Table 1. Summary and Breakdown of DNA Variants

Type	Total Variants	Total High Confidence	Heterozygous High Confidence	Homozygous High Confidence
Total SNVs	3,739,701	3,301,521	1,971,629	1,329,892
Total gene-associated SNVs	1,312,780	1,183,847	717,485	466,362
Total coding/UTR	49,017	44,542	27,383	17,159
Missense	10,592	9,683	5,944	3,739
Nonsense	83	73	49	24
Synonymous	11,459	10,864	6,747	4,117
5′UTR	4,085	2,978	1,802	1,176
3′UTR	22,798	20,944	12,841	8,103
Intron	1,263,763	1,139,305	690,102	449,203
Ts/Tv	—	2.14	—	—
dbSNP	3,493,748	3,167,180	—	—
Candidate private SNV	245,953	134,341	—	—
Indels (−107∼ +36 bp)	1,022,901	216,776	—	—
Coding	3,263	302	—	—
Structural variants (>50 bp)	44,781	2,566	—	—
In 1000G project[a]	4,434	1,967	—	—

High confidence values are from variants identified across multiple platforms (Illumina and CG) and/or Exome and RNA-Seq data. Annotations were based from variant call formatted (vcf) files for heterozygous calls: 0/1, reference (ref)/alternative (alt); 1/2, alt/alt and homozygous calls; 1/1, alt/alt; 1/, (alt/alt-incomplete call). Polyphen-2 was used to identify the location of the SNVs.

[a]1000G (1000 Genomes Project Consortium, 2010).