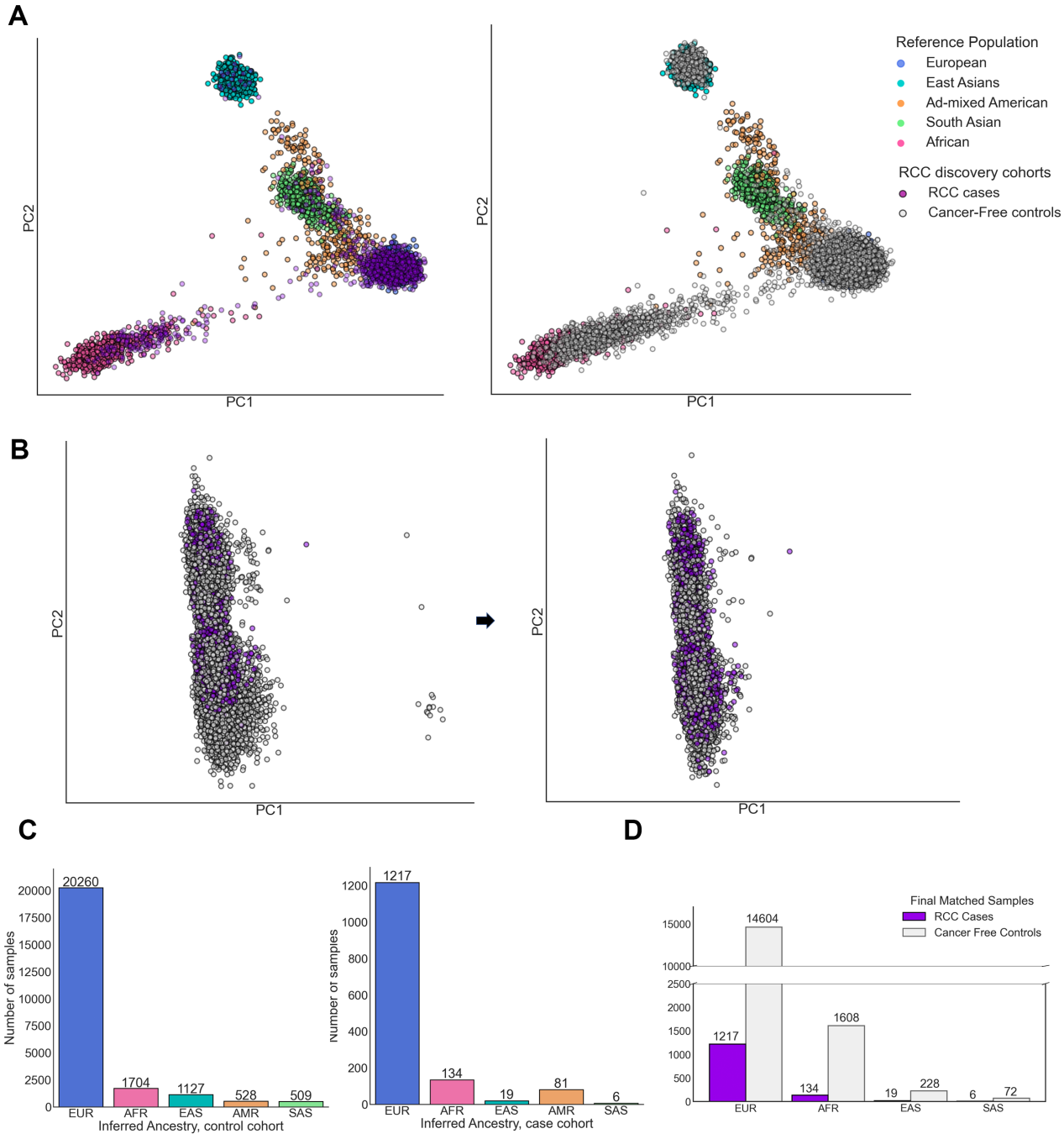
