

## Supplementary Information

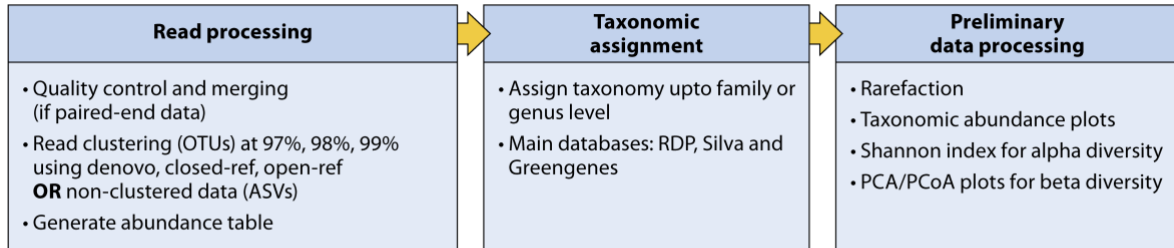
**S1: Popular DNA extraction kits and their lysing mechanisms for bacterial cell wall in different body fluid/tissue categories**

<b>DNA extraction kit</b>	<b>Lysing mechanism in kit protocol</b>	<b>Modification</b>	<b>Body fluid/tissue</b>	<b>References</b>
MoBio PowerSoil kit	Mechanical lysing with beat-beating and enzymatic lysis at RT	Reduce or omit mechanical lysis to prevent overshearing	Fecal samples	(43)
QIAmp DNA mini kit and QIAmp DNA stool kit	Enzymatic lysis with heat at 70°C for 5 min or increase temp to 95°C	Add mechanical bead-beating step with zirconia beads	Fecal samples	(41)
Qiagen DNeasy Blood and Tissue kit	Enzymatic lysis with heat at 56°C for 10 min	Only enzymatic lysis and no mechanical lysis	Vaginal samples	(42)
MoBio PowerLyser Powersoil kit	Mechanical lysing with beat-beating and enzymatic lysis at RT	-	Fecal samples	(44)

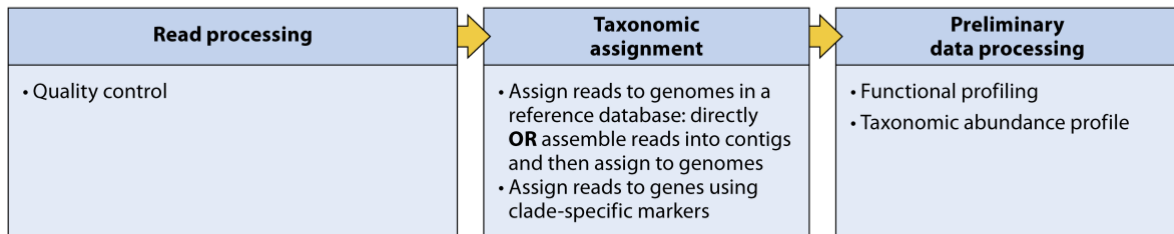
## S2: Bioinformatics pipeline for amplicon sequence data and shotgun data

(A) and (B) depict a basic pipeline and steps involved in processing of amplicon sequence data and shotgun data respectively. (C) Enlists potential options for using machine learning algorithms for making predictions based on amplicon sequence and shotgun data.

### A. Amplicon data



### B. Shotgun data



### C. Predictive modelling

Train prediction classifiers using machine learning algorithms like:

