|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Title  | Origin  | Study accession\*  | Sample accession\*  | Sequencing technology  | #Reads  | Size (bp)  |
| Muddy  | Marine environment  | ERP008945  | **ERS624612**  | Illumina MiSeq | 18 831 892  | 4,3 billion  |
| Sandy  | Marine environment  | ERP008945  | ERS624613  | Illumina MiSeq | 7 271 336 | 1,8 billion  |
| Moose  | Gut metagenome  | ERP010530  | ERS624611  | Illumina MiSeq | 30 966 692  | 6,3 billion  |
| Sea Urchin  | Gut metagenome  | ERP010530  | ERS738393  | Illumina Miseq | 38 716 356 | 6,6 billion  |

**Supplementary Table 1: Datasets used in the present analysis of the two pipelines.**

\* Datasets are available at the European Nucleotide Archive