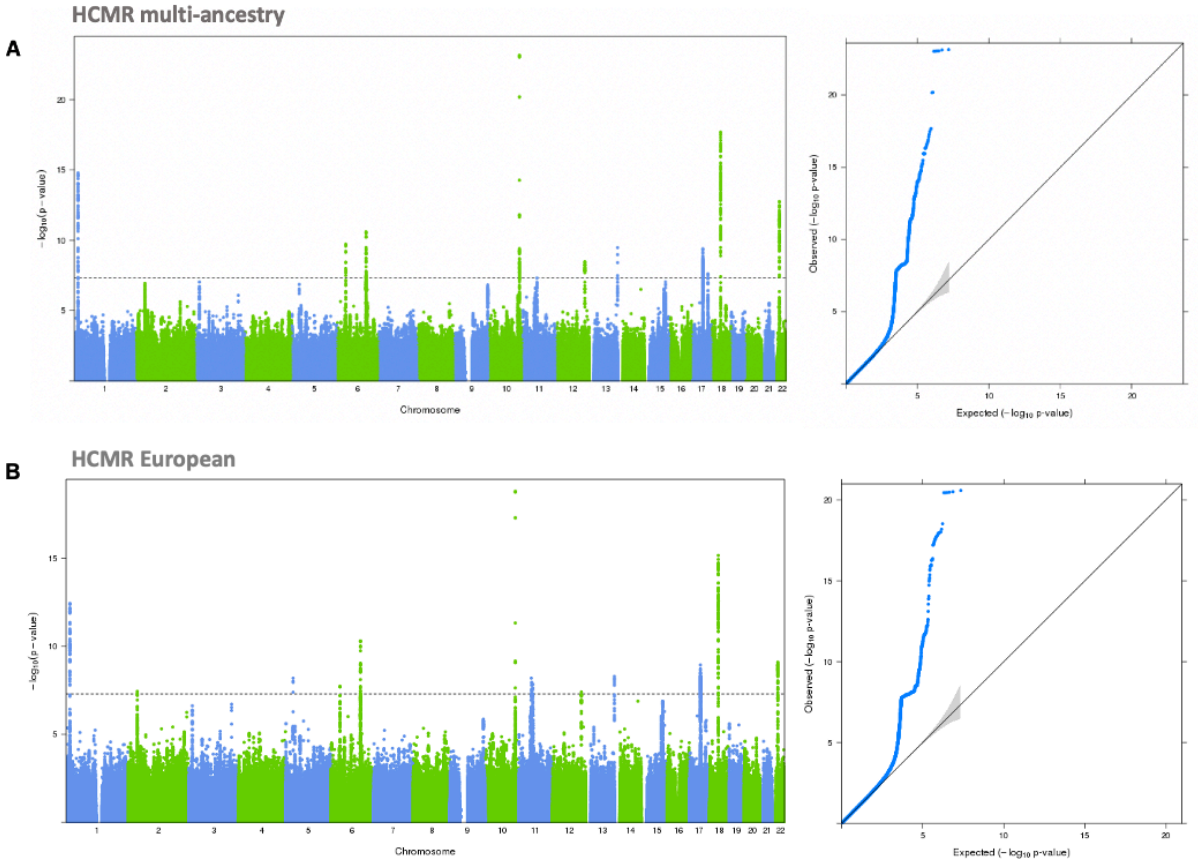
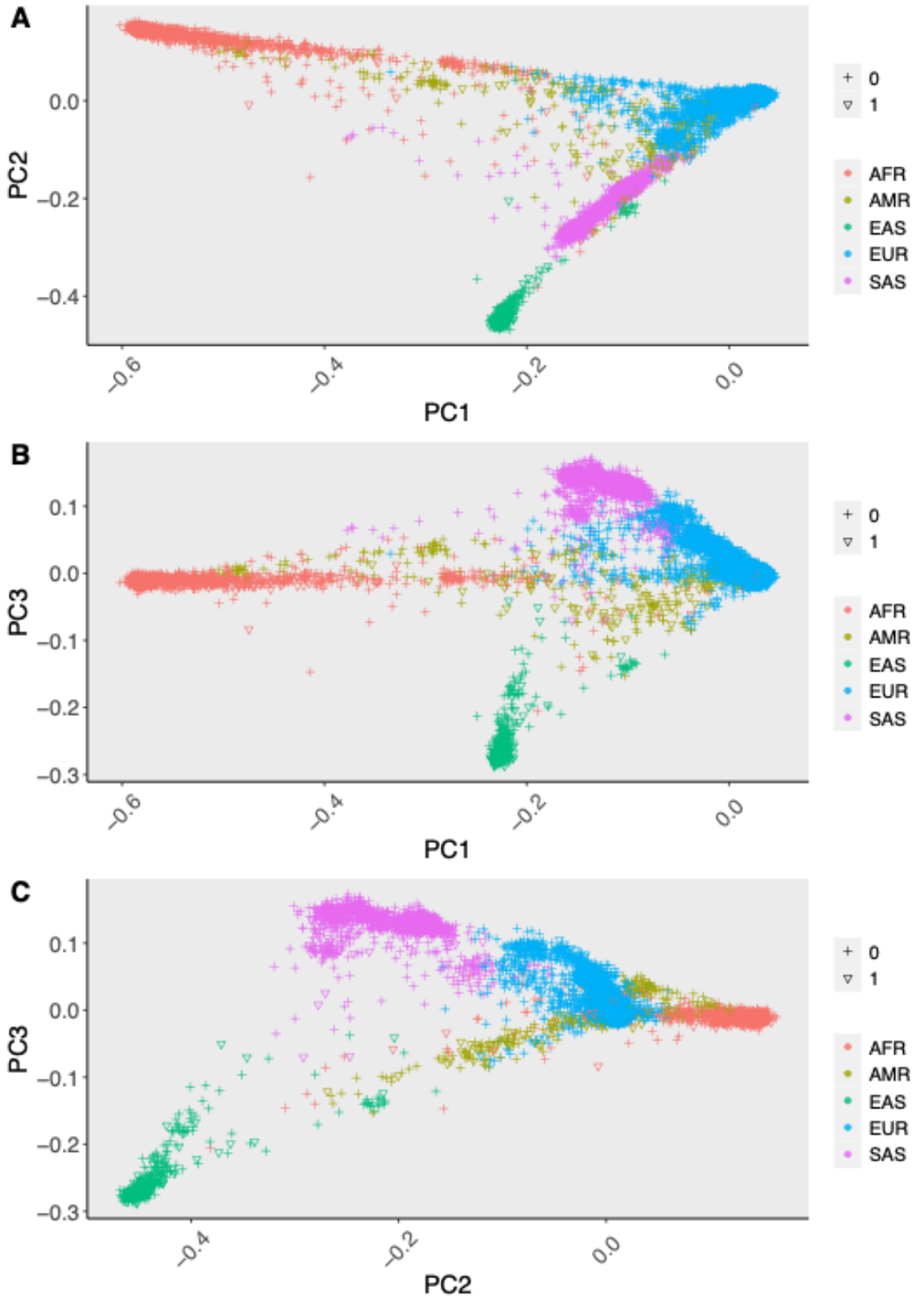
## Supplementary Note

Contents:

