

# Phages with a broad host range are common across ecosystems

---

In the format provided by the  
authors and unedited

1    **Table of contents**

2

3

4    **Supplementary Figures 1-3**

5

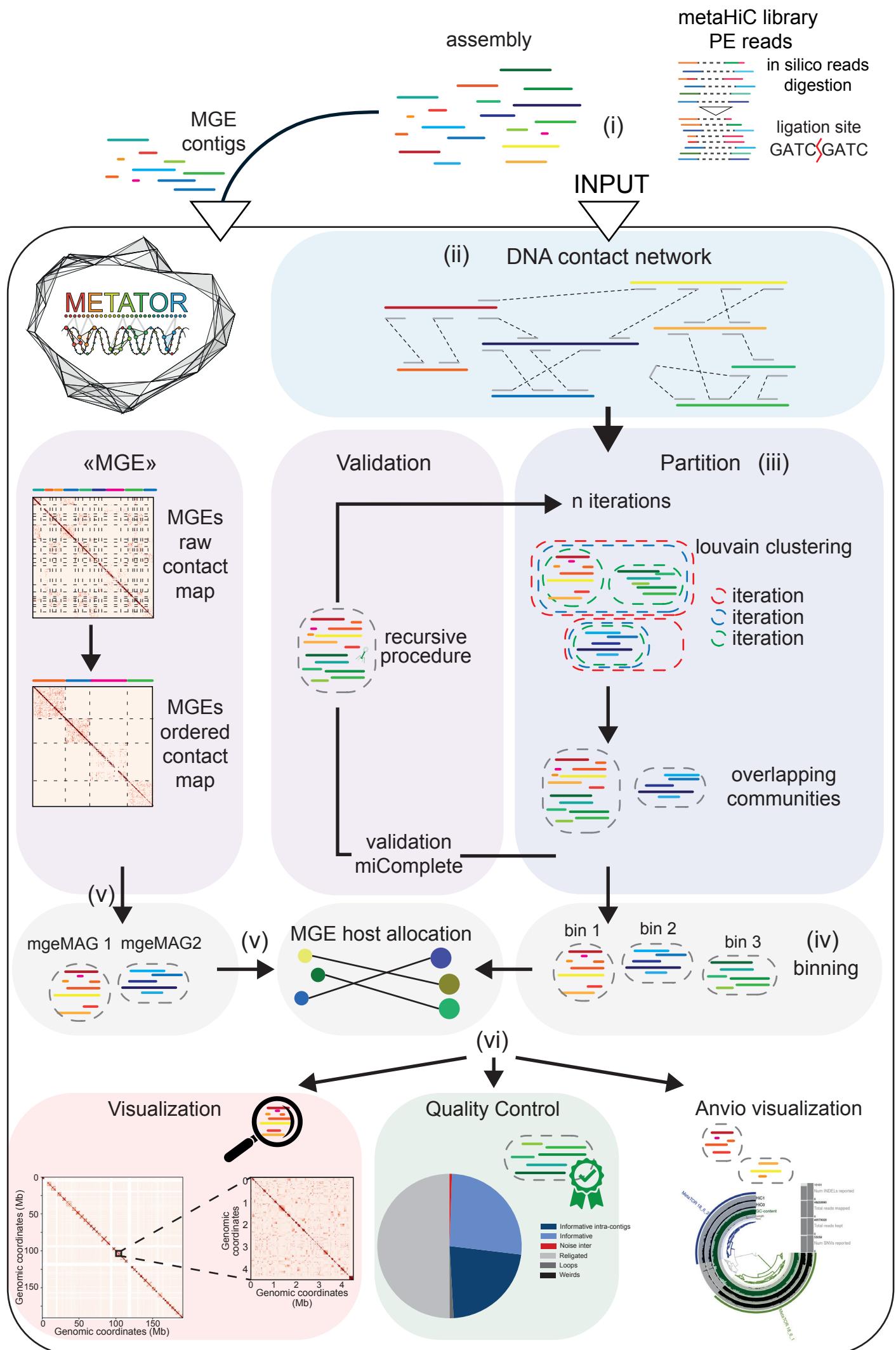
6

7   **Supplementary Figure 1: Overview of the new MetaTOR pipeline.**  
8   *i)* inputs correspond to HiC PE reads and an assembly of the studied sample. A new module  
9   allows to pre-digest the reads based on the restriction enzyme(s) used during the HiC library  
10   generation in order to increase the mapping rate and the 3D signal. *ii)* the contig network of  
11   interactions based on HiC PE reads is computed and, eventually, normalized. *iii)* the network  
12   is partitioned using an iterative procedure of the louvain algorithm and the group of contigs  
13   (overlapping communities) are recovered based on a threshold of the iterative procedure. *iv)*  
14   bins above 500 kb are evaluated using miComplete and, eventually, refined using a recursive  
15   procedure of the louvain algorithm and evaluation of the sub-communities using miComplete.  
16   Final bins with the best ratio completion / contamination are then generated. *v)* a new module  
17   (*i.e.* MetaTOR mge) allows to bin contigs (*i.e.* contigs annotated as MGE in our case) of user  
18   choice based on their intra- and inter-contact signal. MGE bins are then extracted and assigned  
19   to microbial MAGs on the base of the interactions signal. *vi)* different modules allow to 1-  
20   generate contact map of any objects (overlapping communities, final bin, MGE bin...); 2-  
21   generate different metrics of the HiC libraries quality (3D signal, intra vs. inter contigs  
22   signal...); 3- generate different files that serve as input for the Anvio platform.

