

Sample biome	Sub-biome	Project identifiers	Number of assemblies
Aquatic	Freshwater	<i>PRJEB22523, PRJEB22543, PRJEB22524</i>	4
	Glacier	<i>PRJEB22516</i>	3
	Lentic	<i>PRJEB22499</i>	1
	Sediment	<i>PRJEB22502, PRJEB22517, PRJEB22496</i>	7
	Marine	<i>PRJEB22527, PRJEB22092</i>	142
	Coastal	<i>PRJEB22525, PRJEB22509</i>	5
	Hydrothermal vents	<i>PRJEB22510, PRJEB22514, PRJEB22397</i>	14
	Salt marsh	<i>PRJEB22494</i>	1
	Oceanic	<i>PRJEB22493, PRJEB22522</i>	9
	Photic zone	<i>PRJEB22394</i>	11
	Alkaline	<i>PRJEB22526</i>	1
Engineered	Activated sludge	<i>PRJEB22508, PRJEB22512, PRJEB22550, PRJEB22521</i>	10
Host-associated	Human	<i>PRJEB22283, PRJEB22387, PRJEB22388, PRJEB22556</i>	821
	Human, faecal	<i>PRJEB22492, PRJEB22505, PRJEB22513, PRJEB22495, PRJEB22501, PRJEB22498, PRJEB22392, PRJEB22359, PRJEB22360, PRJEB22519, PRJEB22529, PRJEB22395, PRJEB22538, PRJEB22365, PRJEB22542, PRJEB22391, PRJEB22547, PRJEB22549, PRJEB22389, PRJEB22393, PRJEB22552, PRJEB22368, PRJEB22554, PRJEB22362</i>	1108
	Human, large intestine	<i>PRJEB22528</i>	1
	Human, oral	<i>PRJEB22520</i>	1
	Human, skin	<i>PRJEB22555</i>	5
	Rumen	<i>PRJEB22504, PRJEB22623</i>	6
	Soil	<i>PRJEB22518, PRJEB22546, PRJEB22551, PRJEB22533, PRJEB22511, PRJEB22503, PRJEB22544, PRJEB22540, PRJEB22396, PRJEB22366, PRJEB22627, PRJEB22625, PRJEB22624</i>	129
	Agricultural	<i>PRJEB22531</i>	1
	Forest soil	<i>PRJEB22506</i>	5
	Grasslands	<i>PRJEB22497, PRJEB22500, PRJEB22628</i>	5
	Permafrost	<i>PRJEB22530</i>	4
	Sand	<i>PRJEB22545</i>	3
	Wetlands	<i>PRJEB22541</i>	1

Supplementary Table 1. Distribution of EBI Metagenomics assemblies according to source biome. The table lists the source biome, sub-biome (where available) and ENA identifiers of projects that have been assembled and analysed by EBI Metagenomics. The total number of assemblies for each sub-biome is also shown.