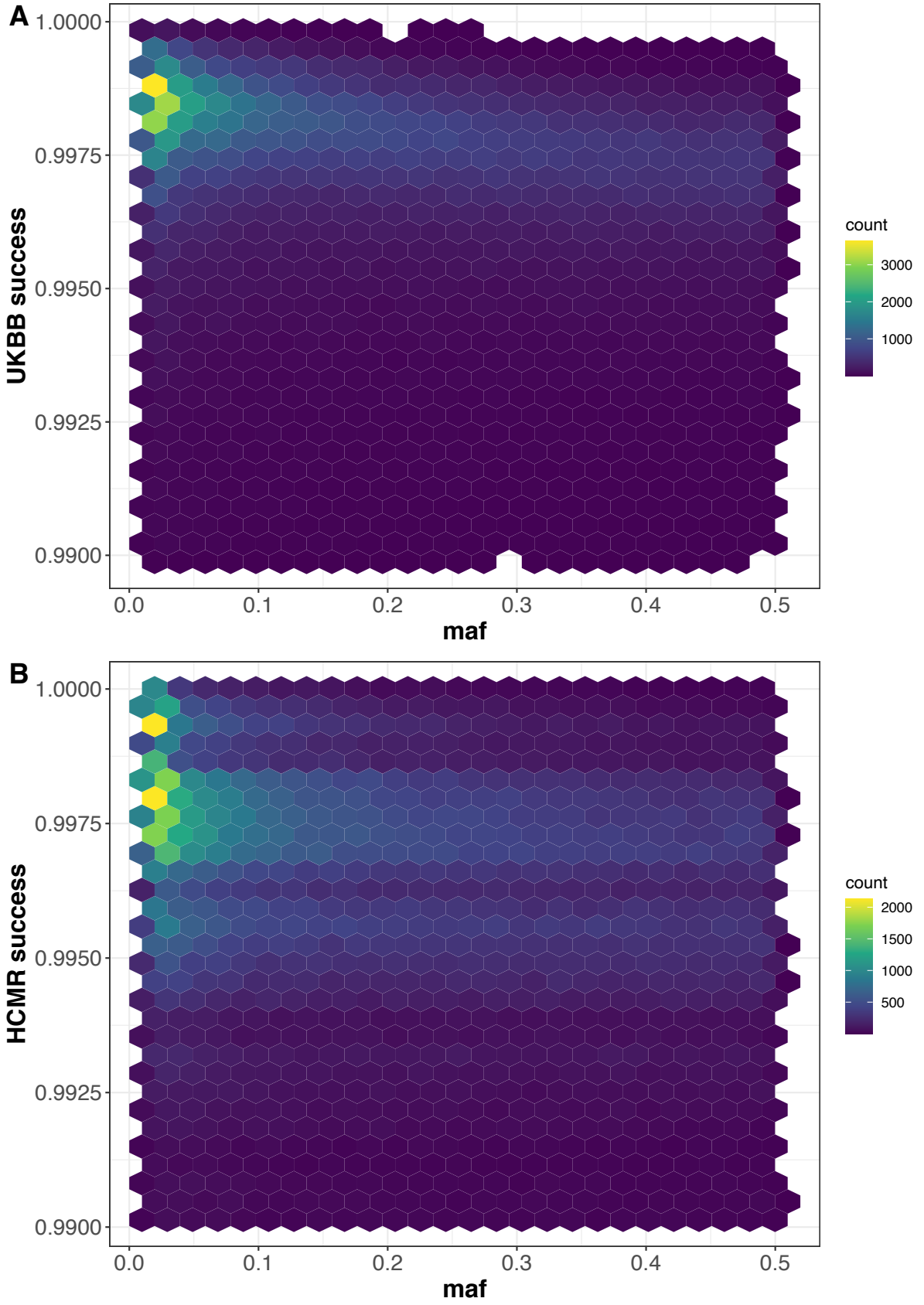
## 1. GWAS studies




**Suppl Note Figure 1: HCMR vs. UKBB GWAS.** Panel A) Multi-ancestry analysis. Following genomic control correction, (original $\lambda$ = 1.191, genomically controlled to a $\lambda$ of 1.0), the multi-ancestry GWAS demonstrated 11 independent genome-wide significant variants, and a further 12 variants <5% false discovery rate (FDR) threshold (p-value: $1.5\times10^{-6}$), associated with hypertrophic cardiomyopathy (Supplementary Table S2 and Suppl Note Table 1). Panel B) European analysis. Following genomic control correction, (original $\lambda$ = 1.148, corrected $\lambda$ 1.0), the multi-ancestry GWAS demonstrated 13 independent genome-wide significant variants, and a further 9 variants <5% FDR threshold. Manhattan plot and accompanying QQ plot. Manhattan plot split by chromosome (x-axis) and -$\log_{10}$(p-value) (y-axis). Dotted line represents genome-wide significance threshold (p-value = 5E-8). QQ plot of expected and observed -$\log_{10}$(p-value) values. Analysis performed using logistic regression to fit an additive case-control association model using SNPTEST v.2.5.4-beta3. *P* uncorrected for multiple testing.



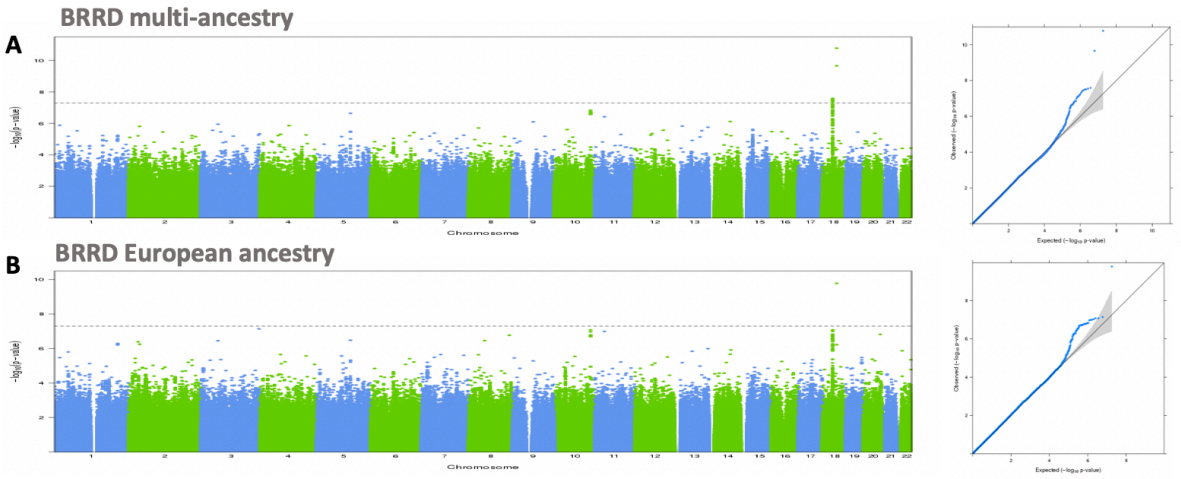

**Suppl Note Figure 2: Principal component analysis for HCMR vs. UKBB cases and controls**. Cases (denoted in the key as "1") and controls (denoted in the key as "0") are colour coded to reflect each individual's average genetic ancestry. This has been informed via principal component analysis, conducted using FlashPCA2, with reference to 1000 Genomes Phase 3 data. Abbreviations: AFR: African; AMR: Ad Mixed American; EAS: East Asian; EUR: European; SAS: South Asian.



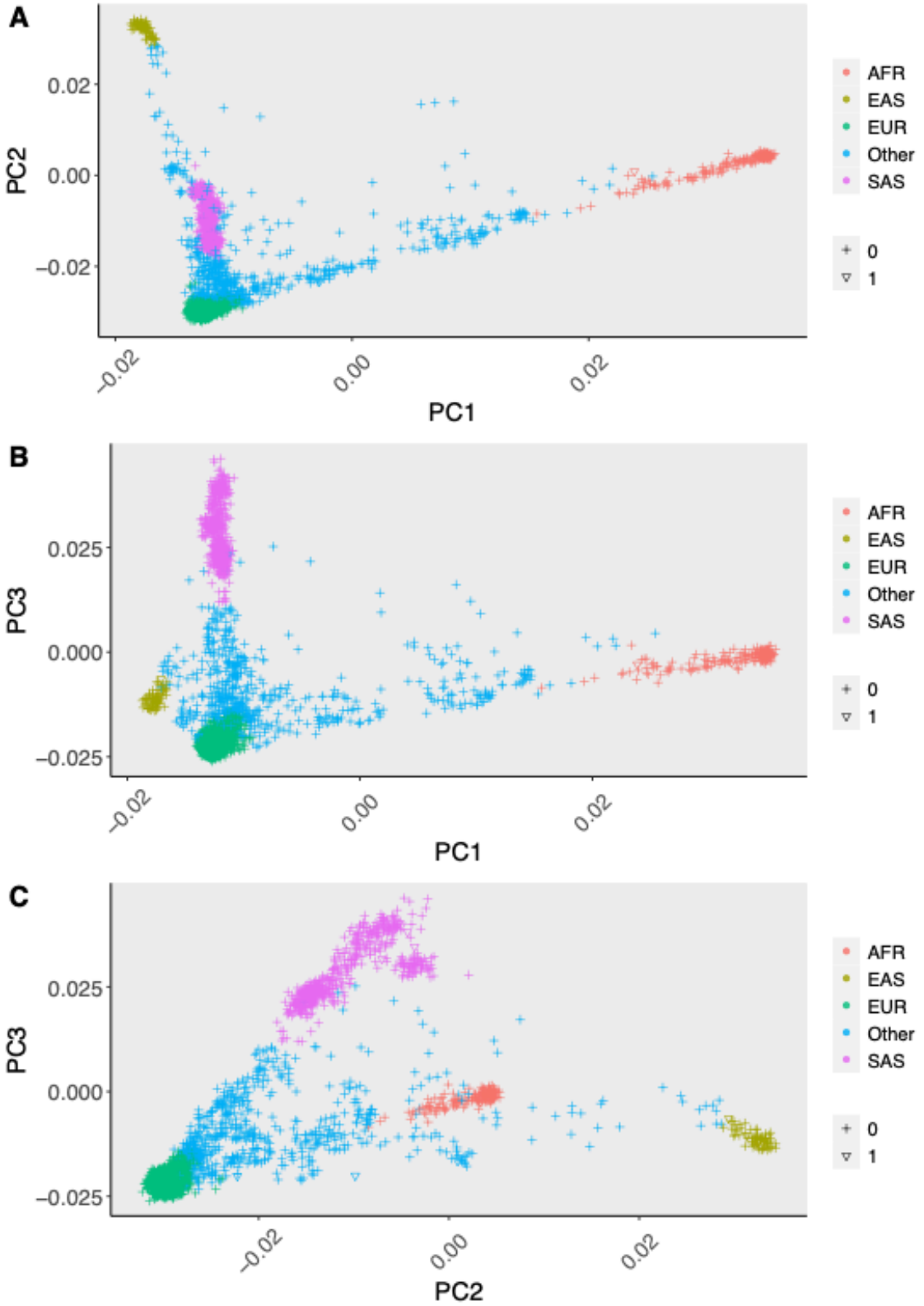

**Suppl Note Figure 3: Genotyping success rate of SNPs used for imputation with the HRC reference panel.** Panel A) UKBB SNPs. Panel B) HCMR SNPs. All 134k SNPs included used for imputation are of high genotyping quality across all minor allele frequency thresholds.

| Locus | Chr | Multi-ancestry | | | | | | European | | | | | | LD | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | SNP | EA | EAF | Summary statistics | | | SNP | EA | EAF | Summary statistics | | | | |
| | | | | | Beta | SE | *P* | | | | Beta | SE | *P* | D' | $R^2$ |
| *BAG3* | 10 | rs72840788 | G | 0.791 | -0.399 | 0.040 | 7.00E-24 | rs2234962 | T | 0.783 | -0.371 | 0.041 | 1.50E-19 | 0.9954 | 0.9863 |
| *HSPB7* | 1 | rs1048302 | T | 0.326 | 0.275 | 0.035 | 1.67E-15 | rs10803407 | T | 0.336 | 0.278 | 0.038 | 3.59E-13 | 0.9154 | 0.8128 |
| *FHOD3* | 18 | rs2644262 | T | 0.715 | -0.325 | 0.037 | 2.08E-18 | rs4477795 | C | 0.656 | -0.303 | 0.038 | 6.66E-16 | 0.6463 | 0.4144 |
| *MMP11* | 22 | rs7284877 | G | 0.219 | 0.293 | 0.040 | 1.81E-13 | rs2267038 | G | 0.192 | 0.273 | 0.044 | 7.93E-10 | 0.9139 | 0.6979 |
| *PLN* | 6 | rs12212795 | G | 0.947 | -0.433 | 0.065 | 2.57E-11 | rs12212795 | G | 0.946 | -0.453 | 0.069 | 4.80E-11 | 1 | 1 |
| *CDKN1A* | 6 | rs762624 | A | 0.718 | -0.228 | 0.036 | 1.91E-10 | rs56100429 | C | 0.720 | -0.227 | 0.040 | 1.82E-08 | 0.622 | 0.155 |
| *ADPRHL1* | 13 | rs41306688 | A | 0.966 | -0.589 | 0.094 | 3.37E-10 | rs41306688 | A | 0.965 | -0.570 | 0.098 | 5.04E-09 | 1 | 1 |
| *SPPL2C* | 17 | rs393838 | G | 0.774 | -0.242 | 0.039 | 4.04E-10 | rs7502280 | T | 0.886 | -0.396 | 0.065 | 1.12E-09 | NA | NA |
| *TBX3* | 12 | rs7300371 | T | 0.267 | -0.240 | 0.041 | 3.45E-09 | rs7300371 | T | 0.265 | -0.247 | 0.045 | 3.86E-08 | 1 | 1 |
| *PRKCA* | 17 | rs7210446 | G | 0.419 | -0.214 | 0.038 | 2.50E-08 | - | - | - | - | - | - | - | - |
| *OR5AK2* | 11 | rs78310129 | C | 0.988 | -0.768 | 0.141 | 5.05E-08 | rs186327160 | T | 0.988 | -0.917 | 0.161 | 1.21E-08 | 0.8997 | 0.5391 |
| *FHOD3* | 18 | rs118060942 | C | 0.988 | -0.773 | 0.119 | 8.61E-11 | rs118060942 | C | 0.987 | -0.804 | 0.122 | 4.57E-11 | 1 | 1 |
| *SLC6A6* | 3 | rs13061705 | C | 0.685 | 0.225 | 0.042 | 9.46E-08 | rs13061705 | C | 0.676 | 0.229 | 0.044 | 2.35E-07 | 1 | 1 |
| *ALPK3* | 15 | rs8033459 | C | 0.529 | -0.182 | 0.034 | 9.34E-08 | rs11073663 | A | 0.545 | -0.194 | 0.037 | 1.32E-07 | 0.9753 | 0.9157 |
| *STRN* | 2 | rs11124555 | G | 0.514 | -0.182 | 0.034 | 1.18E-07 | rs1861435 | T | 0.574 | -0.203 | 0.037 | 3.51E-08 | 0.7624 | 0.3244 |
| *AK098570* | 5 | rs66761011 | A | 0.830 | -0.349 | 0.066 | 1.42E-07 | rs66761011 | A | 0.825 | -0.396 | 0.068 | 6.26E-09 | 1 | 1 |
| *RAPGEF1, POMT1* | 9 | rs734638 | C | 0.711 | -0.196 | 0.037 | 1.53E-07 | rs877373 | C | 0.710 | -0.190 | 0.039 | 1.35E-06 | 0.9197 | 0.5852 |
| *TRDN* | 6 | rs9320939 | G | 0.516 | -0.174 | 0.034 | 3.30E-07 | - | - | - | - | - | - | - | - |
| *ADAMTS7* | 15 | rs8043123 | C | 0.757 | -0.194 | 0.038 | 3.03E-07 | rs1114829 | G | 0.782 | -0.205 | 0.042 | 9.29E-07 | 0.6537 | 0.2237 |
| *SSRP1* | 11 | rs117534260 | A | 0.981 | -0.534 | 0.108 | 8.22E-07 | - | - | - | - | - | - | - | - |
| *FNDC3B* | 3 | rs4894803 | A | 0.596 | -0.186 | 0.038 | 8.32E-07 | rs4894803 | A | 0.585 | -0.209 | 0.040 | 1.91E-07 | 1 | 1 |
| *TCF7L2* | 10 | rs11196085 | T | 0.722 | -0.184 | 0.037 | 8.85E-07 | - | - | - | - | - | - | - | - |
| *CYP2R1* | 11 | rs1390519 | A | 0.666 | -0.203 | 0.042 | 1.14E-06 | - | - | - | - | - | - | - | - |
| *MAPK8IP1P2* | 17 | rs7502280 | T | 0.892 | -0.373 | 0.063 | 3.91E-09 | rs7502280 | T | 0.886 | -0.396 | 0.065 | 1.12E-09 | - | - |
| *SLCO3A1* | 15 | rs143720580 | G | 0.987 | -0.581 | 0.151 | 1.15E-04 | rs143720580 | G | 0.988 | -0.758 | 0.151 | 4.90E-07 | - | - |
| *SPP2* | 2 | rs75944971 | C | 0.950 | -0.348 | 0.080 | 1.25E-05 | rs75944971 | C | 0.948 | -0.408 | 0.082 | 5.67E-07 | - | - |
| *ABCA5* | 17 | rs79000044 | A | 0.985 | -0.556 | 0.135 | 3.64E-05 | rs79000044 | A | 0.985 | -0.652 | 0.136 | 1.60E-06 | - | - |
| *OR9Q1* | 11 | rs78631951 | C | 0.984 | -0.534 | 0.120 | 8.88E-06 | rs78631951 | C | 0.983 | -0.577 | 0.122 | 2.30E-06 | - | - |

**Suppl Note Table 1: HCMR vs UKBB results for multi-ancestry and European only analysis**. Chr = chromosome; SNP = single nucleotide polymorphism; EA = effect allele; EAF = effect allele frequency; SE = standard error; LD = linkage disequilibrium. Analysis performed using logistic regression to fit an additive case-control association model using SNPTEST v.2.5.4-beta3. *P* uncorrected for multiple testing.



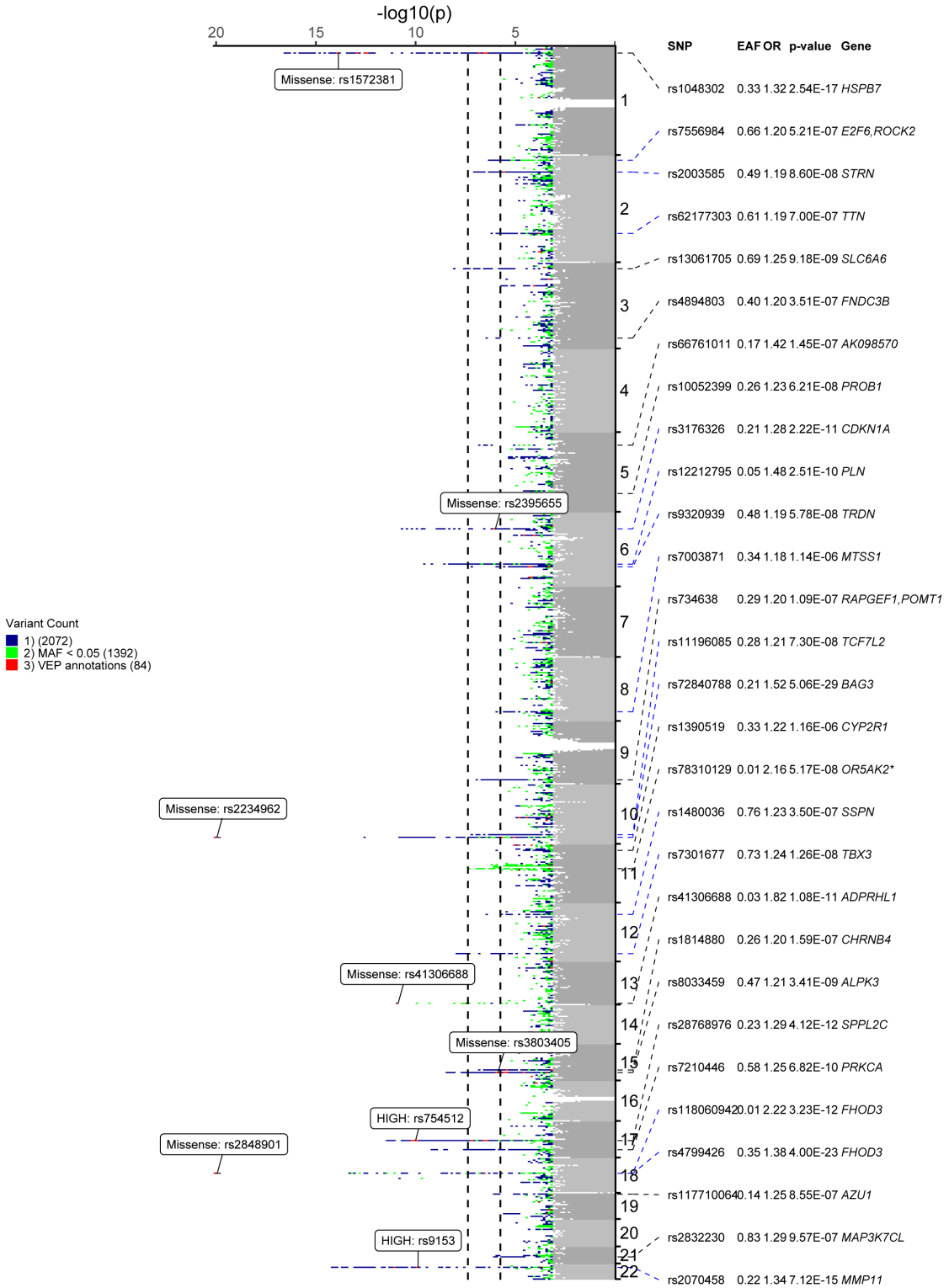

**Suppl Note Figure 4: BRRD vs. BRRD GWAS.** Panel A) Multi-ancestry analysis: There was no evidence of extreme genomic inflation ($\lambda$ = 1.049). Three SNPs (rs1232572641, rs142939703 and rs16968220) demonstrated genome-wide significance and 2 (rs61869036 and rs139472654) were below the 5% false discovery rate (FDR) threshold (Supplementary Table S2 and Suppl Note Table 2). There is no linkage disequilibrium between rs1232572641 and rs142939703 ($r^2$= 0.054 and D'=0.362), but do not contribute towards COJO analyses as they are not present in the linkage disequilibrium backbone pseudo-randomly generated from UK Biobank genotypes. Multi-ancestry Manhattan plot and accompanying QQ plot. Manhattan plot split by chromosome (x-axis) and -$\log_{10}$(p-value) (y-axis). Dotted line represents genome-wide significance threshold ($5\times10^{-8}$). QQ plot of expected and observed -$\log_{10}$(p-value) values. Panel B) European only analysis: There was no evidence of extreme genomic inflation ($\lambda$ = 1.013) analysis. Manhattan plot and accompanying QQ plot. Manhattan plot split by chromosome (x-axis) and -$\log_{10}$(p-value) (y-axis). Dotted line represents genome-wide significance threshold (p-value = $5\times10^{-8}$). QQ plot of expected and observed -$\log_{10}$(p-value) values. Analyses generated using a mixed logistic regression model in SAIGE. *P* uncorrected for multiple testing.



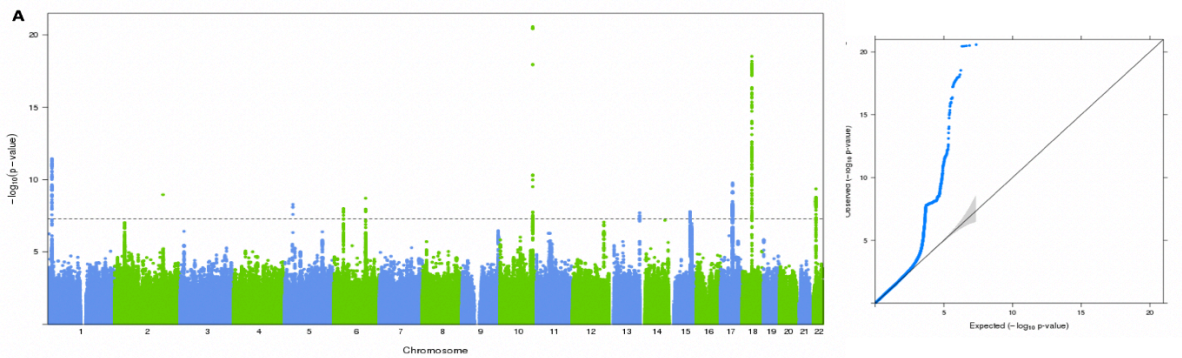

**Suppl Note Figure 5: Principal component analysis for BRRD vs. BRRDD cases and controls.** Cases (denoted in the key as "1") and Control (denoted in the key as "0") are colour coded to reflect each individual's average genetic ancestry. Cases (denoted in the key as "1") and Control (denoted in the key as "0"). Abbreviations: AFR: African; AMR: Ad Mixed American; EAS: East Asian; EUR: European; SAS: South Asian.

| Locus | Chr | SNP | EA | Multi-ancestry | | | | European | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | EAF | Summary statistics | | | EAF | Summary statistics | | |
| | | | | | beta | se | p-value | | beta | se | p-value |
| *FHOD3* | 18 | rs16968220 | C | 0.647 | -0.553 | 0.100 | 2.83E-08 | 0.657 | -0.569 | 0.109 | 1.66E-07 |
| *FHOD3* | 18 | rs12605417 | C | 0.652 | -0.553 | 0.100 | 3.23E-08 | 0.657 | -0.581 | 0.109 | 9.28E-08 |
| *BAG3* | 10 | rs61869036 | G | 0.799 | -0.640 | 0.122 | 1.46E-07 | 0.791 | -0.698 | 0.131 | 1.07E-07 |
| *YTHDC2* | 5 | rs139472654 | C | 0.977 | -1.806 | 0.349 | 2.29E-07 | 0.975 | -1.918 | 0.376 | 3.30E-07 |
| *INTS7* | 1 | rs112269880 | A | 0.976 | -1.432 | 0.329 | 1.34E-05 | 0.975 | -1.861 | 0.370 | 5.07E-07 |

**Suppl Note Table 2: BRRD vs. BRRD results for multi-ancestry and European only analysis.** Chr = chromosome; SNP = single nucleotide polymorphism; EA = effect allele; EAF = effect allele frequency; SE = standard error; LD = linkage disequilibrium. Analyses generated using a mixed logistic regression model in SAIGE. *P* uncorrected for multiple testing.




**Suppl Note Figure 6: Multi-ancestry HCM GWAS meta-analysis.** A fixed effects meta-analysis incorporating 8,585,485 SNPs implemented using GWAMA for 2,780 HCMR/BRRD cases vs. 47,486 UKBB/BRRD controls reveals 13 independent loci at genome-wide significance (p-value < $5 \times 10^{-8}$) and a further 16 SNPs beneath the 5% false discovery rate (FDR) threshold (p-value < $1.82 \times 10^{-6}$) (Supplementary Table S2 and Suppl Note Table 3). Variants with a MAF <0.05 are reported in green. Missense variants and variant of high impact, as annotated by Ensembl's Variant Effect Predictor, are highlighted in red and labelled. No evidence of overdispersion ($\lambda$ = 1.03). Manhattan plot produced using Manhattan ++.[1] **OR5AK2* locus encompassed in region demonstrating long-range LD with pathogenic *MYBPC3* variants (Suppl Note Figure 16 and Suppl Note Table 4), as demonstrated by stack of green (MAF < 0.05) variants. Performed using a fixed-effects inverse-variance meta-analysis in GWAMA. *P* uncorrected for multiple testing.




**Suppl Note Figure 7: European HCM GWAS meta-analysis.** A total of 8,356,574 variants from 2,244 HCM cases, age and sex matched with 42,668 controls, were meta-analysed using two European-only case-control studies (HCMR vs. UKBB and BRRD vs. BRRD). There was no evidence of extreme genomic inflation ($\lambda$ = 1.04). A total of 12 independent genome-wide significant variants and 11 independent variants with a p-value less than the 5% FDR threshold demonstrated associations with risk of HCM (Supplementary Table S2 and Suppl Note Table 3). Performed using a fixed-effects inverse-variance meta-analysis in GWAMA. *P* uncorrected for multiple testing.

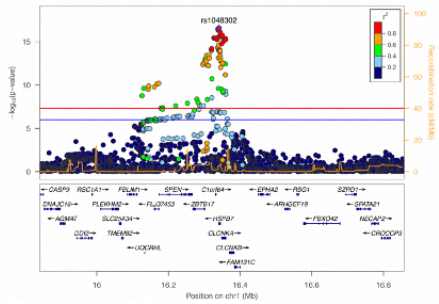
| | | Multi-ancestry | | | | | | | European | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Locus | Chr | SNP | EA | NEA | EAF | Summary statistics | | | SNP | EA | EAF | Summary statistics | | | LD | |
| | | | | | | Beta | SE | *P*-value | | | | Beta | SE | *P*-value | D' | R2 |
| *HSPB7* | 1 | rs1048302 | T | G | 0.325 | 0.277 | 0.033 | 2.54E-17 | rs10803407 | T | 0.335 | 0.276 | 0.036 | 1.97E-14 | 0.9154 | 0.8128 |
| *PRKRA* | 2 | - | - | - | - | - | - | - | rs117491478 | G | 0.988 | -0.782 | 0.146 | 9.12E-08 | - | - |
| *STRN* | 2 | rs2003585 | T | C | 0.490 | 0.174 | 0.032 | 8.60E-08 | rs17443495 | G | 0.576 | -0.192 | 0.035 | 3.69E-08 | 0.9812 | 0.3835 |
| *TTN* | 2 | rs62177303 | C | T | 0.614 | 0.175 | 0.035 | 7.00E-07 | - | - | - | - | - | - | - | - |
| *E2F6, ROCK2* | 2 | rs7556984 | G | A | 0.663 | 0.186 | 0.037 | 5.21E-07 | - | - | - | - | - | - | - | - |
| *SLC6A6* | 3 | rs13061705 | C | T | 0.687 | 0.224 | 0.039 | 9.18E-09 | rs13061705 | C | 0.675 | 0.234 | 0.041 | 1.17E-08 | 1 | 1 |
| *FNDC3B* | 3 | rs4894803 | A | G | 0.597 | -0.179 | 0.035 | 3.51E-07 | - | - | - | - | - | - | - | - |
| *PROB1* | 5 | rs10052399 | T | C | 0.264 | 0.206 | 0.038 | 6.21E-08 | rs10052399 | T | 0.259 | 0.209 | 0.041 | 3.81E-07 | 1 | 1 |
| . | 5 | rs66761011 | A | G | 0.830 | -0.349 | 0.066 | 1.45E-07 | rs66761011 | A | 0.825 | -0.396 | 0.068 | 6.46E-09 | 1 | 1 |
| *PLN* | 6 | rs12212795 | G | C | 0.947 | -0.393 | 0.062 | 2.51E-10 | rs28436726 | G | 0.941 | -0.404 | 0.064 | 3.79E-10 | 0.986 | 0.3478 |
| *CDKN1A* | 6 | rs3176326 | G | A | 0.791 | -0.247 | 0.037 | 2.22E-11 | rs56100429 | C | 0.720 | -0.220 | 0.038 | 6.60E-09 | 0.8516 | 0.5399 |
| *TRDN* | 6 | rs9320939 | G | A | 0.517 | -0.174 | 0.032 | 5.78E-08 | - | - | - | - | - | - | - | - |
| *MTSS1* | 8 | rs7003871 | C | T | 0.336 | 0.169 | 0.035 | 1.14E-06 | - | - | - | - | - | - | - | - |
| *RAPGEF1, POMT1* | 9 | rs734638 | C | G | 0.713 | -0.186 | 0.035 | 1.09E-07 | rs734638 | C | 0.704 | -0.187 | 0.037 | 6.09E-07 | 1 | 1 |
| *TCF7L2* | 10 | rs11196085 | T | C | 0.721 | -0.190 | 0.035 | 7.30E-08 | - | - | - | - | - | - | - | - |
| *BAG3* | 10 | rs72840788 | G | A | 0.792 | -0.421 | 0.038 | 5.06E-29 | rs61869036 | G | 0.784 | -0.403 | 0.039 | 1.77E-24 | 0.9977 | 0.9954 |
| . | 11 | rs1390519 | A | G | 0.666 | -0.203 | 0.042 | 1.16E-06 | - | - | - | - | - | - | - | - |
| *OR5AK2* | 11 | rs78310129 | C | T | 0.988 | -0.768 | 0.141 | 5.17E-08 | - | - | - | - | - | - | - | - |
| *SSPN* | 12 | rs1480036 | T | C | 0.236 | -0.211 | 0.041 | 3.50E-07 | - | - | - | - | - | - | - | - |
| *TBX3* | 12 | rs7301677 | C | T | 0.732 | 0.214 | 0.038 | 1.26E-08 | rs7977151 | A | 0.298 | -0.209 | 0.040 | 2.10E-07 | 0.9432 | 0.6641 |
| *ADPRHL1* | 13 | rs41306688 | A | C | 0.967 | -0.601 | 0.088 | 1.08E-11 | rs41306688 | A | 0.965 | -0.595 | 0.092 | 1.07E-10 | 1 | 1 |
| *CHRNB4* | 15 | rs1814880 | T | C | 0.263 | 0.179 | 0.034 | 1.59E-07 | rs6495317 | A | 0.783 | -0.204 | 0.040 | 2.59E-07 | 0.6123 | 0.3743 |
| *ALPK3* | 15 | rs8033459 | C | T | 0.529 | -0.190 | 0.032 | 3.41E-09 | rs77523699 | C | 0.917 | -0.290 | 0.058 | 6.83E-07 | 0.9772 | 0.0134 |
| *SPPL2C* | 17 | rs28768976 | A | G | 0.772 | -0.255 | 0.037 | 4.12E-12 | rs17769552 | G | 0.778 | -0.264 | 0.040 | 2.55E-11 | 1 | 1 |
| *PRKCA* | 17 | rs7210446 | G | A | 0.419 | -0.220 | 0.036 | 6.82E-10 | - | - | - | - | - | - | - | - |
| *FHOD3* | 18 | - | - | - | - | - | - | - | rs72891799 | C | 0.990 | -0.840 | 0.152 | 3.52E-08 | - | - |
| *FHOD3* | 18 | rs118060942 | C | T | 0.988 | -0.799 | 0.115 | 3.23E-12 | rs118060942 | C | 0.987 | -0.834 | 0.118 | 1.59E-12 | 1 | 1 |
| *FHOD3* | 18 | rs4799426 | A | G | 0.655 | -0.321 | 0.032 | 4.00E-23 | rs4477795 | C | 0.656 | -0.332 | 0.035 | 9.69E-21 | 0.9716 | 0.8369 |
| *AZU1* | 19 | rs117710064 | C | T | 0.865 | -0.222 | 0.045 | 8.55E-07 | rs115172026 | T | 0.859 | -0.245 | 0.047 | 2.23E-07 | 0.9931 | 0.6895 |
| *MAP3K7CL* | 21 | rs2832230 | G | T | 0.832 | 0.251 | 0.051 | 9.57E-07 | - | - | - | - | - | - | - | - |
| *MMP11* | 22 | rs2070458 | A | T | 0.219 | 0.293 | 0.038 | 7.12E-15 | rs2267038 | G | 0.192 | 0.269 | 0.042 | 1.60E-10 | 0.9793 | 0.7868 |

**Suppl Note Table 3: Meta-analysis results for multi-ancestry and European only analysis.** Chr = chromosome; SNP = single nucleotide polymorphism; EA = effect allele; EAF = effect allele frequency; SE = standard error; LD = linkage disequilibrium. Performed using a fixed-effects inverse-variance meta-analysis in GWAMA. *P* uncorrected for multiple testing.

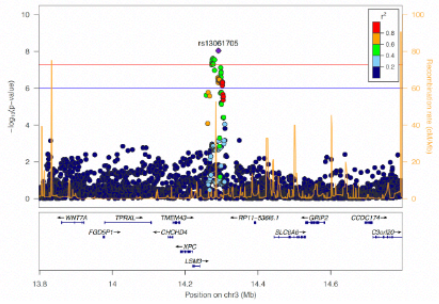
## 2. LocusZoom plots

### Genome-wide significant loci

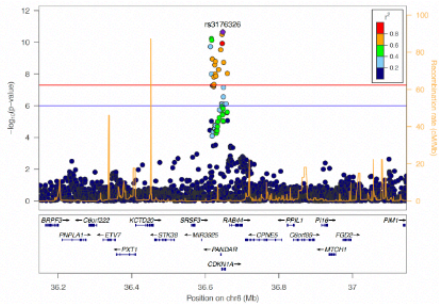
**rs1048302**



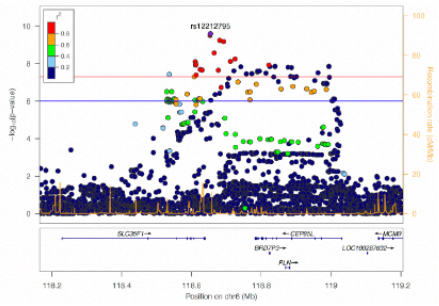
**rs13061705**



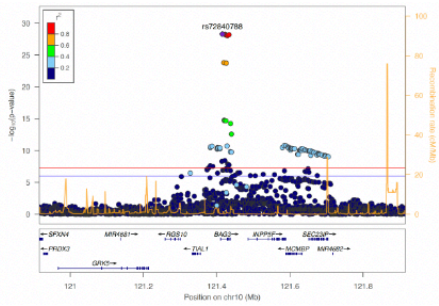
**rs3176326**



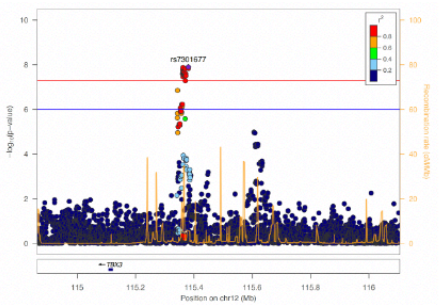
**rs12212795**



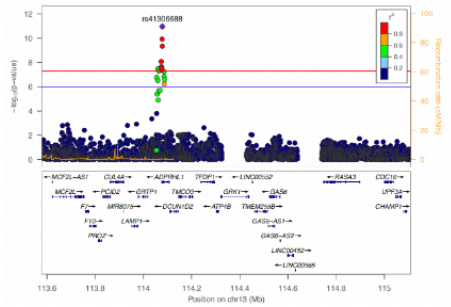
**rs72840788**



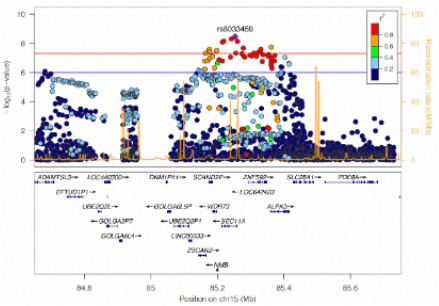
**rs7301677**



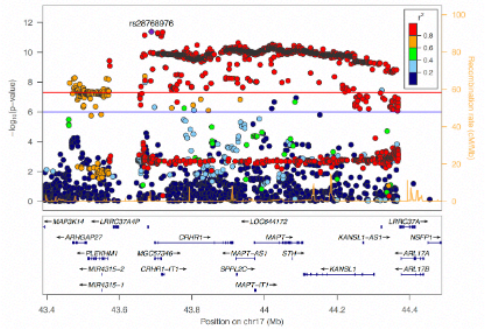
**rs41306688**



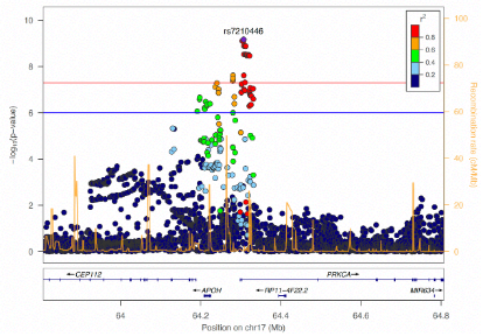
**rs8033459**



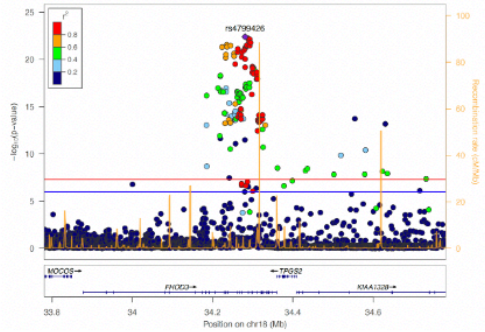
**rs28768976**



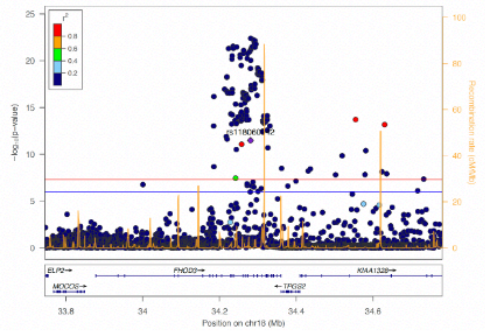
**rs7210446**



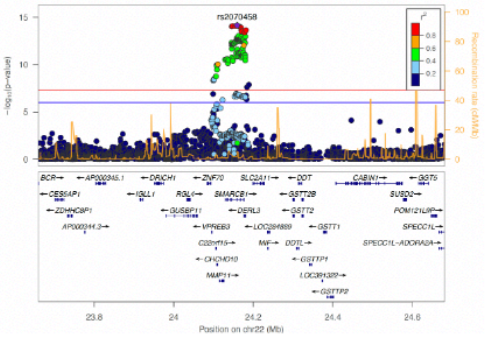
**rs4799426**



**rs118060942**

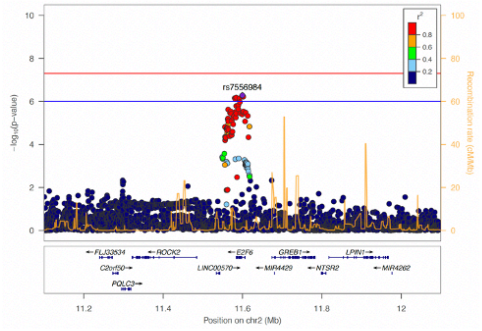


**rs2070458**

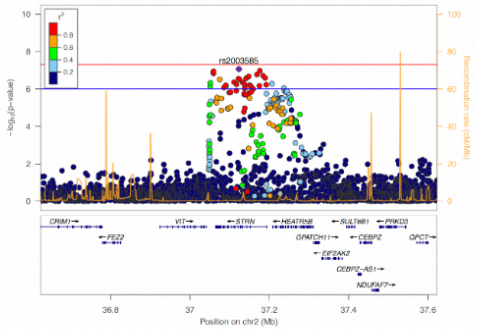


## <5% FDR loci

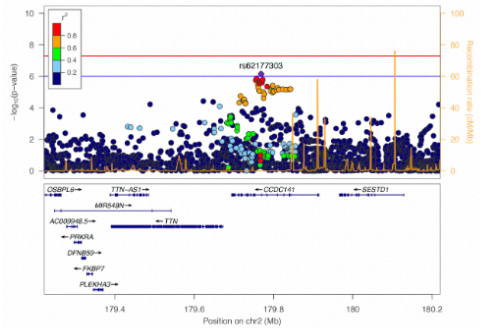
### rs7556984



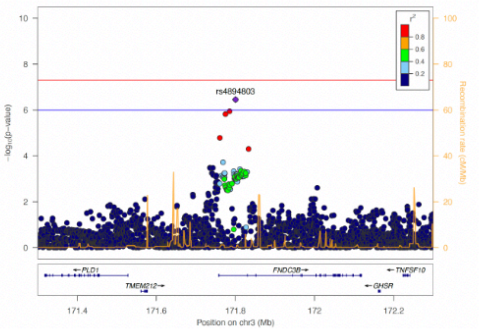

### rs2003585



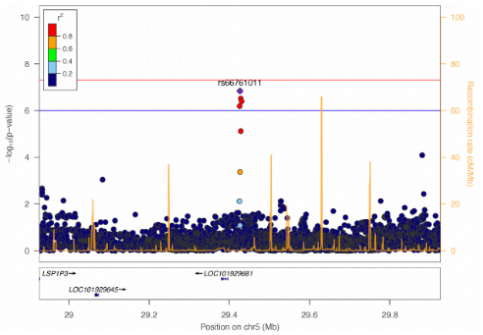

### rs62177303



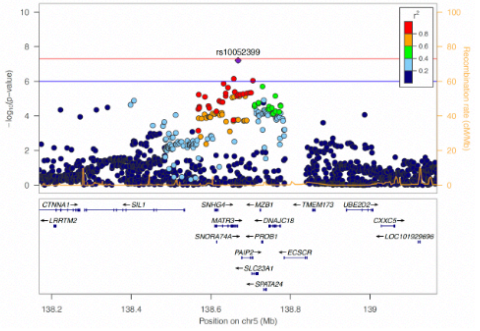

### rs4894803



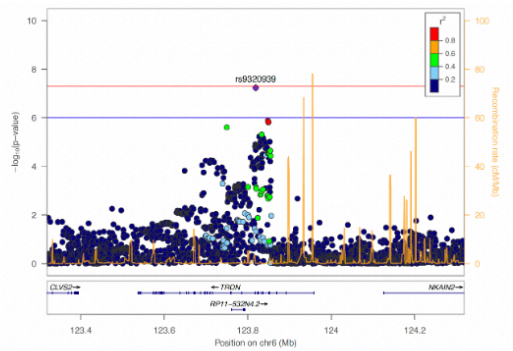

### rs66761011



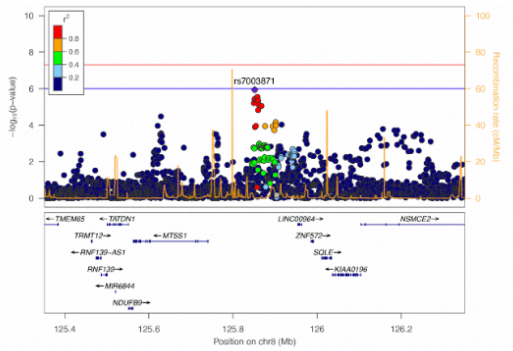

### rs10052399

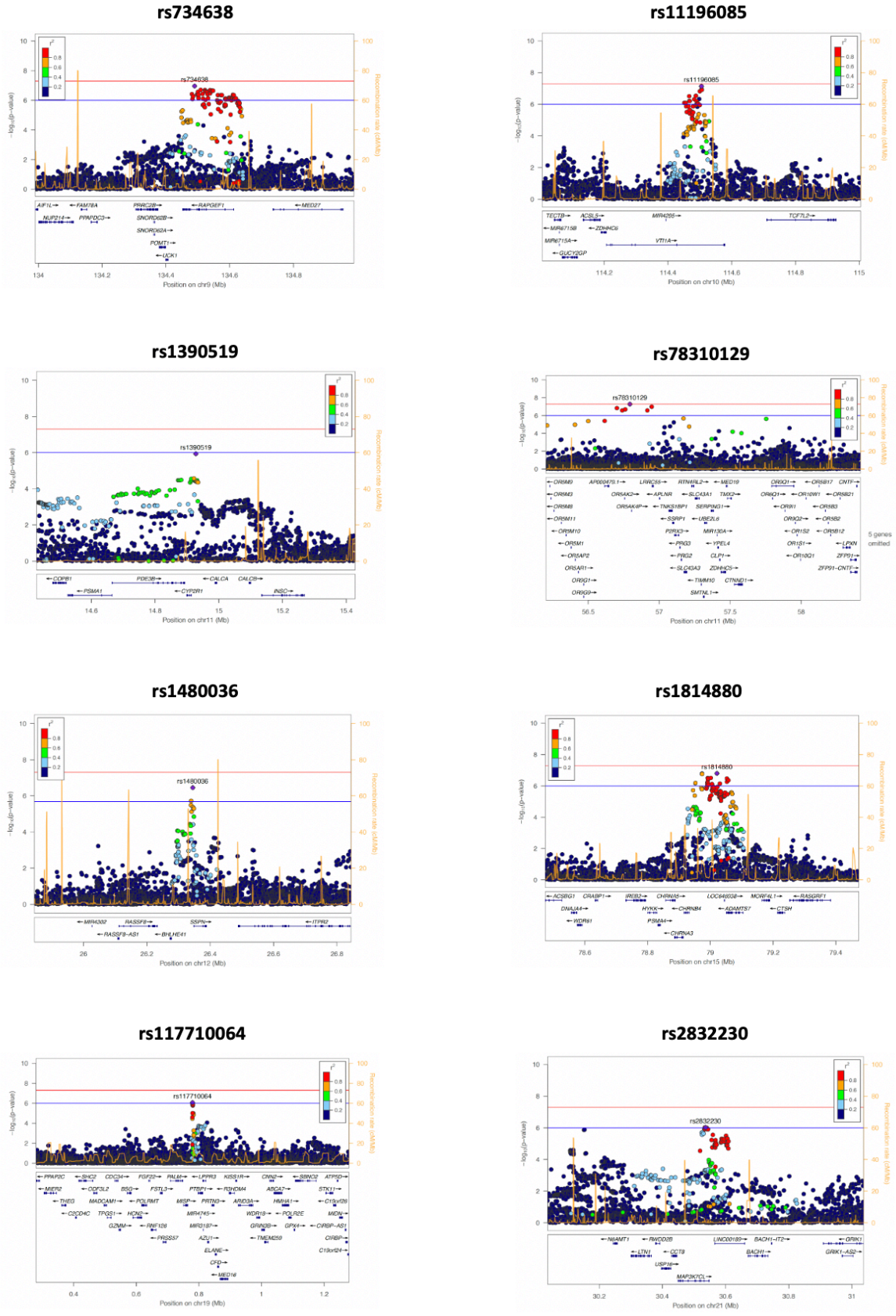



### rs9320939




### rs7003871






**Suppl Note Figure 8:** LocusZoom plots for multi-ancestry all-comer HCM meta-analysis sentinel SNPs (< 5% FDR). Red horizontal line denotes genome-wide significant threshold (p-value = $5 \times 10^{-8}$). Blue horizontal line denotes 5% FDR threshold (p-value = $1.82 \times 10^{-6}$). Performed using a fixed-effects inverse-variance meta-analysis in GWAMA. *P* uncorrected for multiple testing.

