Supplemental Material

SUPPLEMENTAL METHODS

In order to implement ACMG-AMP criteria we automatically extracted relevant information from public available resources, such as MedGen, ClinVar and ExAC.

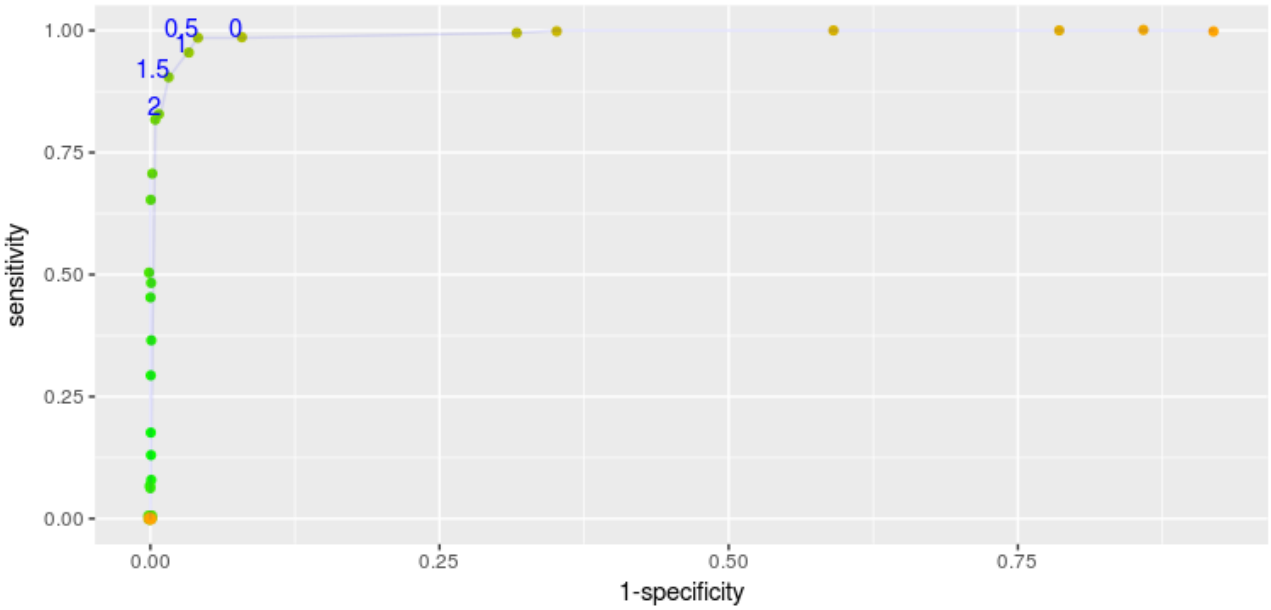
MedGen

The aim of ACMG-AMP guidelines is to interpreted variants associated with a Mendelian disorder. MedGen (https://www.ncbi.nlm.nih.gov/medgen) organizes information related to human genetic conditions into an ontology: a MedGen term, representing a human phenotype, could be related to another phenotype in a "is-a" relationship. For example, Brugada syndrome (MedGen id: C1142166) is a Cardiac arrhythmia (MedGen id: C0003811), and eight more specific conditions are annotated as Brugada syndrome's children. From MedGen files available at ftp://ftp.ncbi.nlm.nih.gov/pub/medgen/ we collected all genes-phenotypes association, together with their inheritance pattern. Therefore, we are able to assign to each variant a precise list of CVDs phenotypes associated with the gene in which the variant occurs.

