**Additional File 2**

**Table S1 | Identification Sensitivity of target and contamination species for QC-Blind-processed simulated datasets.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Sensitivity** | **Species** | **Genus** | **Family** |
| **Simu\_BS\_5%** | target sensitivity | 0% | 100% | 100% |
| **Simu\_BS\_35%** | target sensitivity | 0% | 100% | 100% |
| **Simu\_BS\_65%** | target sensitivity | 0% | 100% | 100% |
| **Simu\_BS\_95%** | target sensitivity | 0% | 100% | 100% |
|  | Average | 0% | 100% | 100% |
| **Simu\_BS\_5%** | contamination sensitivity | 20% | 90% | 100% |
| **Simu\_BS\_35%** | contamination sensitivity | 10% | 90% | 100% |
| **Simu\_BS\_65%** | contamination sensitivity | 20% | 80% | 100% |
| **Simu\_BS\_95%** | contamination sensitivity | 10% | 90% | 90% |
|  | Average | 15% | 88% | 98% |

Identification of target and contamination at family, genus and species level are shown in this table.

**Table S2 | Assembly Statistics for simulated datasets.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **BS 5%** | **BS 35%** | **BS 65%** | **BS 95%** |
| **Velvet N50** | 39369 | 49287 | 47030 | 24697 |
| **Velvet #contig** | 10667 | 6193 | 6232 | 11150 |
| **MEGAHIT N50** | 154200 | 154200 | 154200 | 27129 |
| **MEGAHIT #contig** | 603 | 620 | 599 | 4135 |

N50 and number of contigs were derived from Velvet and MEGAHIT in simulated data.