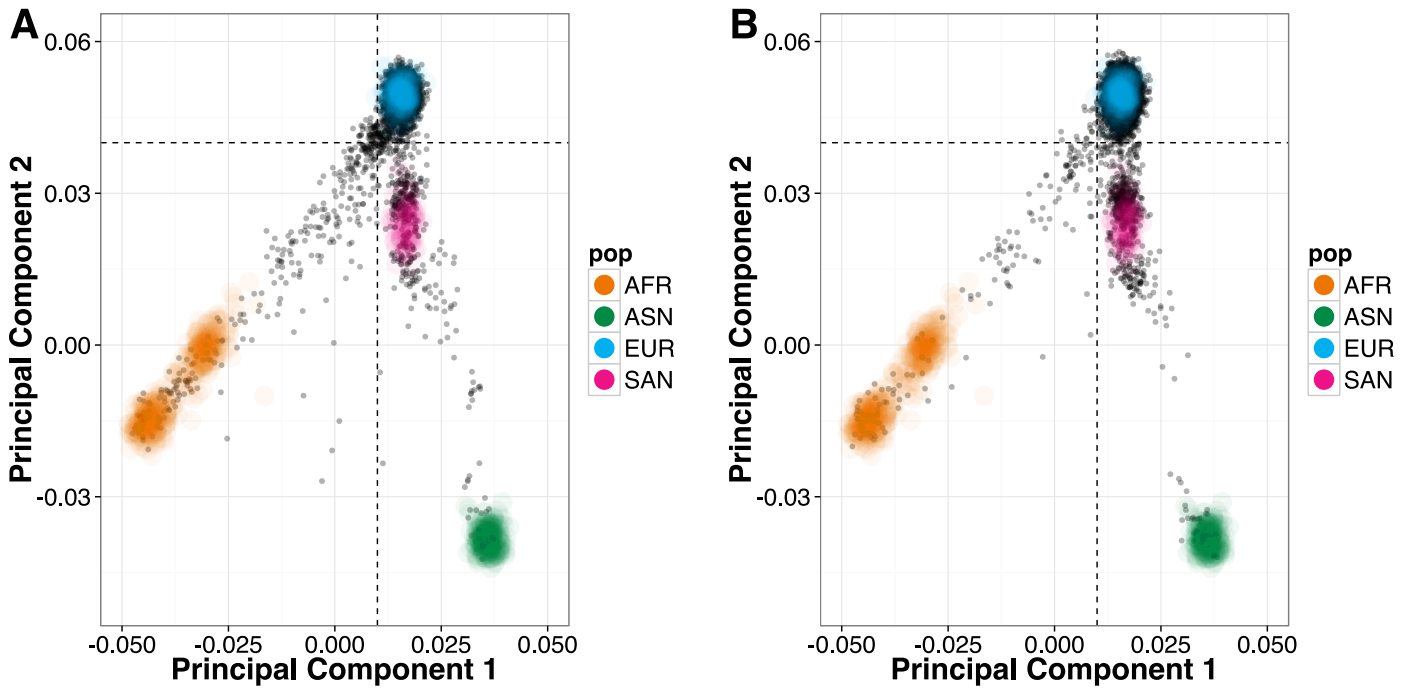


Supplementary Figure 1

Principal component analysis based on a selected set of polymorphic SNPs.

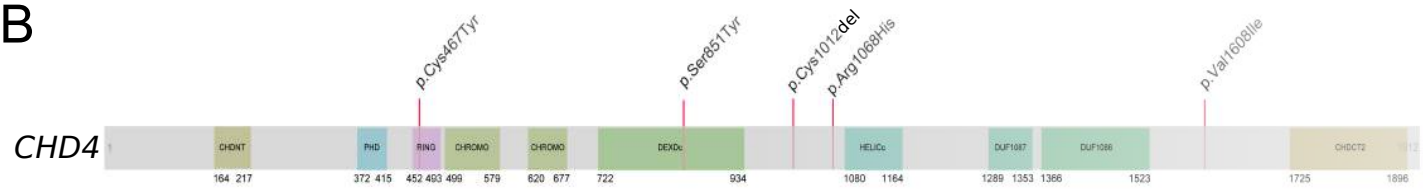
Transparent colored areas represent reference samples of different Hapmap3 populations (African, Asian, European, South-East Asian) as an indicator of ethnic ancestry. Black dots represent samples in the A) cases or B) controls. Dashed lines represent the thresholds used to classify a sample as being of European ancestry (PC1 >= 0.01 and PC2 >= 0.04)

A

	Decipher ID	267060	264040	269294	267459	259179
	Mutation	Cys1012del	Ser851Tyr	Arg1068His	Val1608Ile	Cys467Tyr
Affected Phenotype categories	Cardiac morphology					
	Head and neck					
	Nervous system					
	Genitourinary system					
	Developmental milestones					
	Prenatal or Perinatal development					
	Thoracic cavity					

