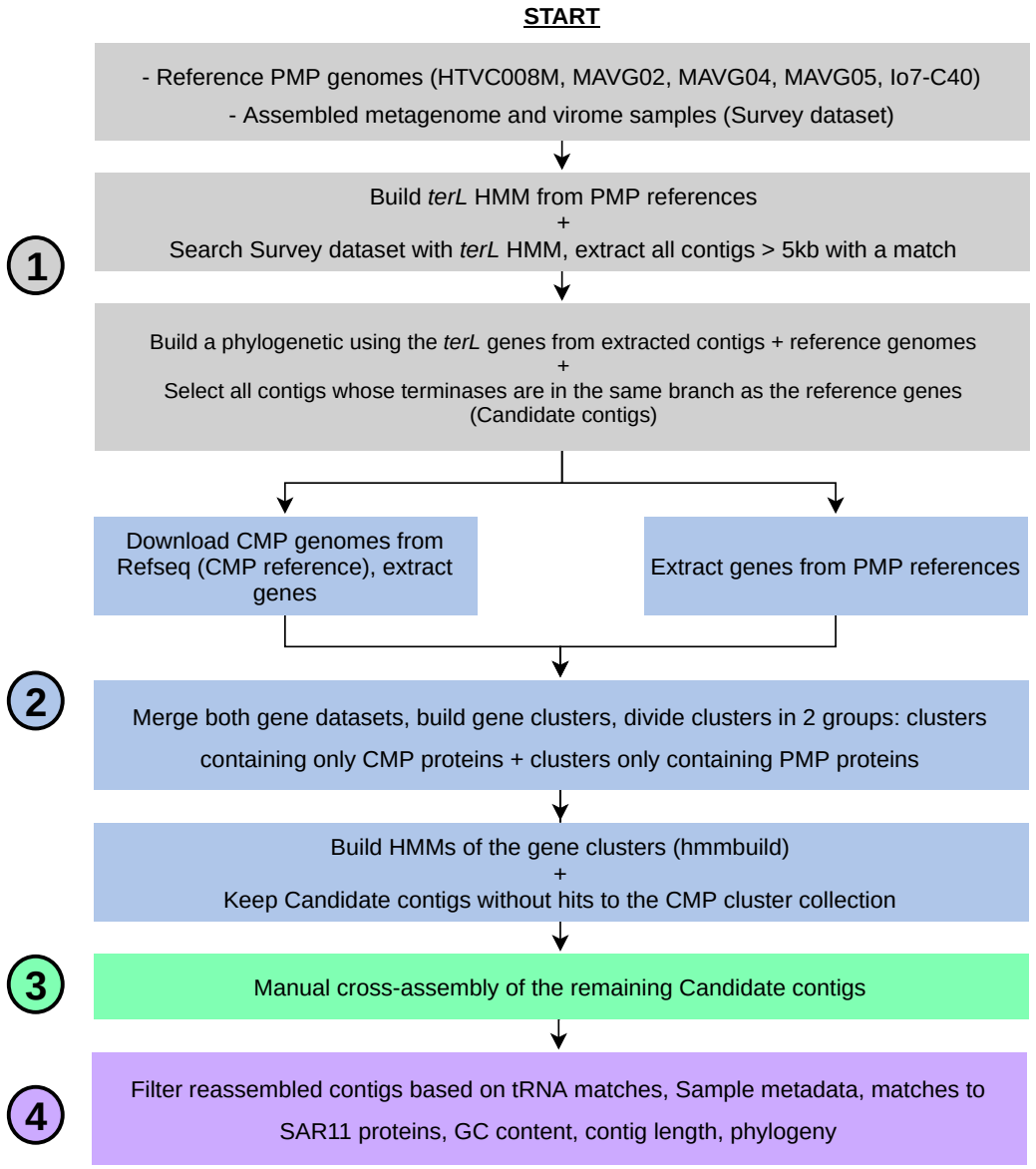
Genome mining flowchart

A



B

	Step	#Contigs	Average contig size (kbp)
1	*terL* hmm filtering	126,467	12.56
2	Gene cluster hmm filtering	30,331	25.89
3	Manual cross-assembly	28,431	26.42
3	Duplicate removal	14,748	28.30
4	Contigs > 100kbp	688	152.05
4	GC% between 30 – 35	197	143.03
4	Phylogeny + Gene filtering	31	137.19