Corpora	# of tokens	# of types	# of genotypes
K1	10006	2767	219
K2	34636	4714	241
K3	31520	5299	262
K0 (K1-3)	76162	10090	304