Supplementary Dataset 6: Identification of LCDV-1 in human enteric virome.

The data for this analysis was obtained from NCBI (SRA ID: PRJNA196801) and the name of the project is “Rapid evolution of the human gut virome - BioProject”. We report the BlastN reads that show a significant alignment with LCDV-1 sequences. No significant hits were identified for LCDV-Sa, GIV and SGIV in this dataset. Human herpes virus and human parechovirus were used as controls. The full length reads that gave a significant hit for GIV are reported at the end of the document. The full length read that gave a significant hit for LCDV-1 is reported at the end of the document.

1- LCDV-1

```
Query= NC_001824.1 Lymphocystis disease virus 1, complete genome

Length=102653
                                                                      Score     E
Sequences producing significant alignments:                          (Bits)  Value

SRR935349.6447861  HWI-ST423:258:C0MLUACXX:6:1216:19027:50697 len...  54.7    0.28


>SRR935349.6447861 HWI-ST423:258:C0MLUACXX:6:1216:19027:50697 length=202
Length=202

 Score = 54.7 bits (29),  Expect = 0.28
 Identities = 34/36 (94%), Gaps = 1/36 (3%)
 Strand=Plus/Minus

Query  57275  ATTG-TAATTTAGATTAAATTTAAATTAAACCATGC  57309
              |||| | |||||||||||||||||||||||||||||
Sbjct  129    ATTGATGATTTAGATTAAATTTAAATTAAACCATGC  94



Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 7966312234334142
```

2- LCDV-Sa

```
Query= NC_033423.1 Lymphocystis disease virus Sa isolate SA9, complete
genome

Length=208501


***** No hits found *****



Lambda      K        H
```

```
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 16084328486396742
```

3- SGIV

```
Query= NC_006549.1 Singapore grouper iridovirus, complete genome

Length=140131


***** No hits found *****



Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 10809070579090602
```

4- GIV

```
Query= AY666015.1 Grouper iridovirus, complete genome

Length=139793


***** No hits found *****



Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 10782991345908366
```

5- Human herpesvirus 7

```
Query= NC_001716.2 Human herpesvirus 7, complete genome

Length=153080


***** No hits found *****
```

```
Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 11808182976950880
```

6- Human parechovirus 1

```
Query= JX575746.1 Human parechovirus 1 isolate CAU10-NN, complete genome

Length=7348


***** No hits found *****



Lambda      K        H
    1.33    0.621     1.12

Gapped
Lambda      K        H
    1.28    0.460    0.850

Effective search space used: 578105862421152


  Database: /n/data1/joslin/cores/bbcore/emrah_virome/PRJNA196801/fasta/gut_micr
obiome_reads.fasta
    Posted date:  Sep 29, 2017  4:06 PM
  Number of letters in database: 96,208,727,862
  Number of sequences in database:  476,280,831



Matrix: blastn matrix 1 -2
Gap Penalties: Existence: 0, Extension: 2.5

BLASTN 2.6.0+


Reference: Zheng Zhang, Scott Schwartz, Lukas Wagner, and Webb
Miller (2000), "A greedy algorithm for aligning DNA sequences", J
Comput Biol 2000; 7(1-2):203-14.



Database: /n/data1/joslin/cores/bbcore/emrah_virome/PRJNA196801/fasta/gut_micr
obiome_reads.fasta
           476,280,831 sequences; 96,208,727,862 total letters
```

The full length read that gave a significant hit for LCDV-1

```
>SRR935349.6447861 HWI-ST423:258:C0MLUACXX:6:1216:19027:50697 length=202
ATATGAATGAAATTGCTAGCTTTATCTTAGATTCAGATGTTGCAGGAGGGAGACACATCGCCAAAGCGTTAAATGACCCT
AAGACTTTAGTTGGCATGGTTTAATTTAAATTTAATCTAAATCATCAATTGTTATTACCTTCTTAGCAGGGGCACCCGTA
CTACCTGCTGTTTTTTTGACCACTGTCTTAGCTGCATTTGGA
```