Counts
5
4
2

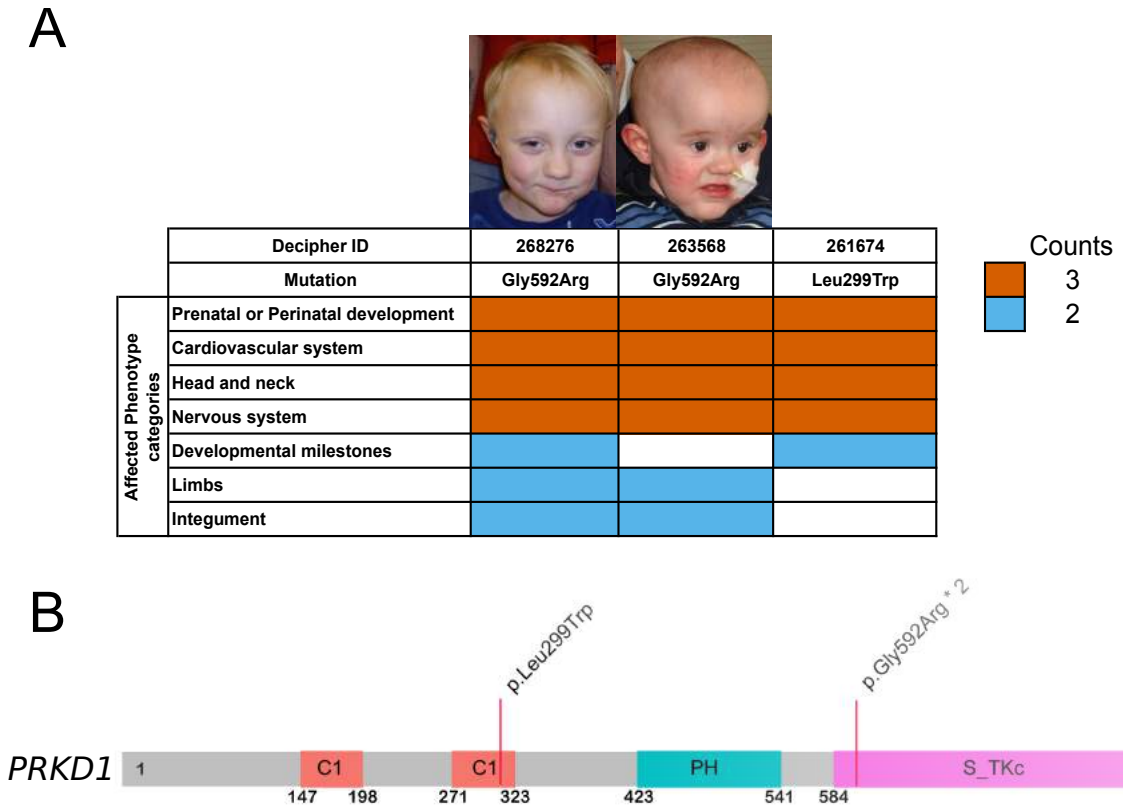
B



Supplementary Figure 2

Probands with *de novo* CHD4 mutations.

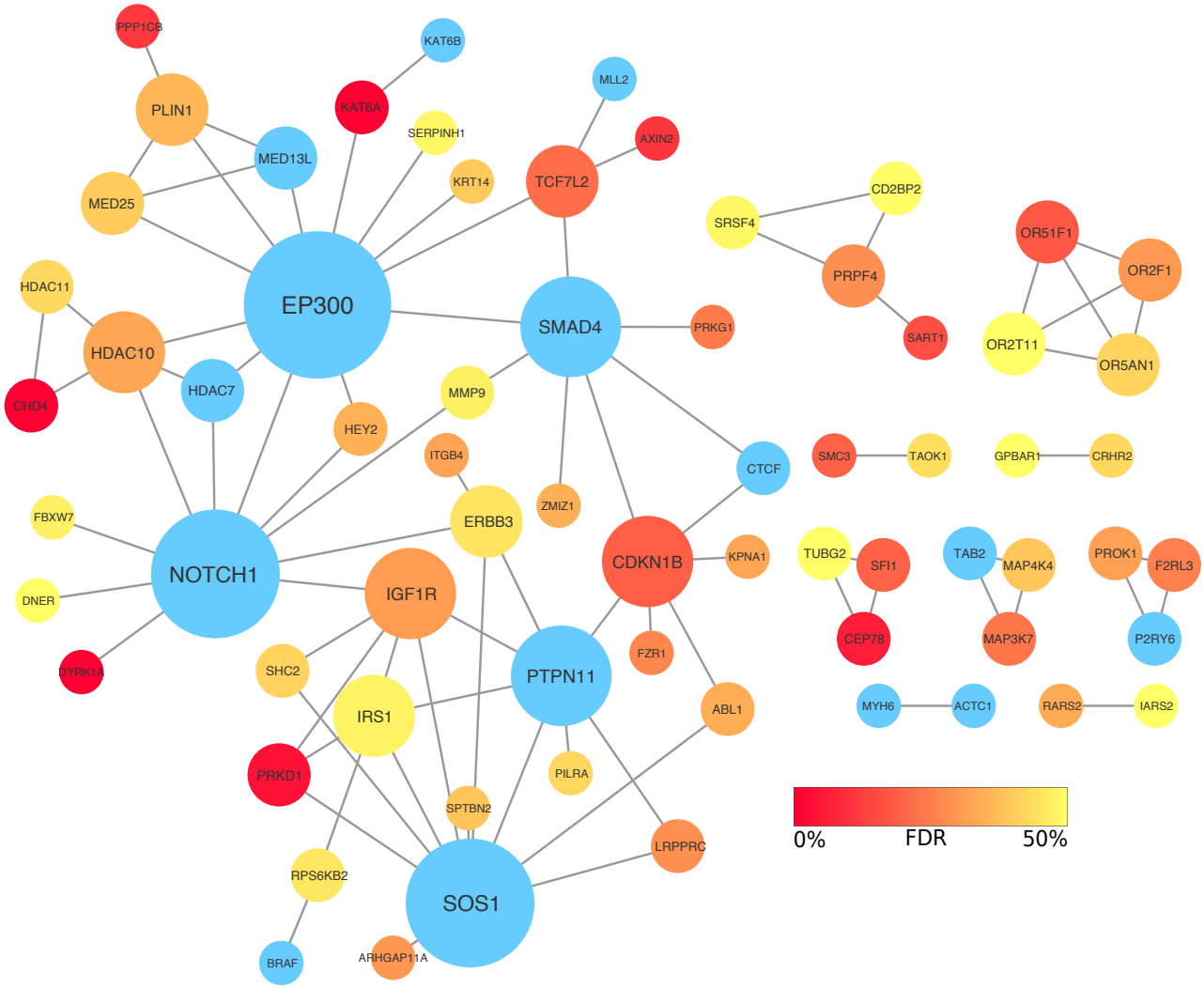
A) Clinical synopsis of the observed phenotypes across patients carrying *CHD4* mutations. Columns represent single probands, shades of cells represent the number of probands sharing a phenotype in the given phenotypic categories. Photographs of affected probands are shown for which consent could be obtained for publication. B) Protein plot showing *CHD4* protein domains and the distribution of *de novo* mutations.