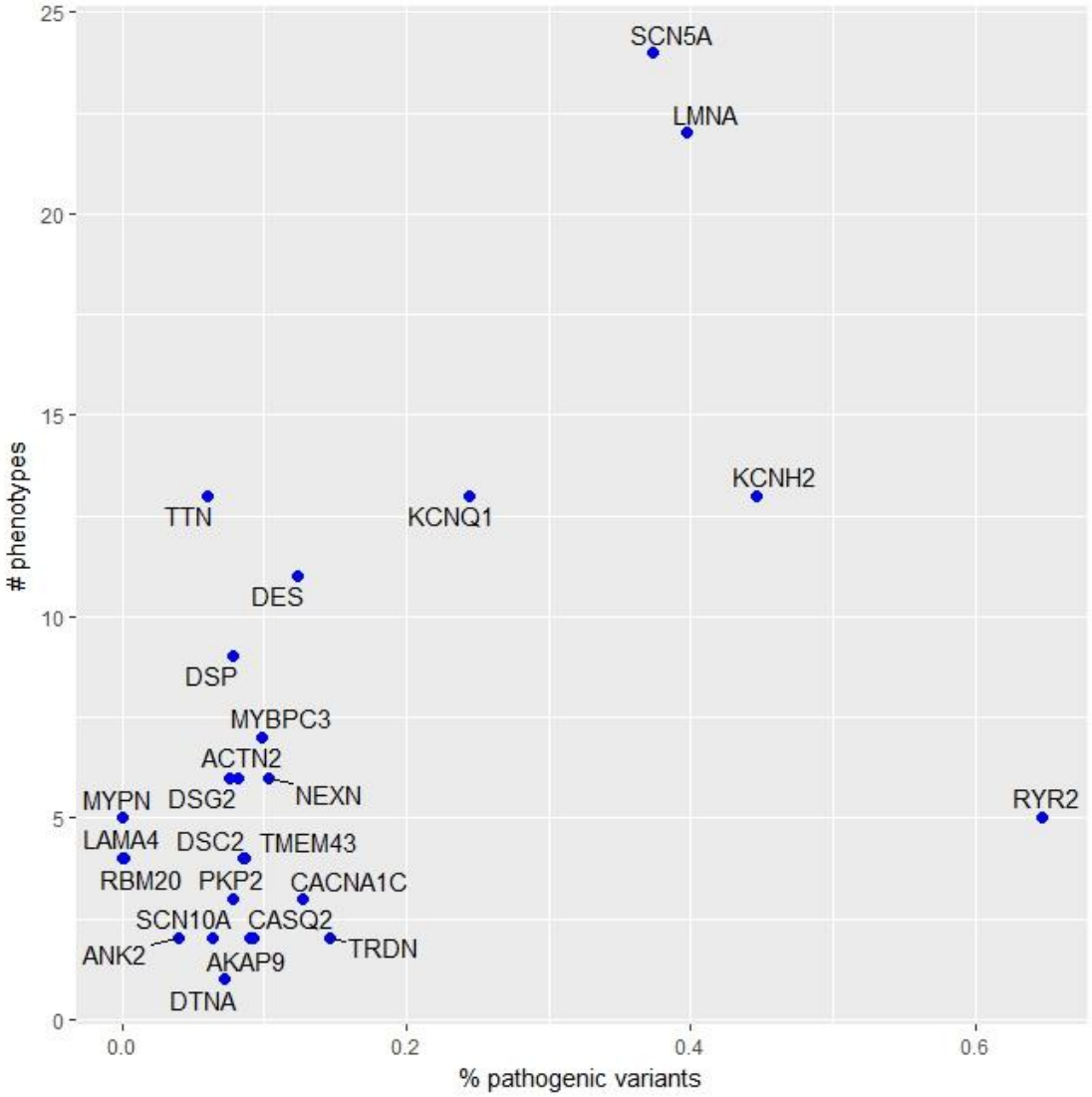
ClinVar

Several ACMG-AMP criteria rely on previous interpretation of sequence variants. For example, PS3 and PP5 criteria, or their benign equivalent BS3 and BP6, explicitly consider a prior classification of the same nucleotide change respect to a Mendelian phenotype, while PS1 concerns variants with the same amino acid change. Bona-fide ClinVar variants could also be helpful for the determination of mutational hotspot (PM1) or to determine the rate of benign missense variation (PP2). We collected such information from ClinVar, a repository of human variants whose clinical significance has been assessed by different laboratories. The entire dataset is provided in a XML form (ftp://ftp.ncbi.nlm.nih.gov/pub/clinvar/xml/). We have developed a pipeline that extracts from ClinVar XML only relevant information for the classification process. ClinVar submitters usually specify their clinical interpretation of a variant-phenotype association. Phenotypes could be identified by ids of different resources, such as MedGen, OMIM or Orphanet. Since we identify phenotypes with MedGen terms, we mapped non-MedGen terms to MedGen terms through Disease Ontology (Kibbe et al., 2015). Conflicting interpretation among submissions are not included in knowledge base. According to our method, a variant has conflicting interpretation if it is labeled as "Benign/Likely benign" and "Pathogenic/Likely pathogenic" by different submitters, even in case they were assessing the variant for different phenotypes.

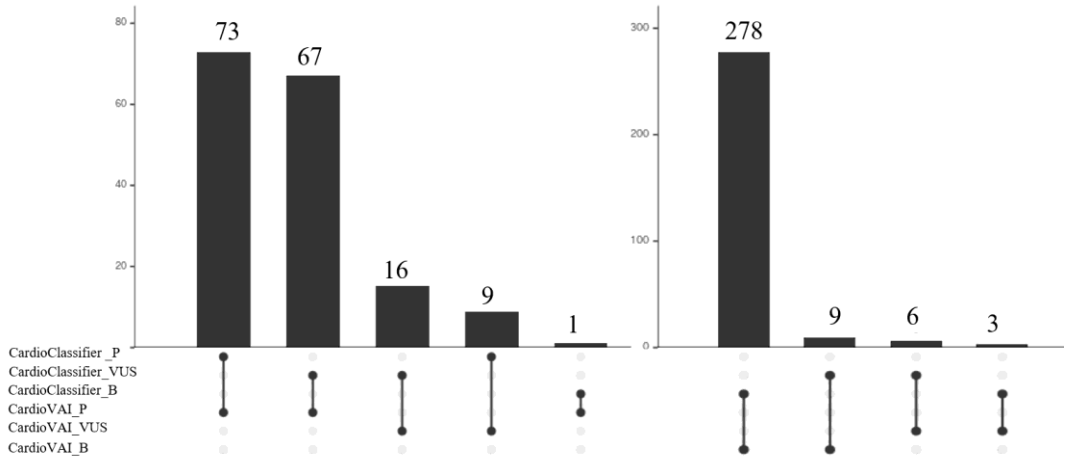
SUPPLEMENTAL FIGURES AND FIGURE LEGENDS



Supp. Figure S1: ROC analysis on 885 CLINVITAE benchmark variants (685 benign/likely benign and 200 pathogenic/likely pathogenic). Since variant-phenotype associations may have different pathogenicity scores with respect to triggered criteria, we assigned to each variant the median of the pathogenicity scores of its variant-phenotype associations. We evaluated the performance of the pathogenicity score as a classification threshold: if a variant has a median pathogenicity score greater than a certain threshold, then it is considered as pathogenic, conversely, it is classified as benign. ranging from -3 to 16.5, We have calculated and reported specificity and sensitivity for different thresholds ranging from -3 to 16.5. E.g. considering a threshold of 1.5, CardioVAI pathogenicity score lead to a specificity and sensitivity of 98% and 90% respectively.



Supp. Figure S2: Correlation between the percentage of interpreted pathogenic variants and the number of gene's phenotype in 23 CVDs genes



Supp. Figure S3: CardioVAI and CardioClassifier comparison on CLINVITAE pathogenic and benign variants

SUPPLEMENTARY TABLES

Class	Level of evidence	Name	ACMG-AMP description	CardioVAI Implementation
Pathogenic	Very Stong +4	PVS1	Null variant in a gene where LOF is a known mechanism of disease. This criterion has been removed in the CMP-EP adaptation since LOF are not a mechanism of disease in MYH7 gene.	Variant is stop gained/frameshift/canonical splicing site. Phenotype is a known LOF for a certain phenotype or ExAC Prec >= 0.9, ExAC Pli >= 0.9 in case of recessive and dominant phenotype respectively. Another pathogenic null variant in the same exon has been reported.
	Strong +3	PS1	Same amino acid change as a previously established pathogenic variant regardless of nucleotide change	The same amino change has been reported as pathogenic for the same phenotype in ClinVar
		PS2	De novo in a patient with the disease and no family history	Not Implemented
		PS3	Well-established in vitro or in vivo functional studies supportive of a damaging effect on the gene or gene product. If applied to MYH7 variants, only functional studies of mammalian knock-in models should be considered, according to CMP-EP.	Variant has been reported as pathogenic for the same phenotype in ClinVar and the submission methods report the presence of functional studies. For MYH7 variants, only mammalian knock-in models are considered.
		PS4	The prevalence of the variant in affected individuals is significantly increased compared with the prevalence in controls	Not Implemented
		PM1	Located in a mutational hot spot and/or critical and well-established functional domain without benign variation	Variant overlaps a known hotspot domain
		PM2	Absent from controls (or at extremely low frequency if recessive) in ESP, 1000 GP, ExAC. Accoring to CMP-EP reccomendations for MYH7 gene, variant should be absent or rare (<0.004%) from large population studies.	Variant is absent from ExAC and 1KGP and ESP frequency <= 1% in case of dominant phenotype. Variant has frequency <= 0.01% in ExAC and 1KGP and ESP frequency <= 1% in case of recessive phenotype. For MYH7 variants, the above frequency thresholds are set to 0.004%.
	Moderate +2	PM3	For recessive disorders, detected in *trans* with a pathogenic variant	Not Implemented
		PM4	Protein length changes as a result of in-frame deletions/insertions in a nonrepeat region or stop-loss variants	Variant is an in-frame deletion/insertion and does not overlap a repetitive region according to Repeat Masker (UCSC source)
		PM5	Novel missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before	A different amino-acid change at the same residue position has been reported in ClinVar as pathogenic
		PM6	Assumed de novo, but without confirmation of paternity and maternity	Not Implemented
	Supporting +1	PP1	Cosegregation with disease in multiple affected family members in a gene definitively known to cause the disease	Not Implemented
		PP2	Missense variant in a gene that has a low rate of benign missense variation and in which missense variants are a common mechanism of disease. This criterion is removed in the CMP-EP adaptations for MYH7 variants.	ExAC missense zscore > 1.96 and more than 5 missense pathogenic variants have been reported in the same gene for the same phenotype. Not considered for MYH7 variants.
		PP3	Multiple lines of computational evidence support a deleterious effect on the gene or gene product	Prediction is “DAMAGING” for PaPI, DANN or dbscSNV
		PP4	Patient’s phenotype or family history is highly specific for a disease with a single genetic etiology	Not Implemented

		PP5	Reputable source recently reports variant as pathogenic, but the evidence is not available to the laboratory to perform an independent evaluation. Criterion is not applied to MYH7 variants according to CMP-EP.	Variant has been previously reported as pathogenic in ClinVar for the same phenotype. Not considered for MYH7 variants.
Benign	Stand-Alone -1.5	BA1	Allele frequency is >5% in Esp, ExAC and 1000 GP. For MYH7 variants, allele frequency is >= 0.1% in ExAC.	Allele is >= 5% in ExAC, 1KGP, ESP. For MYH7 variants, the allele frequency in ExAC, 1KGP or ESP must be >=0.1%.
	Strong -1	BS1	Allele frequency is greater than expected for disorder. Allele frequency threshold for MYH7 variants is 0.02% in ExAC according to CMP-EP.	Phenotype's incidence is drawn from Orphanet. In case of recessive disorder, the carrier frequency is estimated by Hardy-Weinerg equilibrium. In case of dominant disorder, incidence is greater than 0.001% and incidence is greater than ExAC or 1KGP frequency. In case of dominant disorder, incidence is below 0.001% and incidence is greater than ExAC or 1KGP frequency. For MYH7 variants, the allele frequency in ExAC, 1KGP or ESP must be >= 0.02%
		BS2	Observed in a healthy adult individual for a recessive (homozygous), dominant (heterozygous), or X-linked (hemizygous) disorder, with full penetrance expected at an early age	Not Implemented
		BS3	Well-established in vitro or in vivo functional studies show no damaging effect on protein function or splicing	Variant has been reported as benign without conflicting interpretation in ClinVar for the same phenotype and submission method report the presence of functional studies
		BS4	Lack of segregation in affected members of a family	Not Implemented
	Supporting -0.5	BP1	Missense variant in a gene for which primarily truncating variants are known to cause disease. Removed for MYH7 variants according CMP-EP.	Variant is coding and missense and at least 5 null variants in the gene have been reported as pathogenic for the same phenotype in ClinVar. Not considered for MYH7 variants.
		BP2	Observed in *trans* with a pathogenic variant for a fully penetrant dominant gene/disorder or observed in *cis* with a pathogenic variant in any inheritance pattern	Not Implemented
		BP3	In-frame deletions/insertions in a repetitive region without a known function. Removed for MYH7 variants according CMP-EP	Variant is in-frame deletion/insertion , but overlaps a repetitive region (UCSC RepeatMasker). Not considered for MYH7 variants.
		BP4	Multiple lines of computational evidence suggest no impact on gene or gene product	Prediction is "BENIGN" for PaPI, DANN or dbscSNV
		BP5	Variant found in a case with an alternate molecular basis for disease	Not Implemented
		BP6	Reputable source recently reports variant as benign, but the evidence is not available to the laboratory to perform an independent evaluation. Removed for MYH7 variants according CMP-EP	Variant has been previously reported in ClinVar as benign. Not considered for MYH7 variants.
		BP7	A synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site AND the nucleotide is not highly conserved	Variant is synonymous and "BENIGN" for dbscSNV AB and DANN
		BP8	Same amino acid change as a previously established benign variant regardless of nucleotide change	The same amino change has been previously reported in ClinVar as benign

