Supplementary Table S3. Comparison of SNP call sets (unfiltered and filtered) between GATK UnifiedGenotyper and GotCloud for whole exome sequencing data. 1000 exome BAM files were randomly selected from 1000G Phase I data. 10 sample results were averaged over 10 sets of 10 BAM files.

# Sample	Pipeline	Filter	# total SNP	# SNP per sample	# dbSNP (v129)	Nonsyn.		Synonymous	
						Known Ts/Tv	Novel Ts/Tv	Known Ts/Tv	Novel Ts/Tv
10	GotCloud	None	45,403	16,240	32,123	2.15	1.13	5.30	3.23
		SVM	41,479	14,859	31,251	2.19	1.54	5.38	4.17
	GATK	None	41,795	14,376	29,894	2.15	1.03	5.29	3.09
		VQSR	35,918	12,483	28,034	2.21	1.58	5.39	4.32
100	GotCloud	None	127,664	16,397	52,656	2.18	1.34	5.36	3.63
		SVM	109,023	14,121	50,613	2.25	1.94	5.44	4.89
	GATK	None	126,245	15,929	51,806	2.18	1.19	5.45	3.37
		VQSR	96,786	12,305	48,349	2.26	2.13	5.54	5.24
1,000	GotCloud	None	399,811	17,174	69,197	2.24	1.61	5.40	4.05
		SVM	324,381	13,801	65,251	2.33	2.31	5.55	5.49
	GATK	None	446,978	19,196	69,529	2.27	1.54	5.61	3.93
		VQSR	311,452	12,524	64,502	2.35	2.35	5.74	5.51