## 3. Loci description

### BAG3

An intronic variant, rs72840788 (risk allele = A; effect allele frequency = 0.21; odds ratio= 1.52 [95% CI: 1.42 – 1.64]; p-value = $5.06 \times 10^{-29}$), in the *BAG3* (BLC2-associated athanogene 3) locus, demonstrated association with HCM in the multi-ancestry meta-analysis. When conditional analysis was performed, adjusting for the genetic effects of the sentinel SNP (rs72840788), no additional independent SNPs were detected.

Evidence suggests HCM-risk, attributable to rs72840788, acts via BAG3. GTEx (v8.0) indicates BAG3 is expressed in muscle and heart tissue. Rare missense variants in *BAG3*, specifically p.Pro209Leu, have been detected in individuals diagnosed with a myofibrillar myopathy who had evidence of either hypertrophic or restrictive cardiomyopathy.[2,3]

rs72840788 is in linkage disequilibrium (LD) with a BAG3 missense variant, rs2234962 (D' = 1.00 and $r^2$ = 0.99), that demonstrates a phred-scaled Combined Annotation Dependent Depletion (CADD) score of 21.5, suggesting it confers a deleterious impact. Furthermore, whilst supportive *cis*-eQTL or promoter capture Hi-C data implicating rs72840788 with BAG3 is lacking, pQTL data indicates that rs72840788 increases BAG3 protein levels in blood plasma (beta: 0.177; p-value= $5.1\times 10^{-9}$).[4,5]

BAG3 is a member of the BAG family of anti-apoptotic proteins (BAG1-6), that bind and regulate the activity of heat shock protein 70 (Hsp70) via a BAG domain in the C-terminal. Hsp70 is a ubiquitously expressed protein that regulates the quality of protein folding. Given this, it is intriguing that rs1048302, a 3 prime UTR variant in *HSPB7,* is significantly associated with HCM (OR 1.32 [95% CI: 1.24-1.40]; p-value = $2.51\times 10^{-17}$). Within the cardiomyocyte, *BAG3* contributes towards the maintenance of sarcomeric structural integrity during mechanical stress, acts as a homeostatic regulator of filamin, and contributes towards the removal of misfolded or degraded protein products through chaperone-assisted selective autophagy (CASA)[6].