Supplementary Figure 3

Probands with *de novo PRKD1* mutations.

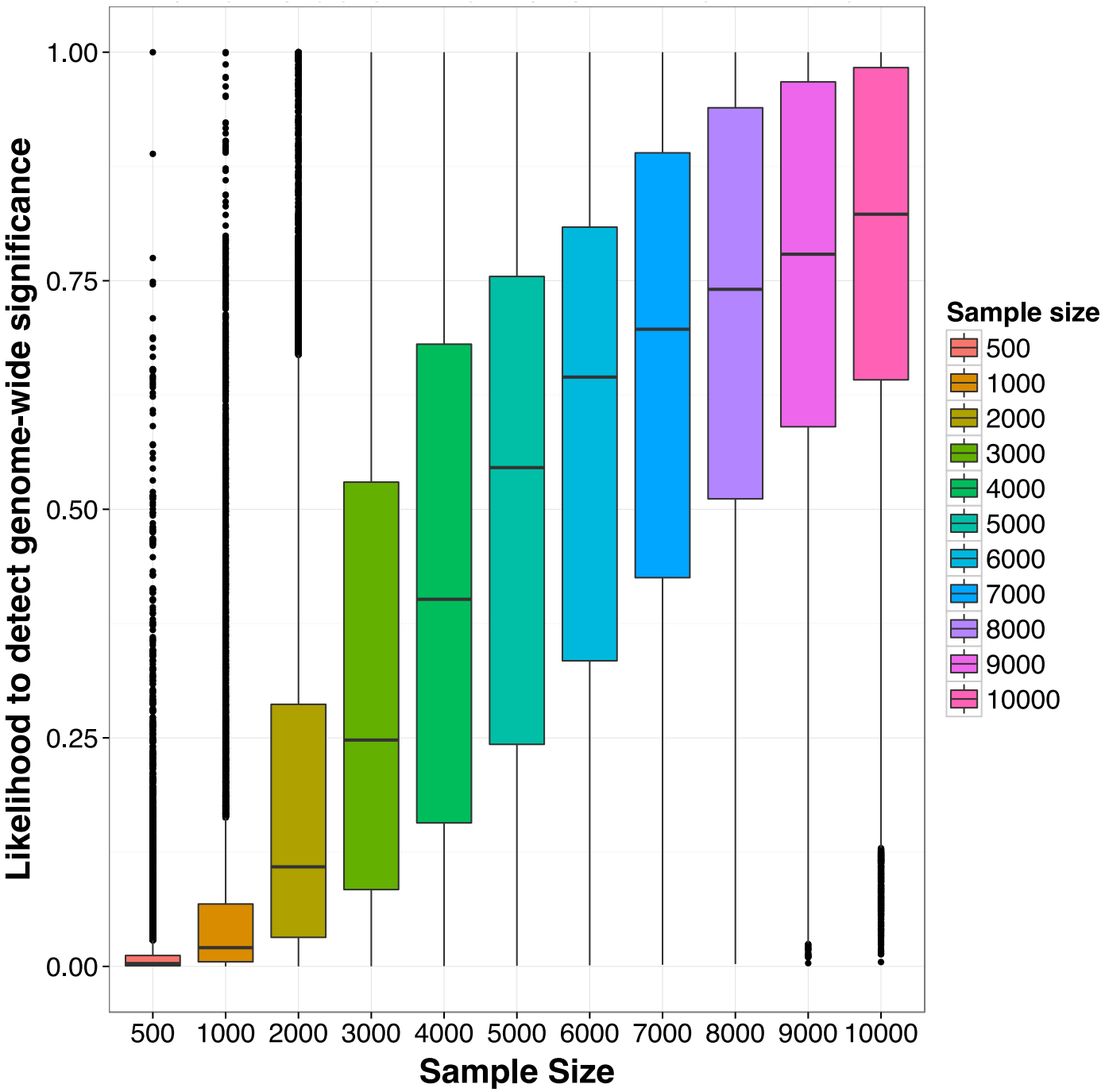
A) Clinical synopsis of the observed phenotypes across patients carrying *PRKD1* mutations. Columns represent single probands, shades of cells represent the number of probands sharing a phenotype in the given phenotypic categories. Photographs of affected probands are shown for which consent could be obtained for publication. B) Protein plot showing *PRKD1* protein domains and the distribution of *de novo* mutations. Two probands share identical missense mutations (Gly592Arg).



Supplementary Figure 4

High confidence interactions found in the STRING PPI database in top-ranking set of CHD-associated genes.

