

Figure S1. Bidirectional clustering heatmap visualizing VIRIDIC-generated similarity matrix for the crAssE-Sib phage, phage phi14:2, and two prophages.

```

TR1  MRLVEQHIIKQSSIIYYNELQDLLHKCKNLNKGLYVVRQHYFQYKNDNTVKYKYLNYYSLEVLKTENDVDYRAL
TR2  MISYKY--NIYHSKKNKRLIKMLRECCFVWNHALQRRYYRLFG----KYIPLNKMQKHF---AKRIKRSIL
      *   :           *   :.* :.*::*   ::*: *   : *::*   :           **  **   .   :   :   *

TR1  PAPVAQQVLMVDRNFKSFFNLFNKKNRGEYSEFVRMPKYLNKDGLFPAVFTTASFSSQKWKIKQGIVKLPKQFSFT
TR2  HSQTAQEILRRLDNAYNRFFKKLSK-----RPPKFKRADKFSFVFTQSGFTLNGNCFTINKGKKRFKFS
      :  .**::*   :*. :. ** :. :.*           * ** : . * : . *** :.* : :           * *  *.** :

TR1  -TRTNKQNIQQLRFPKNGY--IVLEIVYNKKEKDLMSDNGNYLGIDIGLDNLASCVSNNGSCFIINGRPLKSI
TR2  YSRKYEGRIKQVRIVRETCNRFSLIIVTDHNPINNCRKTYDGASTGLDFGLKTYLTKSDGTKIESTLFFKQHQNK
      :*. :  .*:**:* *  .           :::   :*   : :   : *   ***:**.. :   ...   :   :   :.

TR1  NQYYN---KRLAFLKS---KLKDNKHTSKQIRSLTNKRNNKIKDYLHKASRILVNHVVSNGINTIIGHNKCWK
TR2  IRKLHKKLSRAEVGSNNRKRILFDLQQEYRKIRNLRNDFQWKLHSLCKEY---DYIFIEDLN--IEGMKRLWG
      :   :   :   .   :.   * * : :   :***.* * . : * : . * *           ::: :. : *   * * : : *

TR1  QEIN-IGKRNQNFVSIPFNVFISMISYKATLEGINVKIVEESYTSKCSFLDNERICKHESYKGRRTKRGLEKTS
TR2  KKISDLGHSSFINK-----LIFISL-KYDTVVHKID-KWYPSSKTCEGCINKNLSLK-----DRVWICPK
      :.*. :.* : .   *           :***: .*. :. :. * : *   . * *. :. : :   *           . *   : ..

TR1  FGRTINADINGAFNIIRKSEKESF-DVTMLPEGRGFWWN--PVRISV
TR2  CGAINDRDVLAARNILRKGISELESEGNTSSRNTGVPCVSIQESDSL
      *   : * : . * **:*.. *   :   .   . . *   .           * :

```

Figure S2. Alignment of the crAssE-Sib transposases sequences.

```

AR1      MGTNVSNFQNNAKNSNIILTSESNEMEFKSKEVKTVSSFKNSDFGELKIII
AR2      M-----DKVSVFESSDFGELRIIV
          *                               ..** *:*****:**:

AR1      ID-EEPYFIGSPIASFLGYTNPRKAIRDHVEDDDRIMKVPDTQ-GWNET
AR2      DPKGDVWFVASDVAKSLGYINAKDAVKRHVDDDDSIILQVSDNQWGVNQS
          : :*:.* :*. *** *.:.*:: ***:** :::*.*. * * *:

AR1      FRPYTPNTKILIIINESGLYSLIFGSKMDFAKKFKKWVTSEVLPSIRKTGS
AR2      ILKTRYIDNIRIINESGLYSLILSSKLESAKRFKRWITSEVLPSIRKTGE
          :      :* *****:.*:: **:**:*****:*****.

AR1      YSITPKDY-----PSALRALADEIDAKNRAIAERAQAEAERQQAI
AR2      YKTSSGGKGILVPDFSNPADAARAWADQYEAQKAIAEKSQAEAEKQQAL
          *. :. .      ..* ** *: :* :*:*****:*****:**:

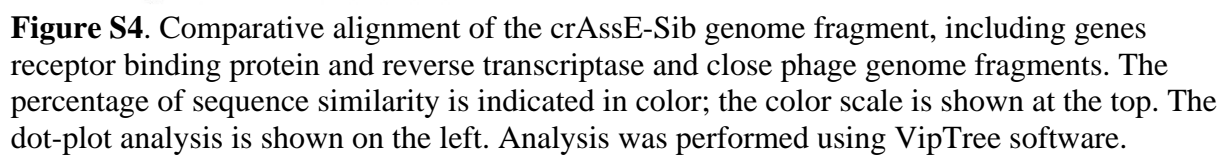
AR1      KTIEEQRPDVEFAESFKKVDHENMWLIRDVAKKLEQNGIIIAEKNLRLFL
AR2      KTIEEHKPDVEFAESFRKVDHNNMWLIRDIAKKLEQNGVIAEKNLRSFL
          *****:*****:*****:*****:*****:***** **

AR1      EEVKFMFRNGQGRWELYSDIVKNKFGVYRSYFVDKYSGERVNQQTIIYMTG
AR2      EEAKFMFRNGLGKWELYSNVVVKGYGVYRSYFIDKYSGDRINQQTIIYMTG
          **.****** *:*****:* : :*****:*****:*:*****

AR1      AGYEVTLNGIKGKCRNTFLNYGKFEGHNF
AR2      SGYEVTLNGIKGLKNVFLKYGKFA----
          :***** :*.**:*

```

Figure S3. Alignment of prophage antirepressor sequences.