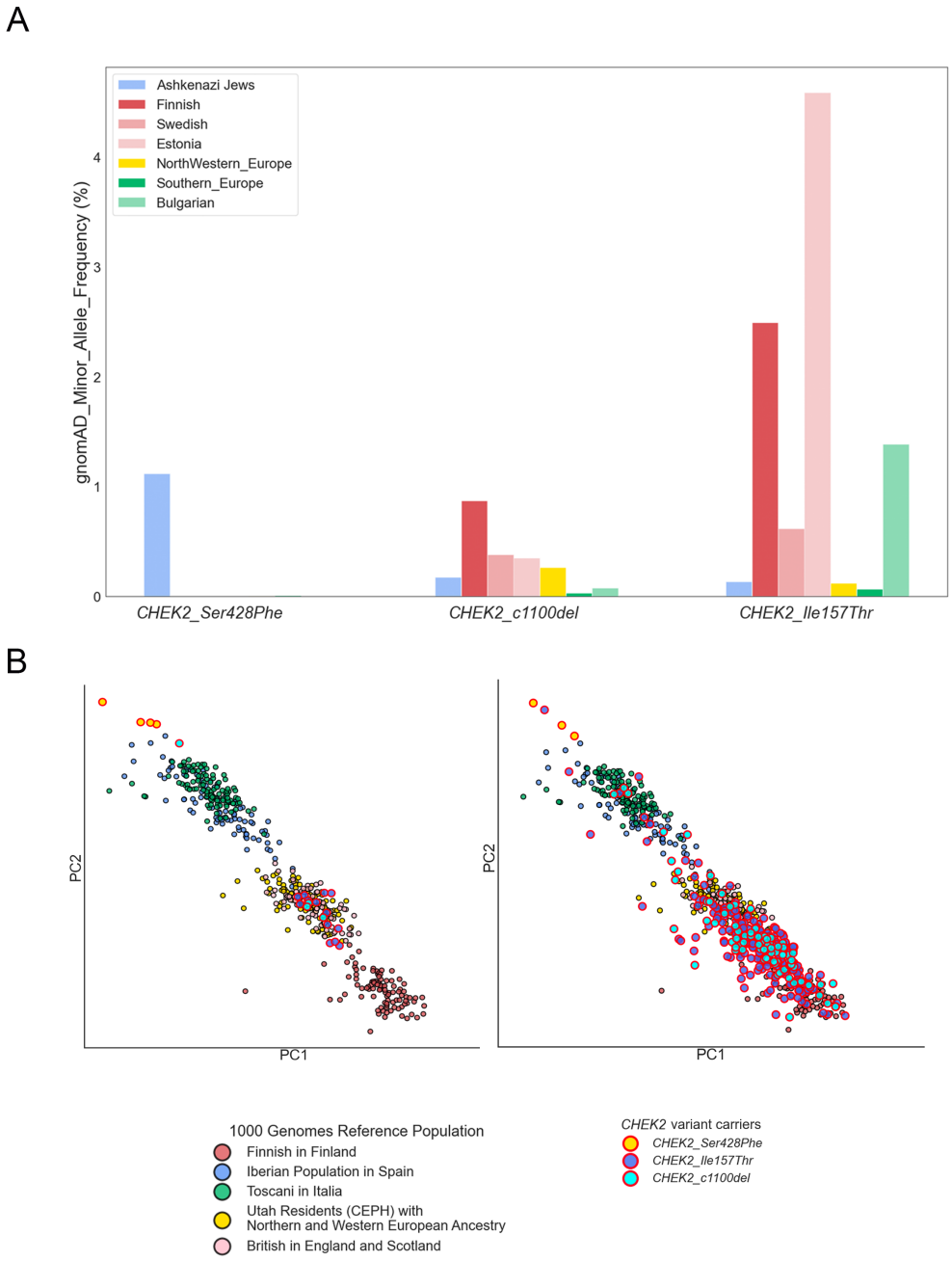