Supp. Table S1: ACMG-AMP criteria and their implementation. The "ACMG-AMP description" column lists criteria description as stated in the ACMG-AMP variant interpretation guidelines. The description may also include how each criterion has been possibly specialized by the CMP-EP for MYH7 variants. The column "CardioVAI implementation" details cardioVAI implementation of ACMG-AMP criteria and CMP-EP criteria, along with all thresholds used.

GENE	PHENOTYPE	MEDGEN CUI	INHERITANCE	LOF/GOF	PMID references
LMNA	Benign scapuloperoneal muscular dystrophy with cardiomyopathy	C0410190	Autosomal dominant	LOF	23853504
KCNH2	Long QT syndrome 2	C3150943	Autosomal dominant	LOF	22529064
ANK2	Long QT syndrome 4	C1833154	Autosomal dominant	LOF	22529064
SCN5A	Brugada syndrome	C1142166	Autosomal dominant	LOF	22529064
CAV3	Familial hypertrophic cardiomyopathy 1	C3495498	Autosomal dominant	LOF	14672715
RYR2	Cathecolaminergic Polymorphic Ventricular Tachycardia	C1631597	Autosomal dominant	GOF	23549275
CACNA1C	Timothy syndrome	C1832916	Autosomal dominant	GOF	20301577
TRPM4	Progressive familial heart block type 1B	C1970298	Autosomal dominant	GOF	21887725
KCNJ2	Andersen Tawil syndrome	C1563715	Autosomal dominant	LOF	24383070
KCNQ1	Short QT syndrome 2	C1865019	Autosomal dominant	GOF	22529064

Supp. Table S2: Known LOF/GOF genes and related phenotypes

Gene	Source	Literature evidence	dbNSFP variants
ABCC9	CardioDB, Illumina	PMID 15034580; MedGen C0007193;	11905
ACTC1	CardioDB, Illumina	PMID 9563954; MedGen C0340427;	2680
ACTN2	CardioDB, Illumina	PMID 1456970;MedGen C0340427;	7065
AKAP9	In-house kb	PMID 18093912; MedGen C2678483;	29552
ALG10	In-house kb	PMID 15280551; MedGen C3150957;	3405
ANK2	CardioDB, In-house kb	PMID 12571597; MedGen C1833154;	29628
ANKRD1	CardioDB, Illumina	PMID 22892539; MedGen C0007193;	2346
CACNA1C	In-house kb	PMID 17224476; MedGen C1142166;	17746
CACNB2	In-house kb	PMID 17224476; MedGen C1142166	5517
CALM1	In-house kb	PMID 23388215; MedGen CN228133;	1166
CALM2	In-house kb	PMID 23388215; MedGen C4015695;	1822
CALM3	In-house kb	PMID 28491681; MedGen C0023976;	1154
CALR3	CardioDB	PMID 17655857; MedGen C0949658;	2870
CASQ2	All	PMID 11704930; MedGen C2677794;	3000
CAV3	CardioDB, Illumina	PMID 17060380; MedGen C0340427;	1079
CRYAB	CardioDB, Illumina	PMID 21920752; MedGen C0340427;	1562
CSRP3	CardioDB, Illumina	PMID 14567970; MedGen C0340427	1424
DES	CardioDB, Illumina	PMID 17720647; MedGen C0340427	3331
DSC2	All	PMID 17033975; MedGen C0349788	6547
DSG2	All	PMID 18678517; MedGen C0340427	7938
DSP	All	PMID 16467215; MedGen C0349788;	20639
DTNA	CardioDB, Illumina	PMID 11238270; MedGen C1858725;	6402
EMD	CardioDB, Illumina	PMID 8248200; MedGen C0410189;	1817
FHL1	CardioDB	PMID 19716112; MedGen C0410189;	2859
FXN	CardioDB	PMID 10543403;	1655