Phenome-wide association analysis for rs72840788, or SNPs in linkage disequilibrium with rs72840788, associate with myocardial biology related traits and diseases, including: heart

failure[7] (rs17617337; in LD with rs72840788 (D' = 1.00 and $r^2$ = 0.99)), dilated cardiomyopathy[8] (rs2234962; in LD with rs72840788 (D' = 1.00 and $r^2$ = 0.99)), reduced left-ventricular end systolic volume[9] (rs72840788) and increased left-ventricular ejection fraction[9] (rs72840788).

**FHOD3**

An intronic variant, rs4799426 (risk allele = G; effect allele frequency = 0.35; odds ratio= 1.38 [95% CI: 1.29 – 1.47]; p-value = $4 \times 10^{-23}$), in the *FHOD3* (formin homology 2 domain containing 3) locus, demonstrated association with HCM. Conditional analysis, that accounted for the genetic effects of the sentinel SNP (rs4799426), revealed rs118060942 an additional, independent SNP (risk allele = T; effect allele frequency = 0.012; odds ratio = 1.79 [ 95% CI: 1.45-2.20]; p-value = $2.35 \times 10^{-8}$) within the *FHOD3* locus.

*FHOD3* appears a strong candidate gene for HCM associated risk, having demonstrated evidence of co-segregation for numerous pathogenic, rare variants, and association with HCM in a previous GWAS[10,11]. Additionally, rs879568, a *FHOD3* intronic variant in LD with the sentinel SNP (rs4799426, $r^2$ = 0.90) has demonstrated an association with QRS-duration (p-value= $8.0 \times 10^{-9}$).[12]

Additionally, rs4799426 is in LD ($r^2$ = 0.84) with rs2303510, a *FHOD3* missense variant, with a CADD score of 22.8. Whilst rs4799426 does not appear to directly influence *FHOD3* expression, SNPs in LD with rs4799426 (rs1495900, $r^2$ = 1.0, GWAS p-value = $2.84 \times 10^{-22}$) do appear to influence *FHOD3* in heart atrial appendage tissue (normalised effect size = -0.15; p-value = $4.2 \times 10^{-8}$).

GTEx (v8.0) data indicates *FHOD3* is highly expressed in cardiac tissue. *FHOD3* contributes towards the organisation of actin and in the maintenance of cardiac function.[13] Functional studies in mice suggest *FHOD3* directly interacts with *MYBPC3*, a mechanism that appears to be important for the regulation of cardiac function.[14]

**HSPB7**

rs1048302, a 3' UTR variant in *HSPB7* (heat shock protein family B (small) member 7) demonstrated association with HCM (risk allele = T; effect allele frequency = 0.33; odds ratio = 1.32 [95% CI: 1.24 – 1.40]; p-value = $2.51 \times 10^{-17}$). Conditional analysis, accounting for the genetic effects of the sentinel SNP (rs1048302), revealed no additional independent SNPs. *HSPB7* is highly expressed in cardiac and muscle tissue, and SNPs in LD with rs1048302 appear to influence HSPB7 expression in atrial appendage tissue.

rs1048302 is in LD with SNPs associated with DCM (rs10927875, $r^2 = 0.80$; p-value = $1.0 \times 10^{-9}$) and there is suggestive evidence supporting an association with heart failure.[7,8,15]

HSPB7 has a critical role in cardiac development, with HSPB7 knockout mouse models demonstrating embryonic lethality. HSPB7 appears to have several roles in maintaining muscle integrity, with roles in modulating the actin thin filament length and consequently suppress actin polymerisation, but also through an interaction with dimerized FLNC.[16,17]

**ADPRHL1**

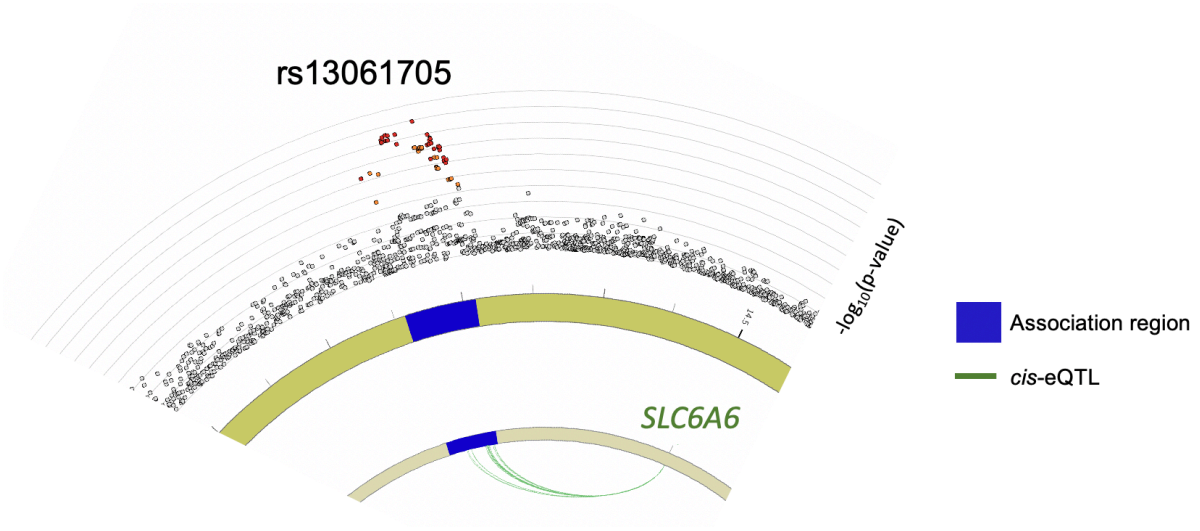
rs41306688, a missense variant with a CADD score of 27.1 in *ADPRHL1,* (ADP-ribosylhydrolase like 1) was proven to be associated with HCM (risk allele = C; effect allele frequency = 0.03; odds ratio = 1.82 [95% CI: 1.53 – 2.17]; p-value = $1.06 \times 10^{-11}$). In a meta-analysis of 22 studies, rs41306688 has been shown to prolong PR interval (beta = 0.10; standard error=0.02; p-value = $7.4 \times 10^{-9}$).[18]

*ADPRHL1* appears highly expressed in cardiac and muscle tissue. Functional studies, performed in *Xenopus,* suggest *ADPRHL1* has a critical role in modifying Z-disc and actin dynamics during cardiac development.[19]

**SLC6A6**

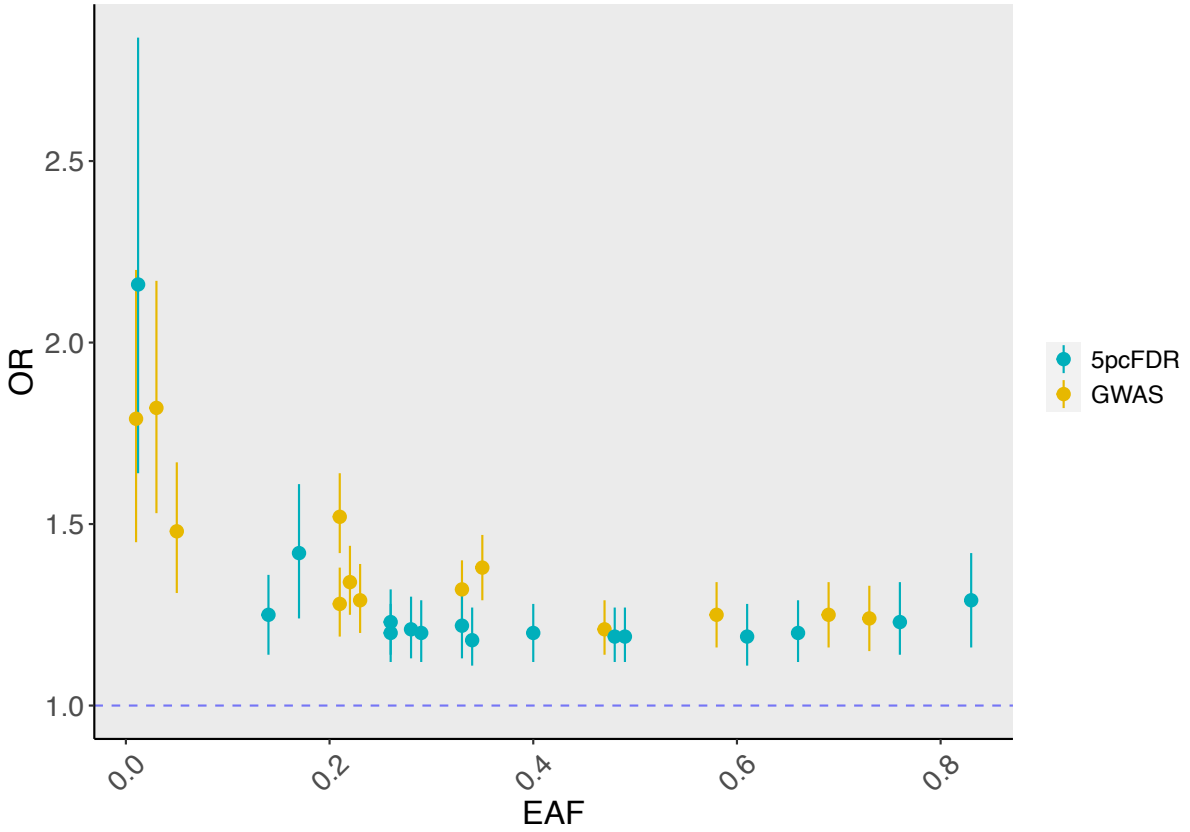
rs13061705 is associated with risk of HCM (risk allele = C; effect allele frequency = 0.69; odds ratio = 1.25 [95% CI: 1.16 – 1.34]; p-value = $9.06 \times 10^{-9}$) and is located in an inter-genic region. *Cis*-eQTL analysis from heart atrial appendage tissue in GTEx (v8) indicates

that SNPs in LD with rs13061705 influence the expression of *SLC6A6* (Supplementary Note Figure 5). For example, rs62231954 (rs13061705, $r^2 = 0.53$) increases *SLC6A6* in heart atrial appendage tissue (normalised effect size = 0.18, p-value = $8.5 \times 10^{-5}$). *SLC6A6* encodes a taurine transporter that appears ubiquitously expressed. Numerous animal studies have demonstrated that taurine deficiency leads to a dilated cardiomyopathy phenotype.[20–23]