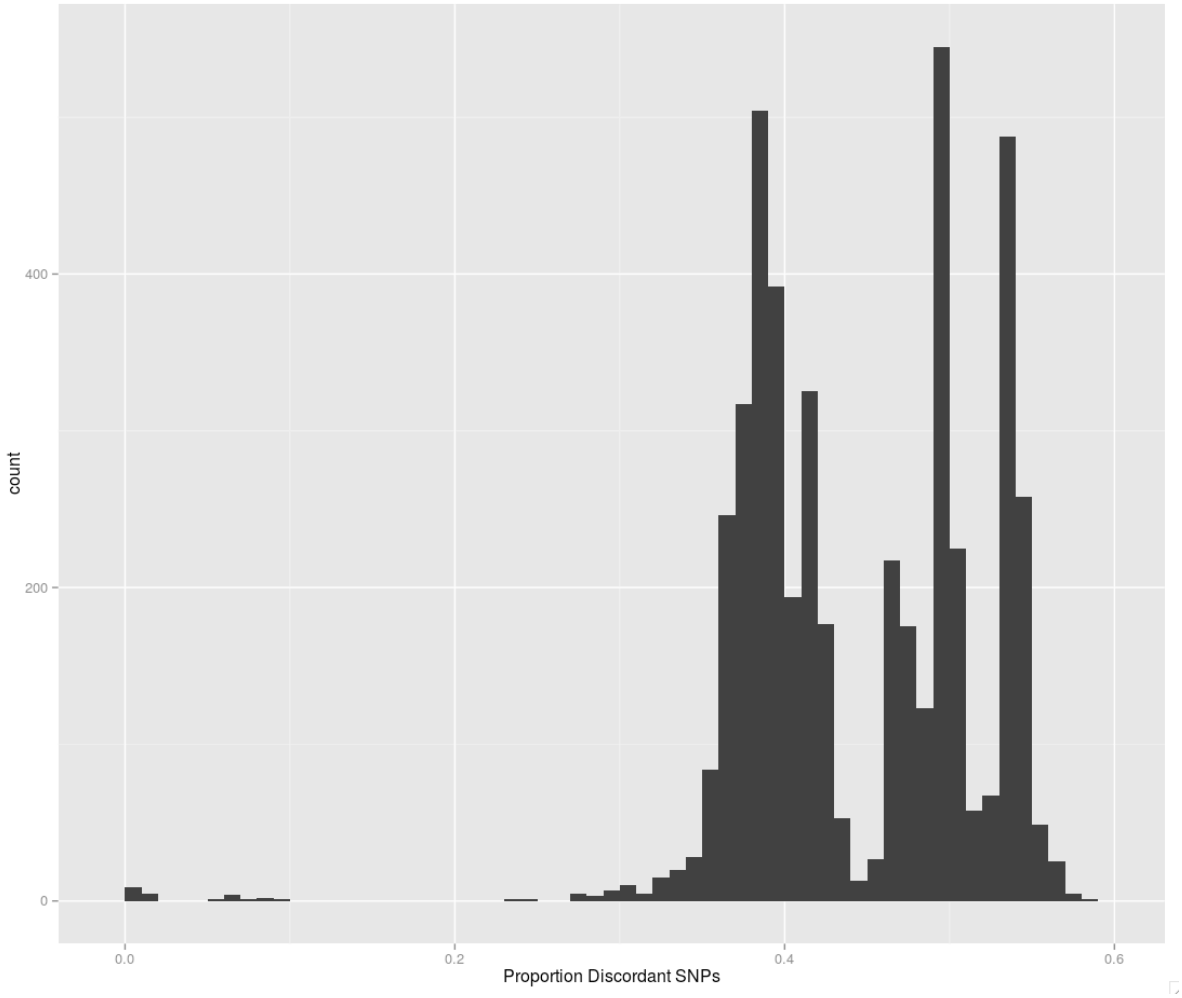
Edges in the graph represent high confidence (STRING score > 0.9) protein-protein interactions between top ranking genes (FDR < 50%) in the integrated *de novo* and inherited variant analysis. Known CHD-associated genes are colored blue, other nodes are scaled according to their association FDR. The size of nodes is scaled by its degree.



Supplementary Figure 5

Saturation analysis for detecting haploinsufficient S-CHD-associated genes.