		MedGen C0016719;	
GLA	CardioDB, Illumina	PMID 2539398; MedGen C1970820;	3123
GPD1L	In-house kb	PMID 9521325; MedGen C1142166;	2530
JPH2	CardioDB	PMID 17509612; MedGen C0949658;	4767
JUP	All	PMID 20130592; MedGen C0349788;	5230
KCNE1	CardioDB, In-house kb	PMID 10400998; MedGen C1867904;	899
KCNE2	CardioDB, In-house kb	PMID 10219239; MedGen C3150953;	889
KCNH2	In-house kb	PMID 11997281; MedGen C3150957;	8829
KCNJ2	CardioDB	PMID 15761194; MedGen C2348199;	3016
KCNJ5	In-house kb	PMID 20560207; MedGen C3150733;	2946
KCNJ8	In-house kb	PMID 23632791; MedGen C1142166;	2965
KCNQ1	CardioDB, In-house kb	PMID 8528244; MedGen C3150944;	5243
LAMA4	CardioDB, Illumina	PMID:17646580; MedGen: C0340427;	13828
LAMP2	CardioDB, Illumina	PMID 10972294; MedGen C0878677;	3594
LDB3	All	PMID 14662268; MedGen C0007193;	5890
LMNA	All	PMID 10580070; MedGen C0340427;	5603
MYBPC3	All	PMID 7493025; MedGen C0949658;	9158
MYH6	CardioDB, Illumina	PMID 15998695; MedGen C0340427;	14093
MYH7	All	PMID 1110671; MedGen C0340427;	14063
MYL2	CardioDB, Illumina	PMID 9535554; MedGen C0949658;	1342
MYL3	CardioDB, Illumina	PMID 8673105; MedGen C0949658;	1439
MYLK2	CardioDB, Illumina	PMID 11733062; MedGen C0949658;	4281
MYOZ2	CardioDB, Illumina	PMID 17347475; MedGen C0949658;	1916
MYPN	CardioDB	MedGen:C0340427;	9687
NEXN	CardioDB, Illumina	PMID 19881492; MedGen C0340427;	5048
PKP2	All	PMID 15489853; MedGen C0349788;	6260
PLN	CardioDB, Illumina	PMID 12610310;	372

