Methods for the VCF-based Clinical Analysis of the *BRCA* region

To confirm our findings using tiling, we converted the 178 PGP participants Complete Genomics Information files to VCF files using cgatools mkvcf with parameters –reference=build37.crr –source-names=masterVar,SV,MEI,CNV and –field-names=GT,FT,HQ,EHQ,GQ. We obtained 69 1000 Genomes Project whole genome VCF files from complete genomics, and 433 1000 Genomes Project whole genome VCF files from NCBI. We obtained the ExAC (Exome Aggregation Consortium) VCF file, which contains all variants and variant annotations in ExAC, from the Broad Institute. We obtained the ClinVar VCF file, which contains ClinVar variants annotated with their pathogenicity, from NCBI. All VCF files were reduced to *BRCA* regions, annotated, and standardized using CAVA. Each annotated VCF file from the Harvard PGP or the 1000 Genomes Project was compared with the ExAC and ClinVar VCF files using bcftools isec. To find variants seen in the annotated VCF file and one of ExAC or ClinVar, parameters –nfiles=2 were used. To find variants seen in the annotated VCF file, ExAC, and ClinVar, parameters –nfiles=3 were used. To find variants seen only in the annotated VCF file, the –complement flag was used.