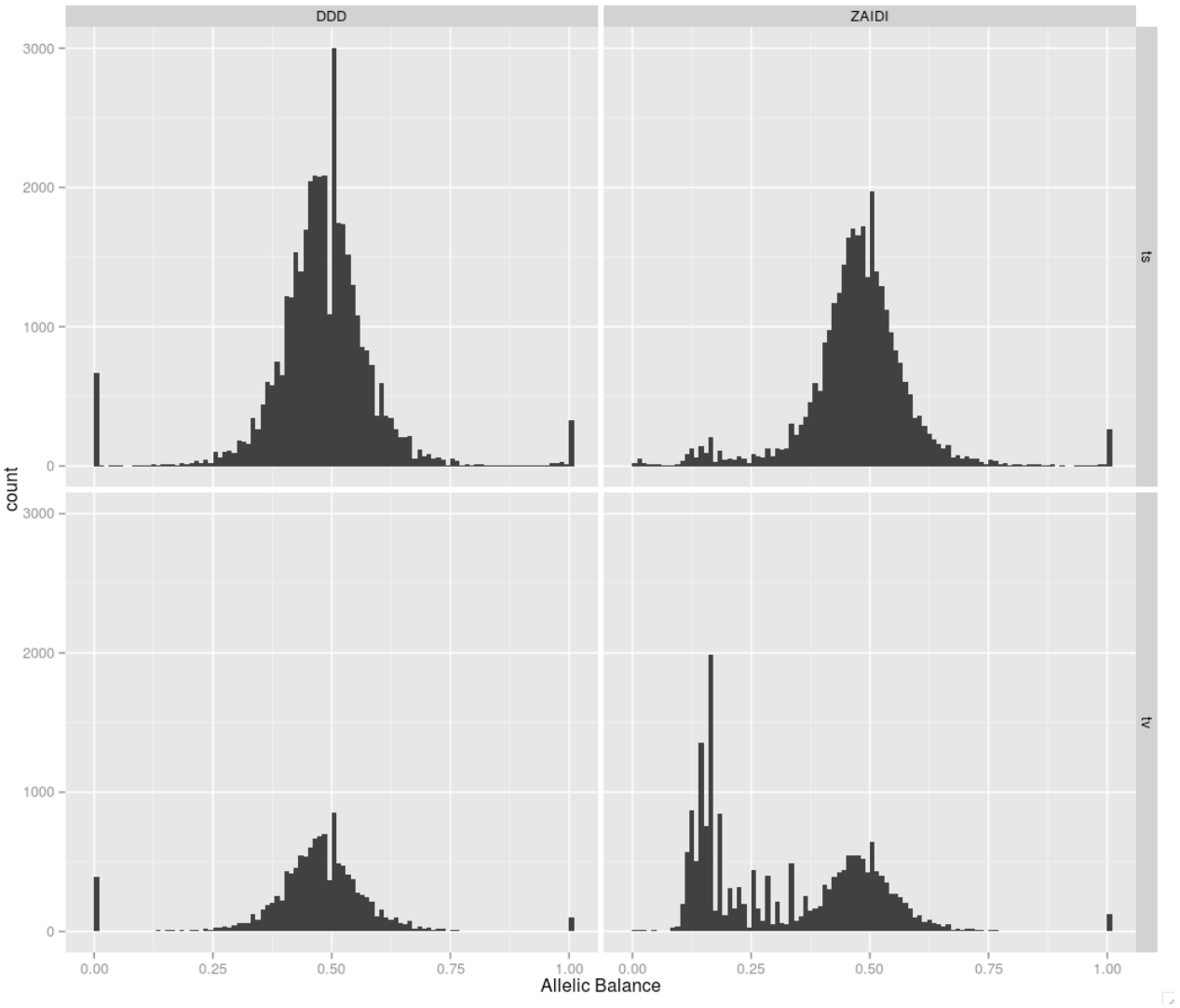
A box plot showing the distribution of statistical power to detect a significant enrichment of PTV mutations across 19252 genes in the genome, for different numbers of trios studied, from 500 trios to 10,000 trios. Line within the box shows the median, box corresponds to the first and third quartiles (the 25th and 75th percentiles) and whiskers correspond to most extreme values within 1.5 times the interquartile range from the box.



Supplementary Figure 6

Distribution of pairwise-discordance between any sample pair in the cohort.

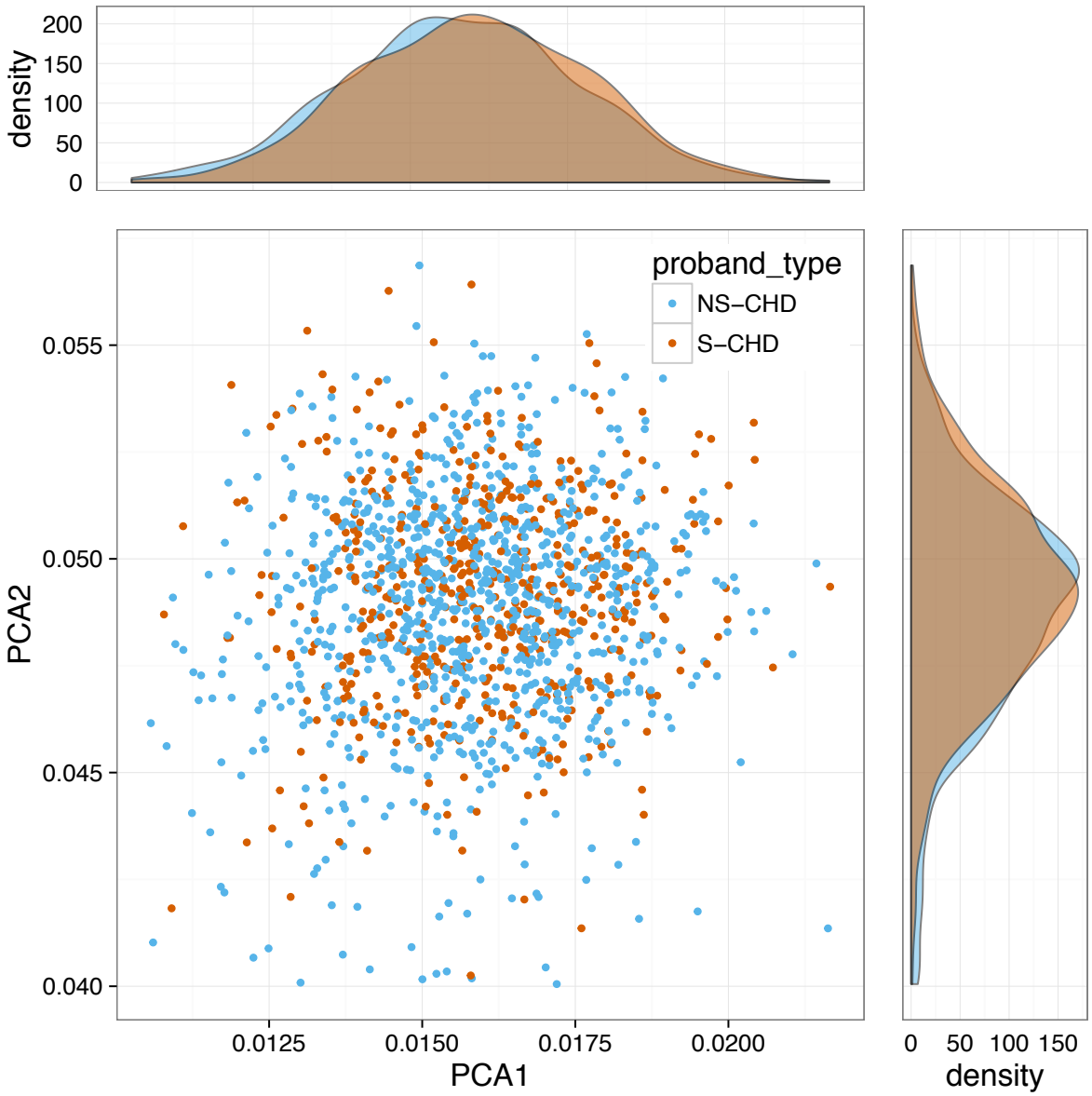
Samples with the lowest median coverage in sample-pairs with a discordance below 20% were excluded from further analysis.



