

## S2 Table. Overview of reference sequence databases and associated download information.

Reference sequence information can be obtained from the URL's shown in 'Download information'.

<b>Database name</b>	<b>Download information</b>
<b>MetaHitAssembly<sup>2</sup></b>	<a href="http://www.ncbi.nlm.nih.gov/nuccore/?term=PRJEB674">http://www.ncbi.nlm.nih.gov/nuccore/?term=PRJEB674</a> (PRJEB674 - PRJEB1046)
<b>Bacteria<sup>2</sup></b>	<a href="ftp://ftp.ncbi.nih.gov/genomes/genbank/bacteria">ftp://ftp.ncbi.nih.gov/genomes/genbank/bacteria</a>
<b>Human</b>	<a href="http://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.25/#/def_asm_Primary_Assembly">http://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.25/#/def_asm_Primary_Assembly</a> Human genome ver. GRCh37.p13
<b>Invertebrates</b>	<a href="ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/invertebrate">ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/invertebrate</a>
<b>Protozoa</b>	<a href="ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/protozoa">ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/protozoa</a>
<b>Virus<sup>1</sup></b>	<a href="ftp://ftp.ncbi.nih.gov/genomes/virus">ftp://ftp.ncbi.nih.gov/genomes/virus</a>
<b>Virus_NCBI<sup>1</sup></b>	<a href="http://www.ncbi.nlm.nih.gov/genbank/">http://www.ncbi.nlm.nih.gov/genbank/</a>
<b>Virus_Vipr<sup>1</sup></b>	<a href="http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3245011">http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3245011</a>
<b>Plants</b>	<a href="ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/plant">ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/plant</a>
<b>Fungi</b>	<a href="ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/fungi">ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/fungi</a>
<b>Parasites</b>	<a href="ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/parasites">ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/parasites</a>
<b>Vertebrates</b>	<a href="ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/vertebrates_mammals">ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/vertebrates_mammals</a> <a href="ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/vertebrates_other">ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/vertebrates_other</a>

<sup>1</sup> All viral databases were combined into one. <sup>2</sup> All bacterial databases were combined into one.