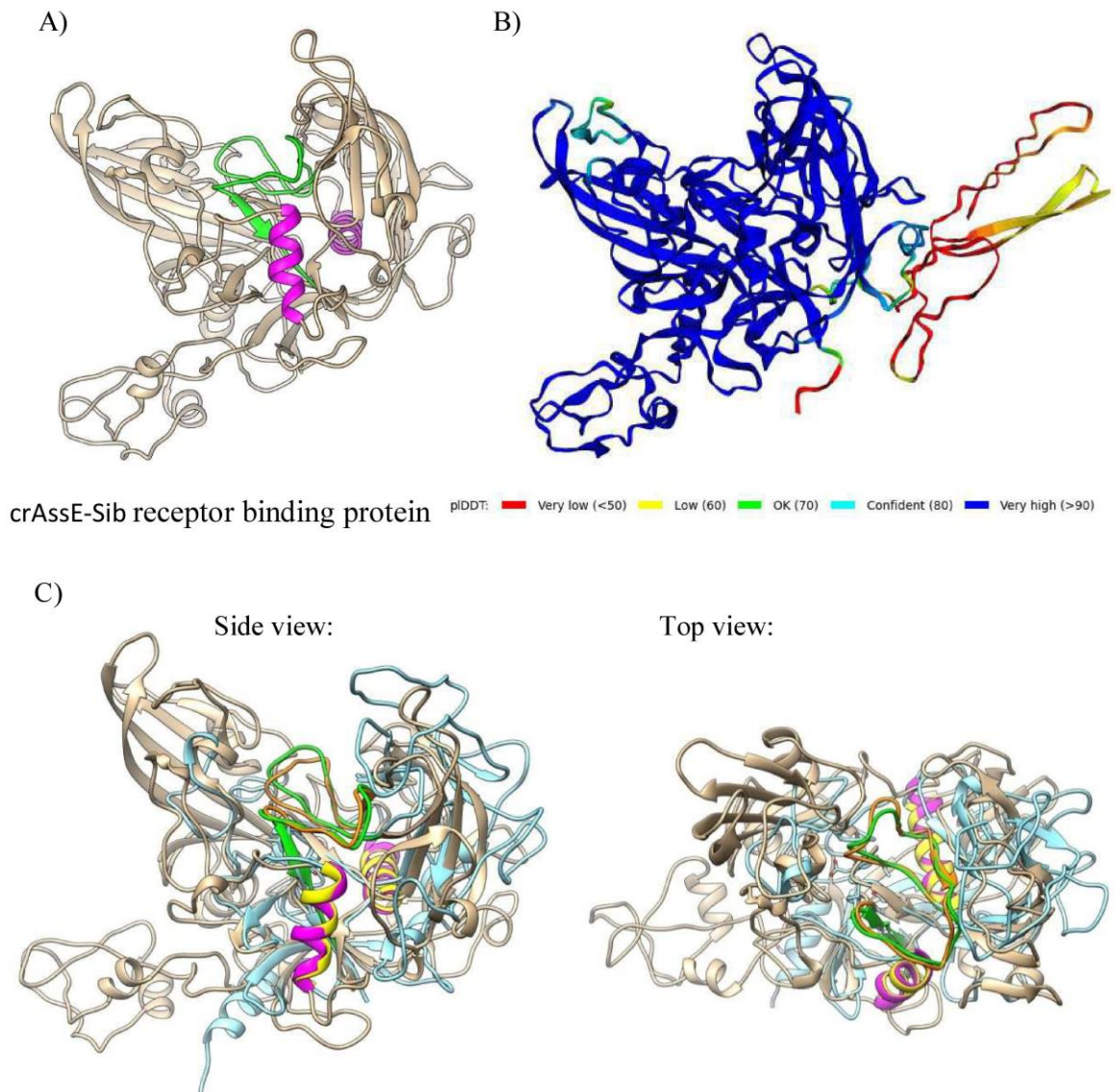


Figure S5. 3D model of crAssE-Sib receptor binding protein (RBD). A) 3D model of crAssE-Sib RBD from 85 aa to 678 aa; VR region is colored in green. B) AlphaFold2 model of this protein with probability that is shown by color according to pLDDT scale. N-terminal region is shown. C) 3D model of crAssE-Sib RBP (shown in tan) superimposed at TaqVP crystal structure pdb 5VF4 (shown in light blue); VR regions of RBD and TaqVP are in green and orange, respectively. Alpha helices (colored in magenta and yellow) show similar orientation of the molecules.