Supplementary Figure 7

Comparison of candidate *de novo* mutations called by DeNovoGear between publicly available datasets of the DDD study in and Zaidi et al. study.

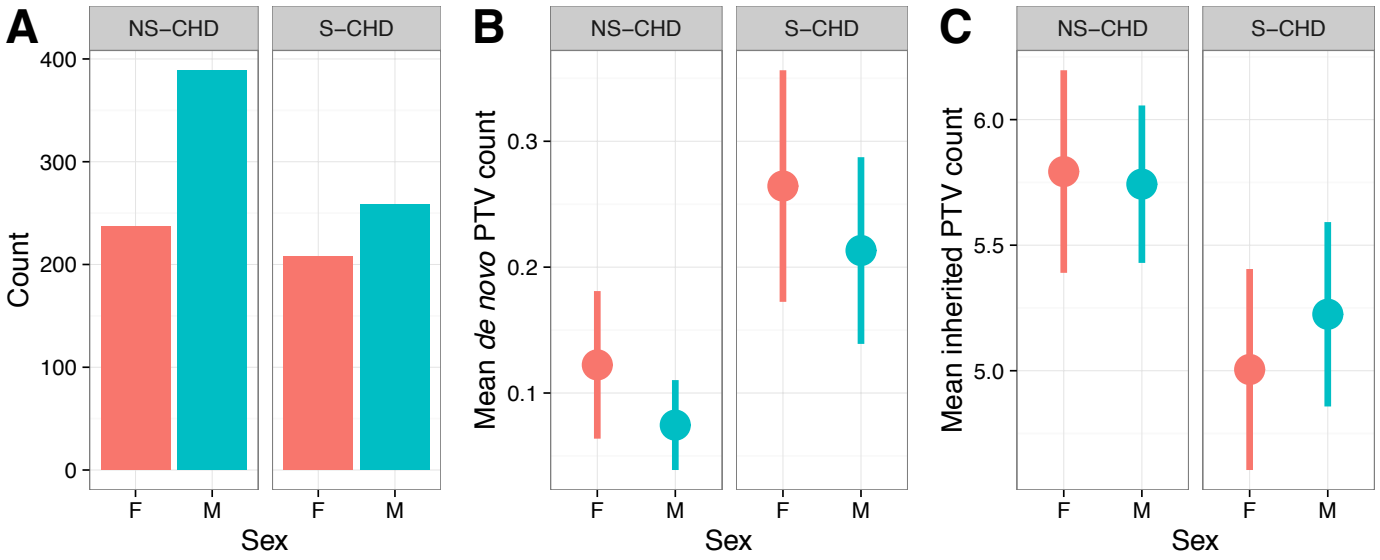
Allelic balance distributions are shown for transitions (ts) and transversions (tv) in both datasets. An overrepresentation of transversion events is found in the Zaidi et al. dataset at low allelic balances, primarily made up of G>T transversions.



Supplementary Figure 8

Principal component analysis of common SNPs for syndromic CHD probands and non-syndromic CHD probands.

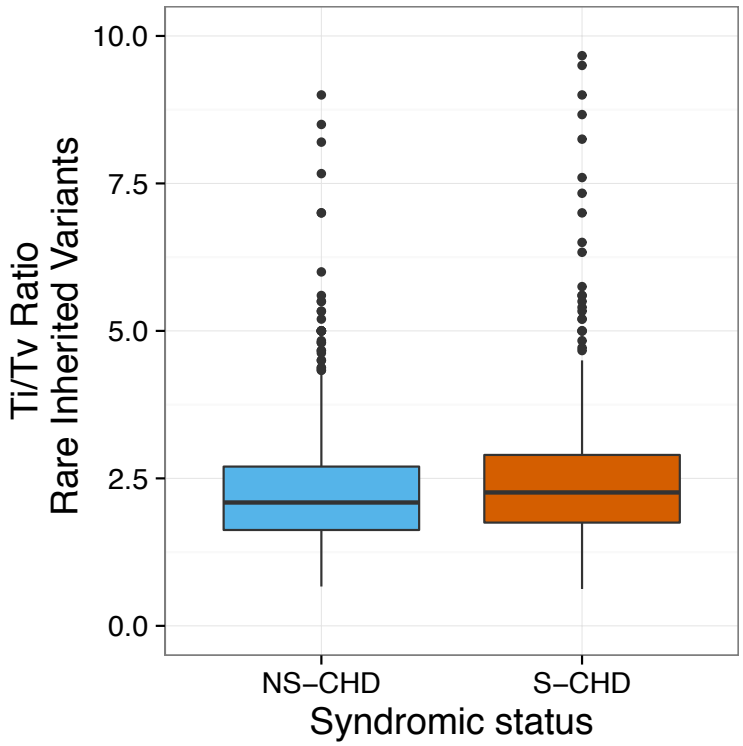
First two components of PCA analysis, blue dots represent non-syndromic CHD cases while orange dots represent syndromic CHD cases. Sideplots represent the marginal distributions of the principal components, stratified by syndromic status.



