Supplemental Table S1 | taxMaps databases.

Name	Included databases	N sequences (N taxa)	Total Length (Gb)	*K*-mer Length (nt)	*K*-mers (G)	Compressed *k*-mers (G)	Compression ratio	Required RAM (GB)
refseq_model_organisms	Selection of RefSeq genomes, including human	322129 (14)	16.683	75	16.66	14.49	1.15	29.0
				125	16.64	14.94	1.11	24.7
				150	16.63	15.05	1.11	24.3
				300	16.59	15.35	1.08	23.5
refseq_complete_genomes	Complete archaea, bacterial and viral genomes from RefSeq	16022 (9460)	18.323	75	18.32	9.70	1.89	23.0
				125	18.32	10.23	1.79	21.1
				150	18.32	10.42	1.76	20.7
				300	18.32	11.17	1.64	20.3
blast_nt	BLAST nt Database, excluding "other sequences" and "unclassified sequences"	34904805 (1109518)	114.518	75	111.95	62.07	1.80	166.1
				125	110.23	64.96	1.70	147.8
				150	109.38	65.77	1.66	143.1
				300	104.43	67.56	1.55	137.8
refseq_microbial	Archaea, bacteria, fungi, protozoa and viral genomes from RefSeq	6516024 (35641)	242.673	75	242.19	51.82	4.67	146.3
				125	241.87	56.30	4.30	134.9
				150	241.70	57.94	4.17	131.6
				300	240.76	64.68	3.72	131.4
combined_ncbi	*refseq_microbial* + *refseq_model_organisms* + *blast*_nt	41742958 (1130634)	373.873	75	370.80	106.77	3.47	304.6
				125	368.74	114.19	3.23	270.4
				150	367.72	116.67	3.15	262.6
				300	361.77	125.23	2.89	257.0
refseq_invertebrate	Invertebrate genomes from RefSeq	2388261 (101)	39.647	75	39.47	32.97	1.20	63.5
				125	39.35	33.67	1.17	56.6
				150	39.29	33.78	1.16	55.4
				300	38.95	33.80	1.15	53.4
refseq_plant	Plant genomes from RefSeq	931136 (68)	33.222	75	33.15	25.96	1.28	61.6
				125	33.11	27.35	1.21	50.0
				150	33.08	27.65	1.20	48.0
				300	32.96	28.27	1.17	44.8