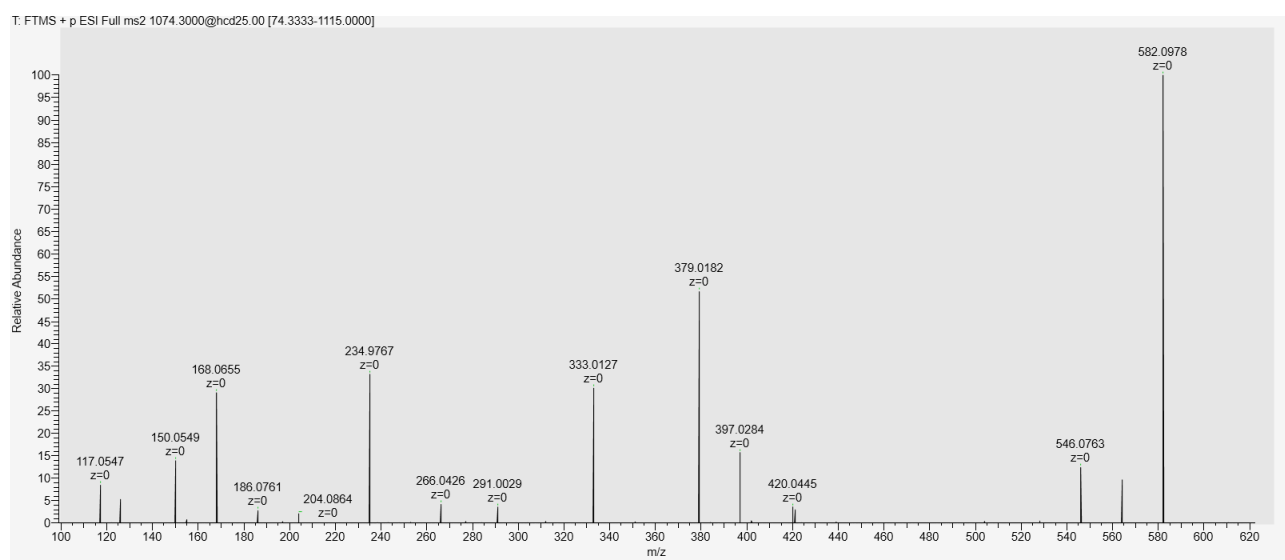


Figure S3. Portion of the MS/MS spectrum of precursor ion 535.6 (i.e., 1072.3, $z = 2$) in negative mode, fragmented with a normalized collision energy (NCE) of 30. A phosphate ion was detected.

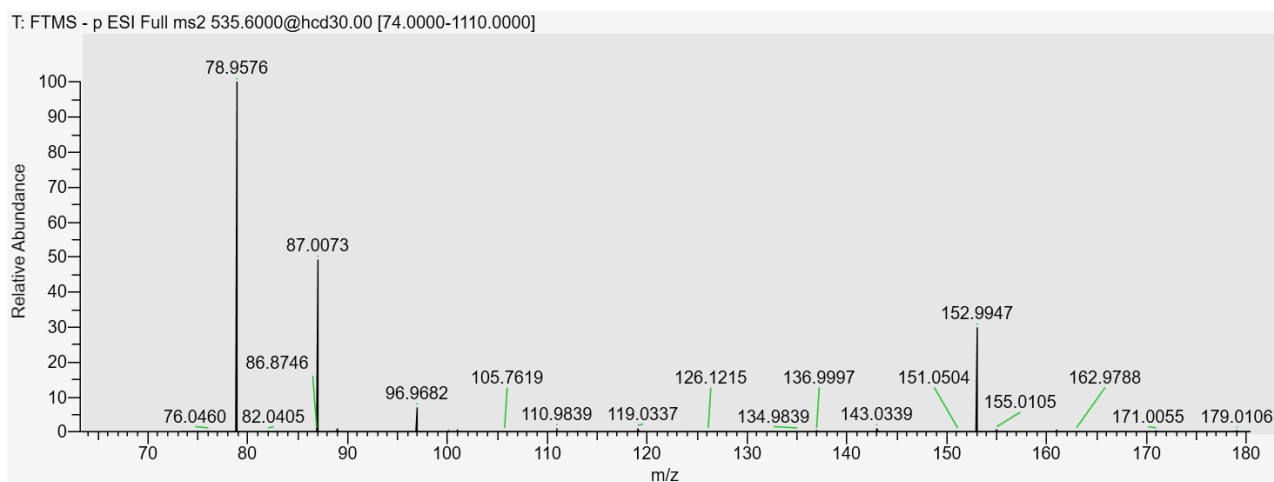


Figure S4. Genome map of the *Cobetia* bacteriophage Carin-5. Light pink arrows and light green arrows in the outermost circle indicate the directions and sizes of the genes and tRNAs respectively. A selection of genes and tRNAs are annotated. The two innermost circles represent GC skew $[(G-C)/(G+C)]$ and G+C content. The black circle represents the G+C content. An outward direction indicates that the G+C content of that region is greater than the average G+C content of the whole genome, and an inward direction indicates that the G+C content of that region is lower than the average G+C content. The green/magenta circle represents the GC skew $[(G-C)/(G+C)]$. The dark green color (outward direction) indicates that the GC skew $[(G-C)/(G+C)]$ is larger than zero, and the dark magenta color (inward direction) indicates that the GC skew $[(G-C)/(G+C)]$ is lower than zero. The scale units are base pairs. Circular maps with CDS, GC skew, and GC content were generated using the online Proksee visualization system [22].