Supplementary Figure 9

Number of cases and variants by sex stratified by syndromic status.

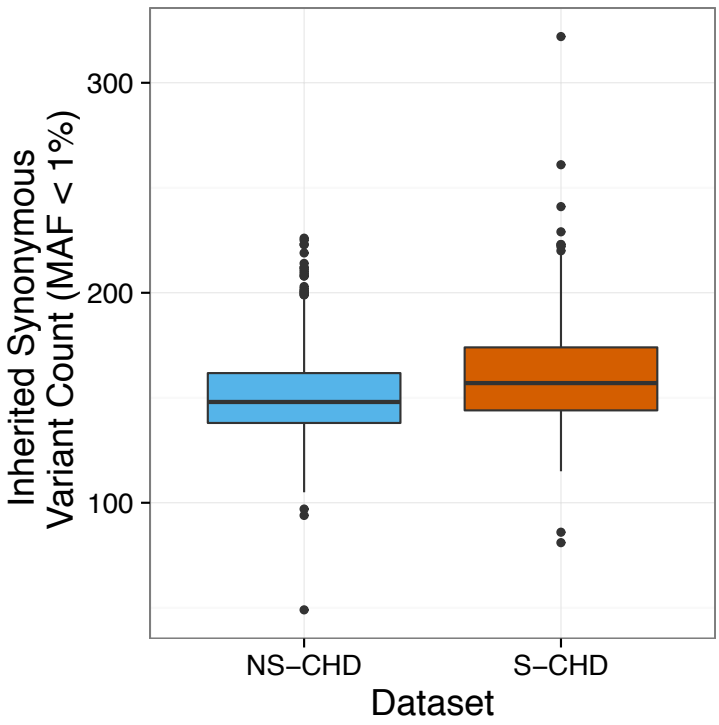
A) Counts by sex and syndromic status of samples included in the burden analysis. B) Mean number of B) *de novo* and C) inherited PTV variants stratified by sex and syndromic status. Error bars represent 99% confidence intervals on the measured means.



Supplementary Figure 10

Boxplots for Transition/Transversion (Ti/Tv) ratios of rare exomic inherited variants for samples used in the burden analysis.

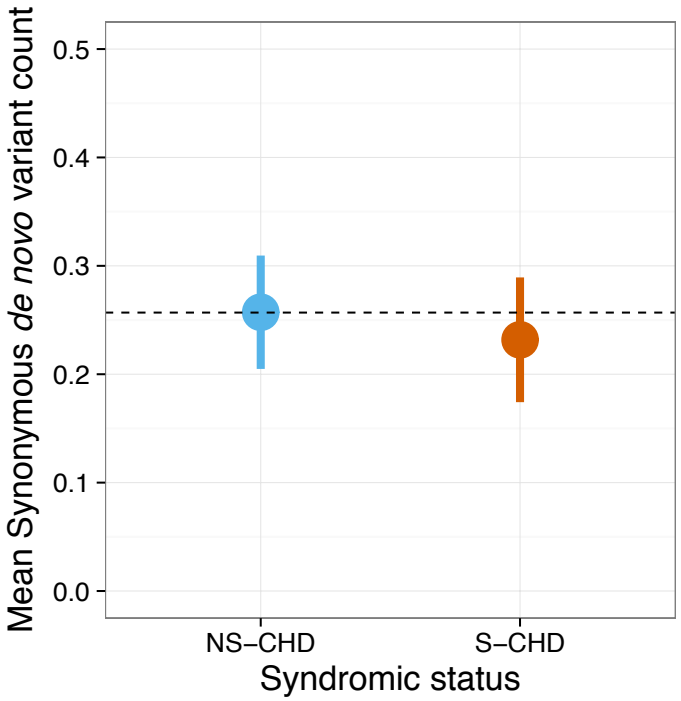
Samples are stratified by syndromic status. Whiskers represent the highest/lowest value within 1.5 times the interquartile range. Upper and lower hinges represent the 25th and 75th percentiles, respectively.



Supplementary Figure 11

Boxplots showing counts of rare inherited synonymous variants for samples used in the burden analysis.