23

24

# Supplementary Figure 1



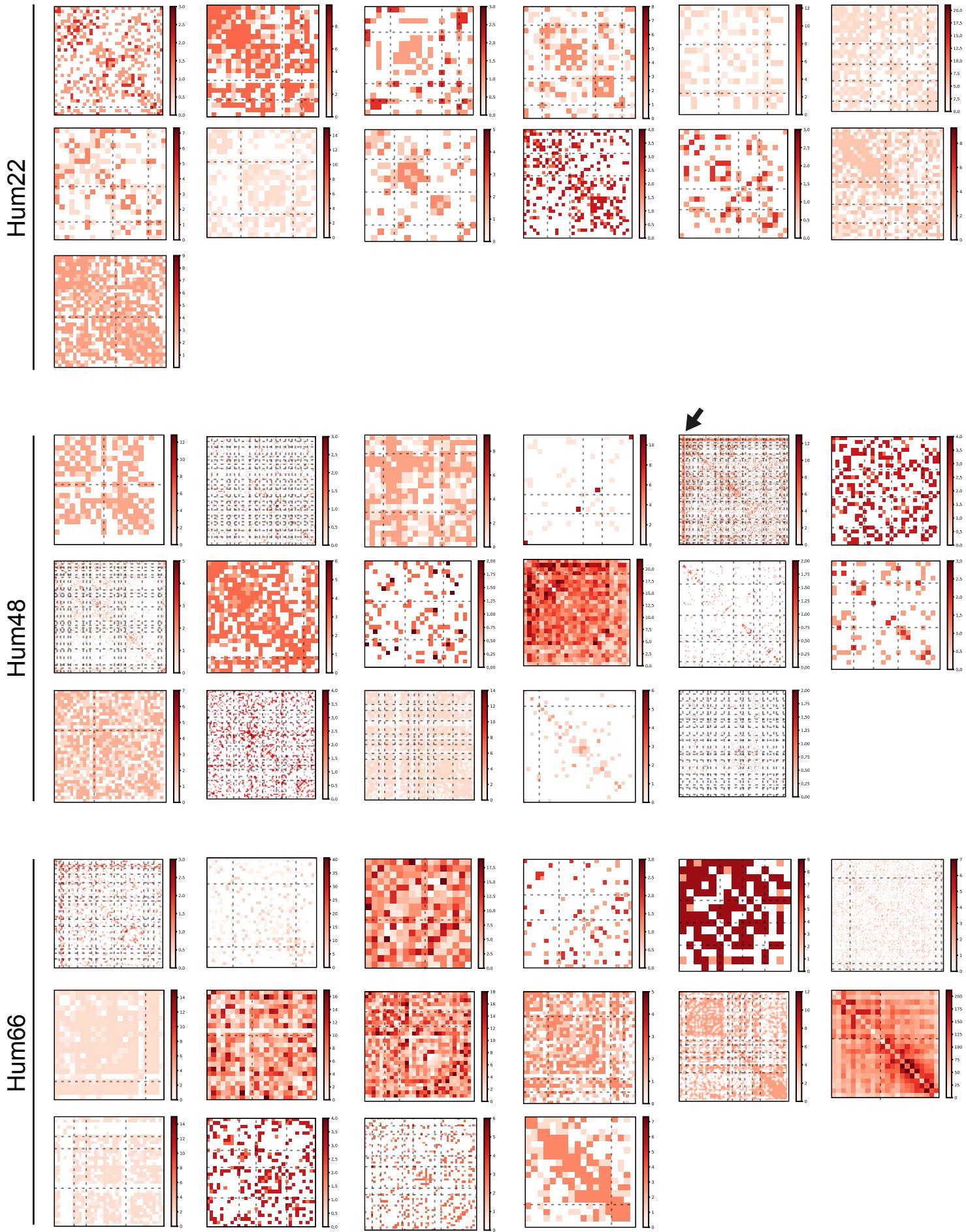
25 **Supplementary Figure 2: Contigs contact map of the complete and HQ vMAGs**  
26 **encompassing several contigs obtained for the Hum22, Hum48 and Hum66 samples.**

27 Raw contact map of the 46 vMAGs obtained on samples Hum22, Hum48 and Hum66 assessed  
28 as complete or high-quality by CheckV and encompassing several contigs. Dashed lines  
29 represent borders of the different contigs. Samples are indicated on the left while scale bars are  
30 present on the right. Black arrow points to the only vMAGs exhibiting an aberrant signal. Main  
31 diagonal was set to zero prior to binning and plotting of the 3D signal.

32

33

## Supplementary Figure 2



34 **Supplementary Figure 3: Benchmark of MetaTOR pipeline.**

35 **a.** Bar plot of the number of complete, HQ and MQ viral genomes assessed by CheckV for the  
36 raw contigs and the vMAGs obtained using MetaTOR, SemiBin or ViralCC for 3 human gut  
37 datasets (Hum22, Hum48 and Hum66). **b.** Pie chart of the proportion of circular viral contigs  
38 either left alone or binned with other contigs. **c.** Example of contact map of ViralCC and  
39 SemiBin bins encompassing circular viral contigs. Black arrows indicate circular signals in the  
40 contact map.

41

42

43

# Supplementary Figure 3

