Execution steps:

A) **Prerequisites and input**

1. To run on the local computer, we need to install python (ok with ***Python version 3.6*** and above).
2. Create a directory, say the name: **AAA** (it can be any name)
3. For inferring host-virus interactions, we need two source fasta format files (CDS, nucleotide sequences): host (say, ***H.fasta***) and virus (say, ***V.fasta***), inside the same directory, **AAA.**
4. Save the below python script file (without any change) as the name specified: **X) *main.py***, and also keep it inside the same directory, **AAA.**
5. There are several python modules (***Bio (SeqIO), CAI (RSCU), scipy.stats***) called from inside the ***main.py*** script; make sure those are already installed with python.

B). Execution:

1. The only ***main.py*** is needed for execution (or RUN); by default, **p-value is set to p<.001** for edge (host-viral interaction) selection with strong correlation score. Further, one can modify/add a single code for setting the correlation value range/cut off.
2. The only ***main.py*** is needed for execution (or RUN); by default, **p-value is set to <.001**, which can be set as needed (strong correlation for lower p-value cut-off) from the ***main.py*** file. Further, one can modify/add a single code for setting the correlation value range/cut off.

C) Output:

3. Output results will be saved as ***result.csv,*** where each row contains four entries (***host id; virus id, correlation value, p-value).*** The host id and virus id are taken as it is from the fasta files. After modifying the csv file, it can be utilized for host-virus network reconstruction using Cytoscape and other software tools.

X) main.py

```
from Bio import SeqIO

from CAI import RSCU

import scipy.stats

N=['A','T','C','G']

C=[]

for i in range(4):

    for j in range(4):

        for k in range(4):

            C.append(N[i]+N[j]+N[k])

[C.remove(x) for x in {'ATG','TAA','TAG','TGA','TGG'}]

fid=open('Result.csv','w')

for seq1 in SeqIO.parse('H.fasta','fasta'):
```

```
    ref1=[]
    ref1.append(seq1.seq)
    a=RSCU(ref1)
    x=[0 for i in range(59)]
    for i in range(59):
        x[i]=round(a[C[i]],3)
    for seq2 in SeqIO.parse('V.fasta','fasta'):
      ref2=[]
      ref2.append(seq2.seq)
      b=RSCU(ref2)
      y=[0 for i in range(59)]
      for i in range(59):
        y[i]=round(b[C[i]],3)
      r,p=scipy.stats.pearsonr(x,y)
      if p<.001:
        fid.write('\n%s;%s;%s;%s'%(seq1.id,seq2.id,round(r,3),p))
        print(seq1.id,seq2.id,round(r,3),p)
fid.close()
```