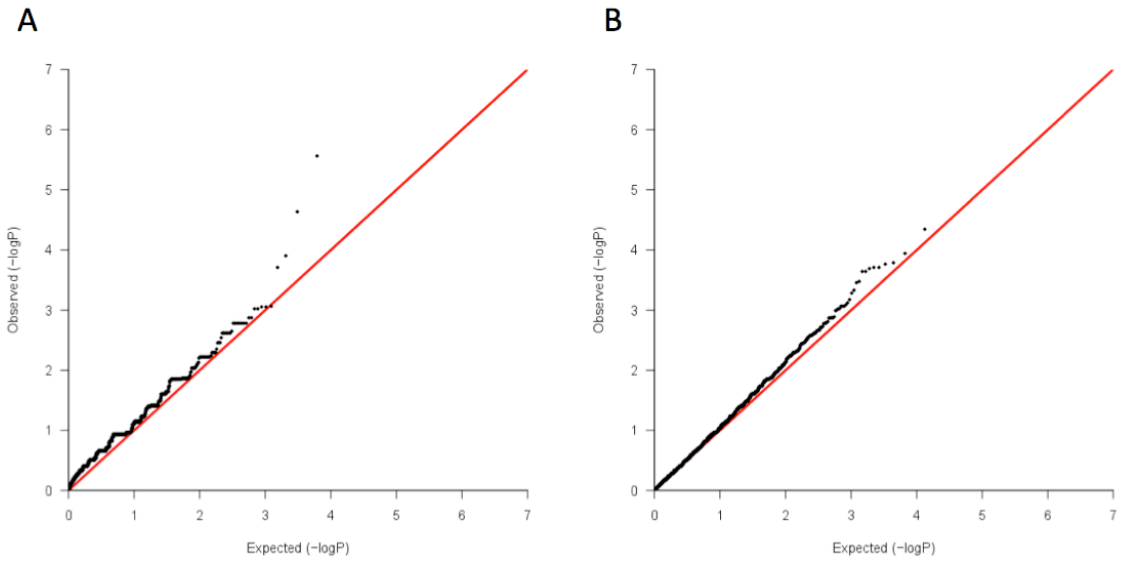
Samples are stratified by syndromic status. Whiskers represent the highest/lowest value within 1.5 times the interquartile range. Upper and lower hinges represent the 25th and 75th percentiles, respectively.



Supplementary Figure 12

Mean number of autosomal *de novo* synonymous mutations stratified by syndromic status.

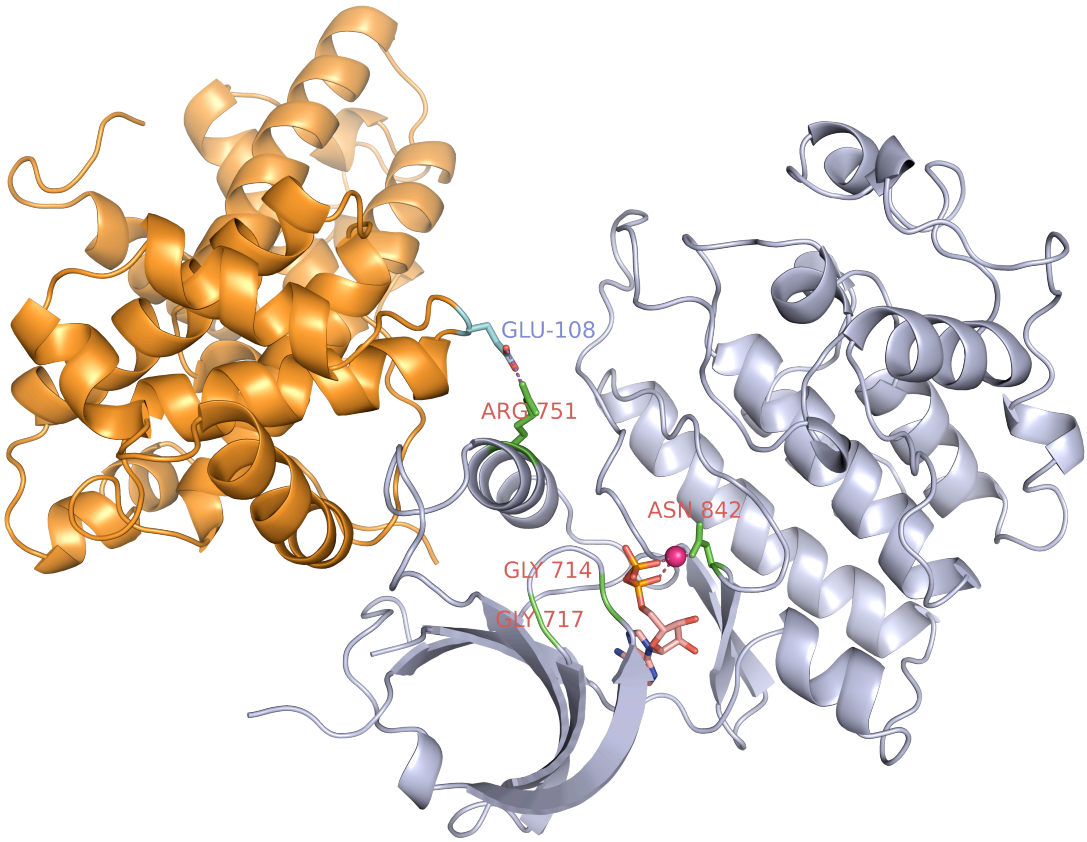
Dashed line represents the expected synonymous mutation rate (Samocha et al., 2014). Error bars represent the 99% confidence interval for the measured mean.



Supplementary Figure 13

Quantile-quantile plots for case/control counts using Fisher's exact test statistics.

QQ-plots considering case/control A) PTV counts or B) missense counts using Fisher's exact test statistics to test for systematic deviations from the null hypothesis.



Supplementary Figure 14

***CDK13* protein structure by homology modeling.**

CDK13 protein structure by homology modeling on the known *CDK12* structure with co-crystallized ATP-substitute AMP ligand: mutated residues are shown in green, catalyzing Magnesium ion is shown in magenta, Cyclin domain is shown in orange with it's interacting residue shown in cyan.