		MedGen C0340427;	
PRKAG2	CardioDB, Illumina	PMID 11371514; MedGen C0949658;	4341
RBM20	CardioDB, Illumina	PMID:22466703,20590677; MedGen:C0340427,C0007193;	8666
RYR2	All	PMID 11159936; MedGen C0349788;	36453
SCN10A	In-house kb	PMID 23872634; MedGen C1142166;	13944
SCN4B	In-house kb	PMID 17592081; MedGen C2678484;	1621
SCN5A	CardioDB, In-house kb	PMID 9521325; MedGen C1142166;	14658
SGCD	CardioDB, Illumina	PMID 10974018; MedGen C0340427;	2250
SNTA1	In-house kb	PMID 18591664; MedGen C2751830;	3451
TAZ	CardioDB, Illumina	PMID 23031367; MedGen C0340427;	2549
TCAP	CardioDB, Illumina	PMID 15582318; MedGen C0340427;	1160
TGFB3	CardioDB	PMID 15639475; MedGen C0349788;	2959
TMEM43	All	PMID 18313022; MedGen C0349788;	2871
TNNC1	CardioDB, Illumina	PMID 16302972; MedGen C0949658;	1237
TNNI3	CardioDB, Illumina	PMID 9241277; MedGen C0949658;	1968
TNNT2	All	PMID 8205619; MedGen C0949658;	2433
TPM1	CardioDB, Illumina	PMID 7898523; MedGen C0949658;	3973
TRDN	In-house kb	PMID 22422768; MedGen C3809536;	5949
TRPM4	In-house kb	PMID 19726882; MedGen C1970298;	8487
TTN	CardioDB, Illumina	PMID 10462489; MedGen C0949658;	183819
TTR	CardioDB, Illumina	PMID 12050338; MedGen C3151471;	1361
VCL	CardioDB, Illumina	PMID 11815424; MedGen C0340427;	8063

Supp. Table S1: List of 72 genes associated with CVDs according to CardioDB, Illumina . The last column represents the number of non-synonymous variants collected from dbNSFP in each gene (total number of variants 624095).

GENE	CLINVITAE	CARDIODB
NEXN	yes	no
CASQ2	yes	yes
LMNA	yes	yes
ACTN2	yes	no
RYR2	yes	yes
CACNB2	yes	no
MYPN	yes	no
VCL	yes	no
LDB3	yes	yes
ANKRD1	yes	no
RBM20	yes	no
KCNQ1	yes	no
CSRP3	yes	no
MYBPC3	yes	yes
CRYAB	yes	no
CACNA1C	yes	no
ABCC9	yes	no
PKP2	yes	yes
MYL2	yes	no
MYH6	yes	no
MYH7	yes	yes
ACTC1	yes	no
TPM1	yes	no
TCAP	yes	no
JUP	yes	yes
KCNJ2	yes	no
DSC2	yes	no
DSG2	yes	yes
TTR	yes	no
TNNI3	yes	no
TTN	yes	no
DES	yes	no
MYLK2	yes	no
SNTA1	yes	no
JPH2	yes	no
KCNE1	yes	no
CAV3	yes	no
TMEM43	yes	yes
GPD1L	yes	no
SCN5A	yes	yes
SCN10A	yes	no
MYL3	yes	no
TNNC	yes	no
ANK2	yes	no
SGCD	yes	no

DSP	yes	yes
LAMA4	yes	no
TRDN	yes	no
AKAP9	yes	no
KCNH2	yes	no
PRKAG2	yes	no
FXN	yes	no
GLA	yes	no
LAMP2	yes	no
FHL1	yes	no
EMD	yes	no
TAZ	yes	no
TNTT2	no	yes

Supp. Table S4: CVDs-related genes in CLINVITAE and CardioDB benchmark datasets. Only 12 genes are both on CLINVITAE and CardioDB, 45 are found in CLINVITAE (but not in CardioDB) and 1 gene is found in CardioDB only.

Criteria	Resources
PVS1	-ExAC (pREC, pLI) -MedGen (inheritance) -ClinVar -Custom LOF/GOF gene-disease (see Supplemental Table 1)
PS1, PM5, PP5, BP6, BS3, BP8	-ClinVar -MedGen -Disease Ontology
PM2, BA1, BS1	-ExAC -1000 Genomes -Exome Sequencing Project (ESP) -Orphanet (incidence/prevalence)
PM1	-ClinVar (pathogenic hotspots) -Custom Functional Domains (see Table 3)
PP2	-ExAC (missense zscore)
PP3, BP4, BP7	-PaPI -DANN -dbscSNV
BP3	-UCSC RepeatMasker

Supp. Table S5: Resources used for criteria implementation.

SUPPLEMENTAL REFERENCES

Kibbe, W. A., Arze, C., Felix, V., Mitraka, E., Bolton, E., Fu, G., … Schriml, L. M. (2015). Disease Ontology 2015 update: an expanded and updated database of human diseases for linking biomedical knowledge through disease data. *Nucleic Acids Research*, *43*(Database issue), D1071-1078. https://doi.org/10.1093/nar/gku1011