

Methods :

Dataset		TARA		HMP	
Number of samples		139		835	
Reads		Tara filtered reads		Hmp filtered reads	
Database		freeze11	mOTUs	freeze11	mOTUs
Aligner		BWA			
	Alignment length	>= 75			
	Similarity	97 %			
	Multimappers	No			
SNP calling		metaSNV			
	Vertical coverage	10X (Avg)			
	Horizontal coverage	40% (Avg)	60% (Avg)	40% (Avg)	60% (Avg)
	Allele prevalence	90 %			
	Min. number of samples meeting the above conditions	5		20	

Allele prevalence : percentage of samples where the position was observed at coverage >=5X