Supplementary Figure 1. Genetic ancestry inference and ancestry matching of cases and controls All available RCC patients were ancestry-matched with cancer-free controls, and patients with RCC subtypes other than clear cell, papillary, or chromophobe RCC were excluded from the final matched set of cases and controls.
(A) Projection of cases and controls on the first two principal components (PCs) along with five broad continental reference clusters formed using the samples of the 1000 Genomes project.
(B) Ancestry pair-matching of the cases and controls identified as European in the inference.

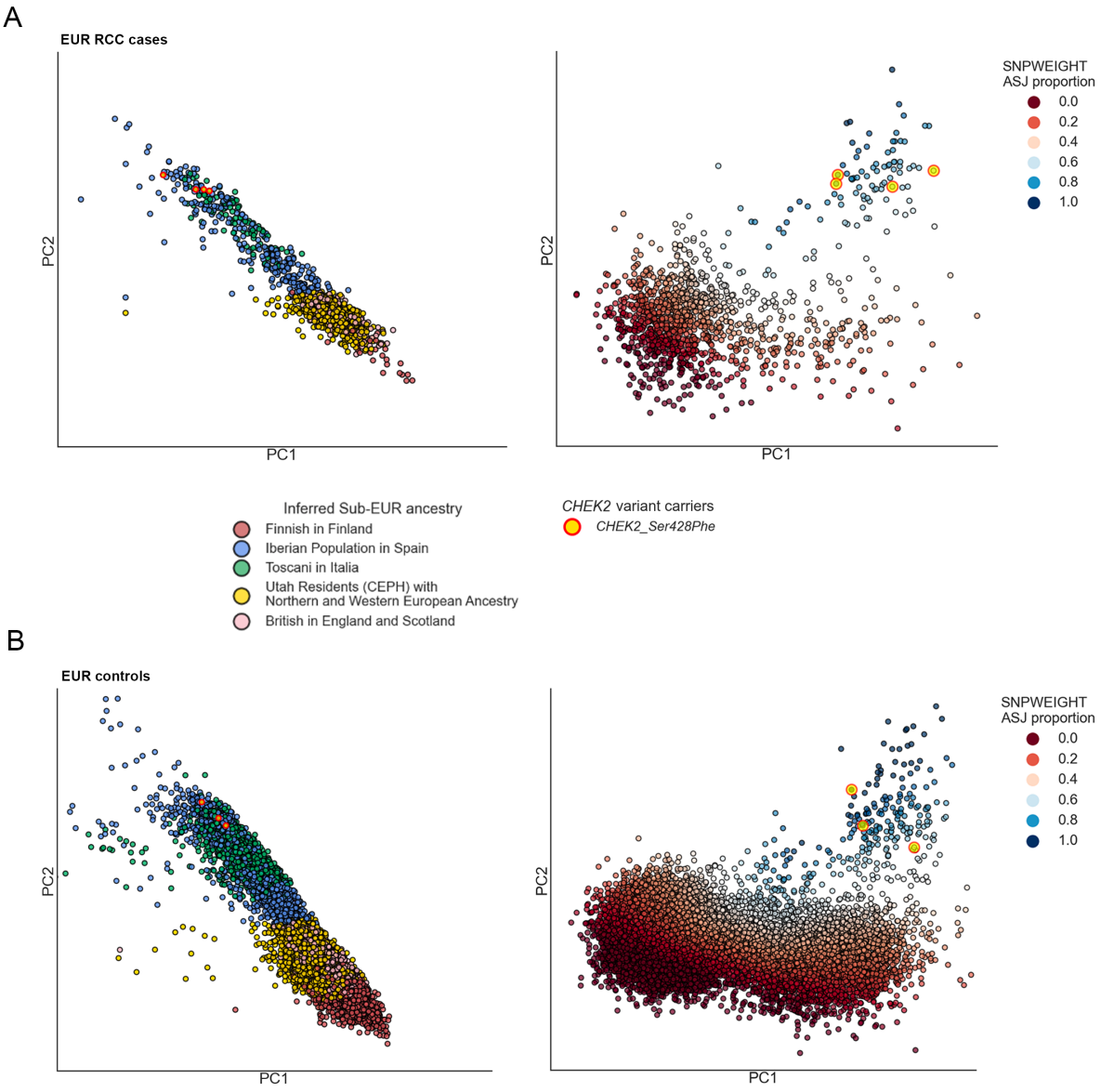
Control samples closest to each case were identified using the top 10 PCs and the rest was excluded
(C) Continental ancestry assignment of cases and controls. Each sample was assigned to one of the five continental ancestries (EUR: European, AFR: African, EAS: East Asian, AMR: Admixed American, SAS: South Asian) after the first round of PCA and random forest. (D) Cases and controls were matched with a 1:12 ratio in each ancestry group. To secure the largest ratio possible between cases and controls, AMR cases and controls were excluded



