

**Table S1: Biome-averaged genome length estimated by GAAS for the metagenomes of each environment. The numbers reported are: mean (median)  $\pm$  standard deviation**

| Biome              | Sub-biome     | Average viral genome length (kb) | Average bacterial and archaeal genome length (kb) | Average protist genome length (kb) |
|--------------------|---------------|----------------------------------|---|------------------------------------|
| Aquatic (total)    | -             | 79.9 (61.4) $\pm$ 59.4           | 3,020 (2,970) $\pm$ 1,150                         | 2,690 (807) $\pm$ 6,060            |
| Aquatic            | Ocean         | 102 (103) $\pm$ 53.3             | 2,580 (2,150) $\pm$ 1,120                         | -                                  |
| Aquatic            | Hypersaline   | 91.9 (66.2) $\pm$ 72.2           | 3,250 (3,430) $\pm$ 925                           | -                                  |
| Aquatic            | Freshwater    | 42.0 (41.4) $\pm$ 8.23           | 4,240 (4,130) $\pm$ 313                           | -                                  |
| Aquatic            | Hot spring    | -                                | 2,820 (2,560) $\pm$ 1,340                         | -                                  |
| Aquatic            | Stromatolites | 24.1 (17.3) $\pm$ 23.6           | 4,560 (4,500) $\pm$ 157                           | -                                  |
| Sediments          | -             | 85.6 (85.6) $\pm$ 4.28           | 4,370 (4,370) $\pm$ 803                           | -                                  |
| Terrestrial (soil) | -             | -                                | 5,910 (5,930) $\pm$ 218                           | -                                  |
| Host-associated    | -             | 33.2 (39.8) $\pm$ 21.2           | 3,150 (3,190) $\pm$ 1,420                         | -                                  |