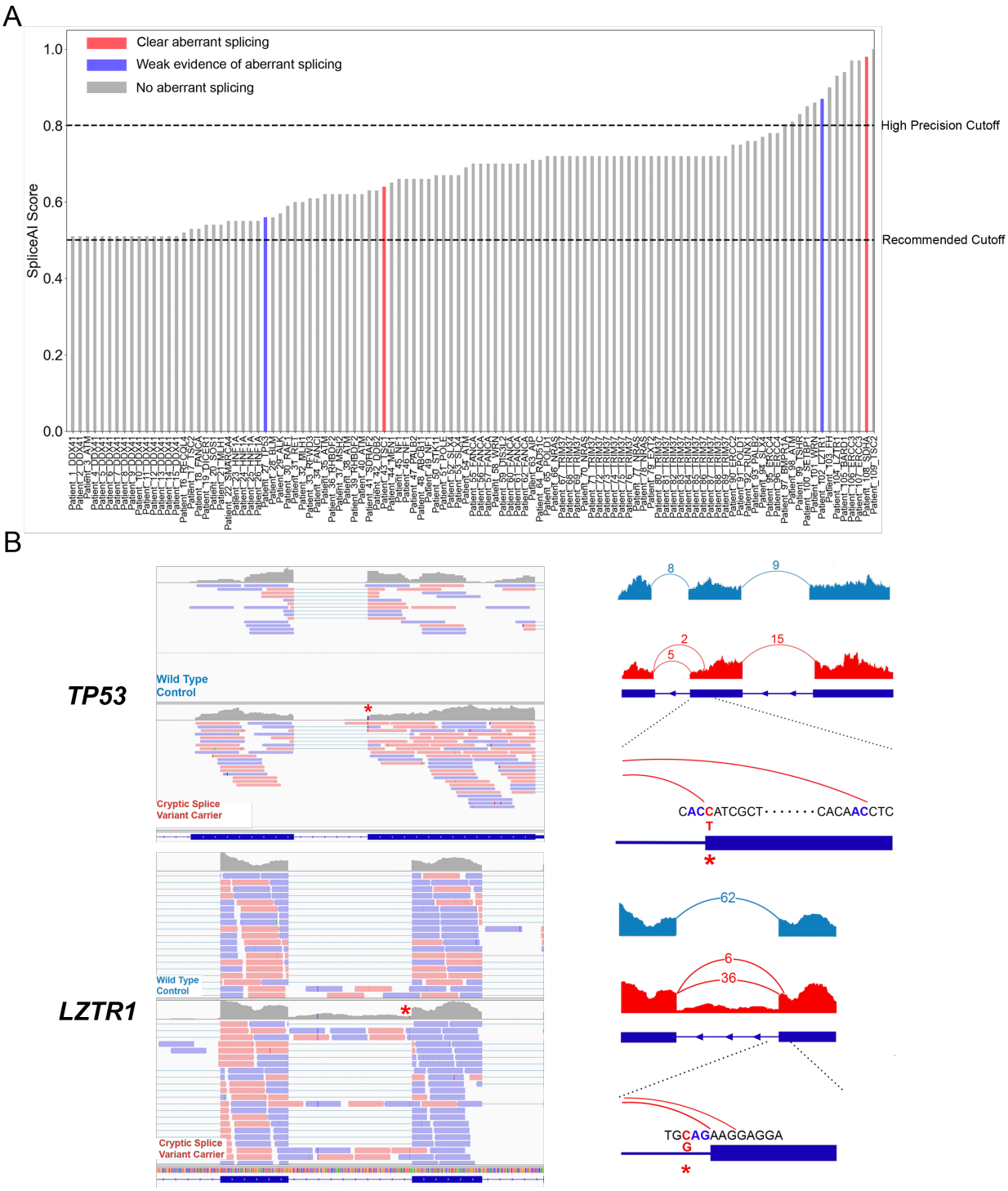
Supplementary Figure 2. Substantial variation in population allele frequencies of *CHEK2* founder variants in different European populations

A) gnomAD minor allele frequency in % for the three *CHEK2* founder variants indicating the wide range of population frequency in different European groups

B) Distribution of the three *CHEK2* founder variants in RCC cases (Left) and in cancer-free controls (Right). Reference sub-European clusters were formed using 1000 Genomes European samples with corresponding sub-European labels.



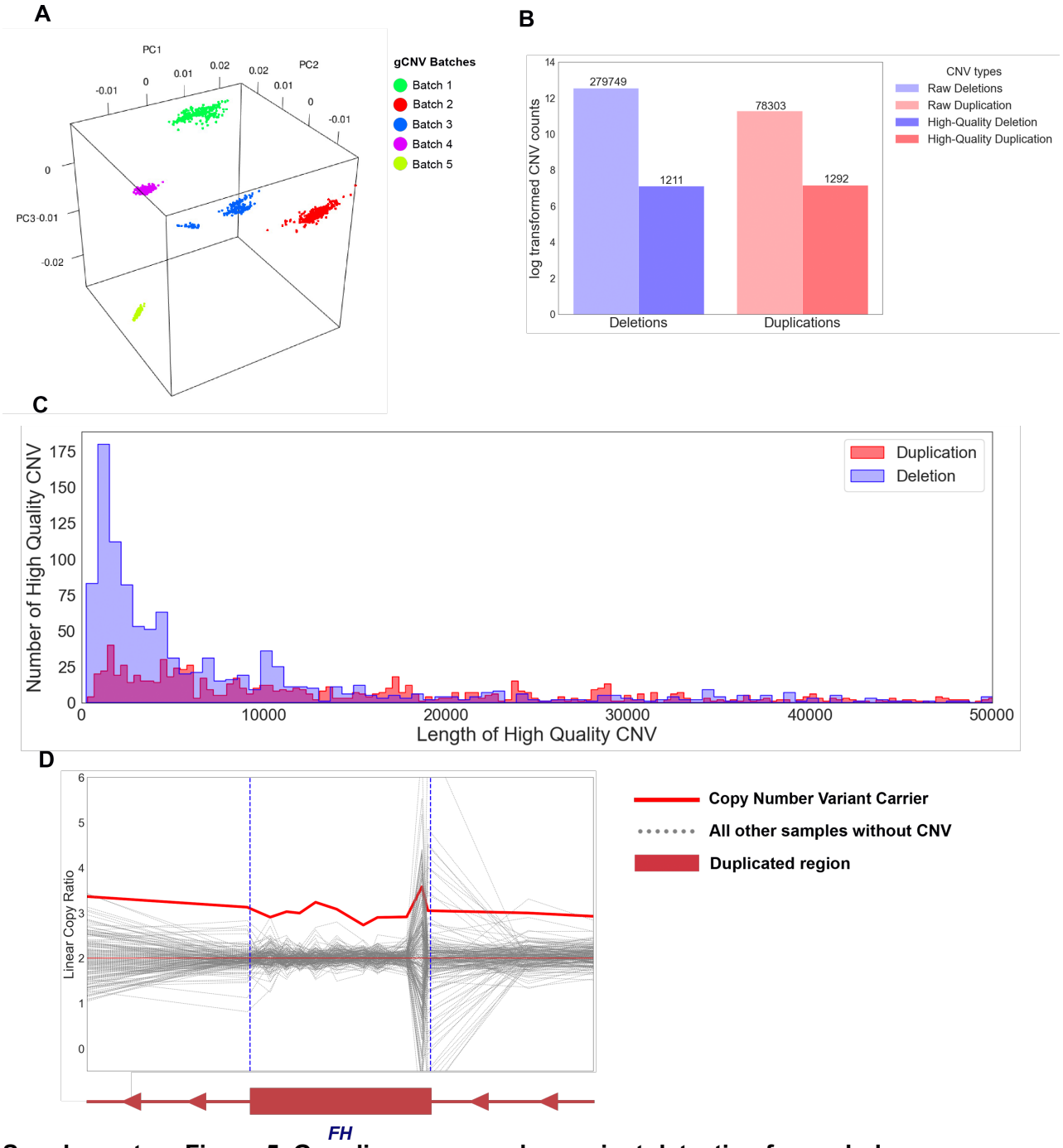
Supplementary Figure 3. SNPWEIGHSTs sub-European ancestry inference
Ashkenazi Jewish individuals, which were unable to be identified using PCA and random forest with 1000 Genomes were identified using SNPweights.
A) Left - Projection of the first two PCs for all European RCC cases color coded with inferred sub-EUR ancestry. All 4 carriers of *CHEK2* p.Ser428Phe clustered with Southern European individuals. Right - Projection of the two PCs from SNPweights with color coding indicating the level of inferred Ashkenazi Jewish proportion. All *CHEK2* p.Ser428Phe carriers fell within a cluster with a high (>0.6) ASJ proportion
B) Same as A for the European controls.



Supplementary Figure 4. Validation of cryptic splice variants identified in the TCGA and CHECKMATE samples

A) mRNA sequencing data of 89 TCGA and CHECKMATE samples with identified cryptic splice variants were evaluated to confirm the effect of the variants. Most of the putative cryptic splice variants didn't show evidence of aberrant splicing

B) A cryptic splice variant in *TP53* and another in *LZTR1* shows aberrant splicing induced by donor loss and acceptor loss respectively, but the number of junction split reads were too low to confidently support the presence of aberrant splicing.



Supplementary Figure 5. Germline copy number variant detection from whole exome sequencing data

A) PCA on sequencing baits formed 5 different clusters to run GATK-gCNV

B) Bar plots indicate the number of raw deletions and duplications as well as high-quality CNVs after stringent filtering steps

C) Histogram summarizing the length of high-quality duplications and deletions

D) The large duplication of 56kbp covering 7 different genes including *FH*