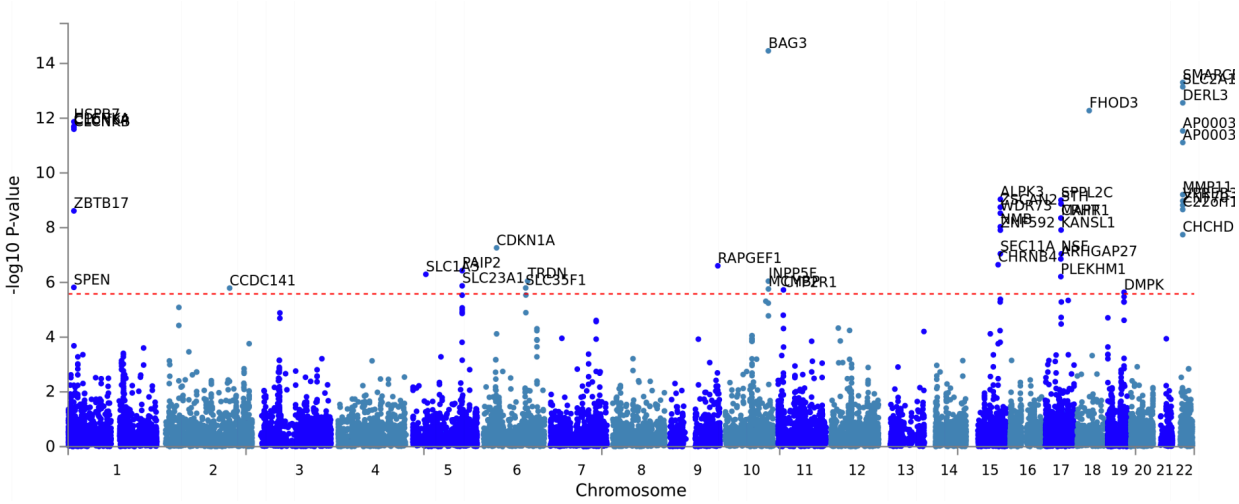

**Suppl Note Figure 9: cis-eQTL analysis performed in FUMA** (v1.3.5e) (https://fuma.ctglab.nl/snp2gene) using GTEx (v8.0) expression data from heart tissue (atrial appendage and left ventricle). For a sentinel SNP, rs13061705, there is evidence of a *cis*-eQTL with *SLC6A6* (rs13061705, $r^2 = 0.53$) in atrial appendage tissue.

4. Effect size to allele frequency relationship

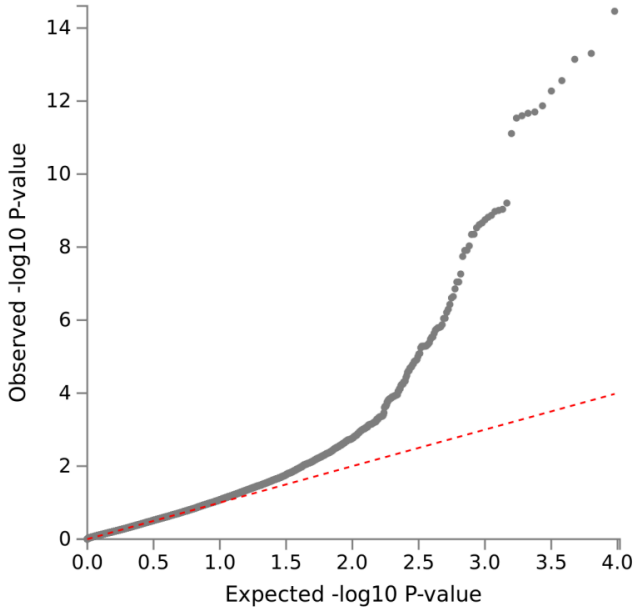



**Suppl Note Figure 10: Relationship between effect allele frequency (EAF) and odds ratio (OR) for sentinel variants.** Data derived from the multi-ancestry all-comer meta-analysis (2,780 HCM cases vs. 47,486 controls). Error bars reflect 95% confidence intervals.

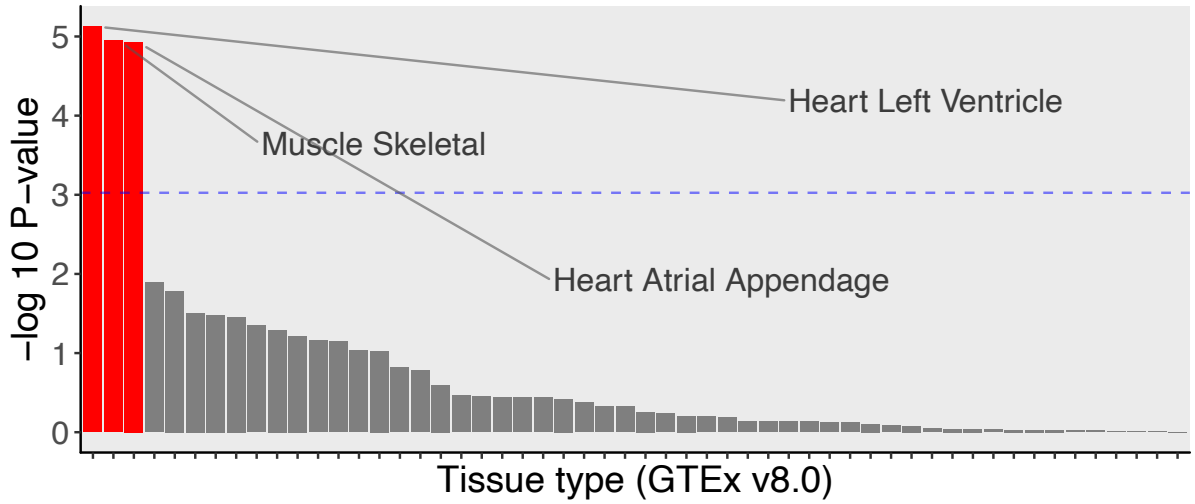
5. Gene-based analysis




**Suppl Note Figure 11: Manhattan plot reporting gene-level analysis performed in FUMA** (v1.3.5e) (https://fuma.ctglab.nl/snp2gene). Significance threshold denoted by red dotted horizontal line ((α = 0.05/19026(total number of genes tested) = $2.63 \times 10^{-6}$). Data generated using summary statistics from multi-ancestry all-comer HCM meta-analysis. MAGMA uses gene windows that include SNPs 35 kb upstream and 10 Kb downstream. Gene symbols annotated as per Ensembl version 92. *P* uncorrected for multiple testing.

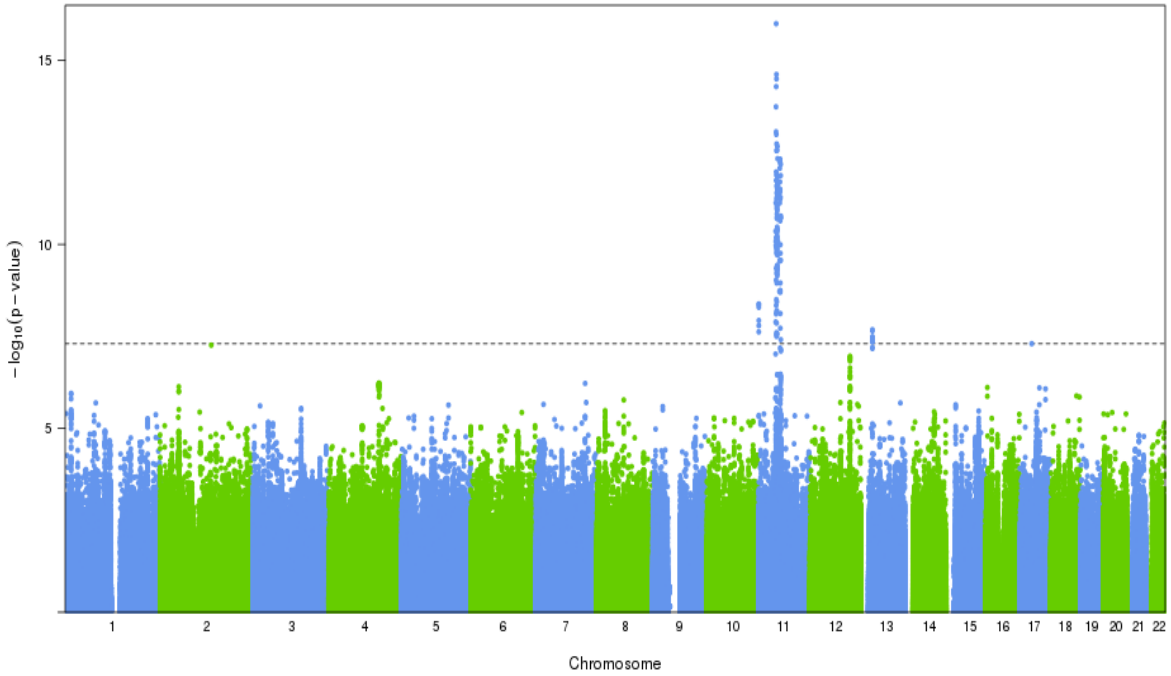



**Suppl Note Figure 12: QQ plot associated with gene-level analysis performed in FUMA** (v1.3.5e) (https://fuma.ctglab.nl/snp2gene) as presented in Suppl Note Figure 11. *P* uncorrected for multiple testing.

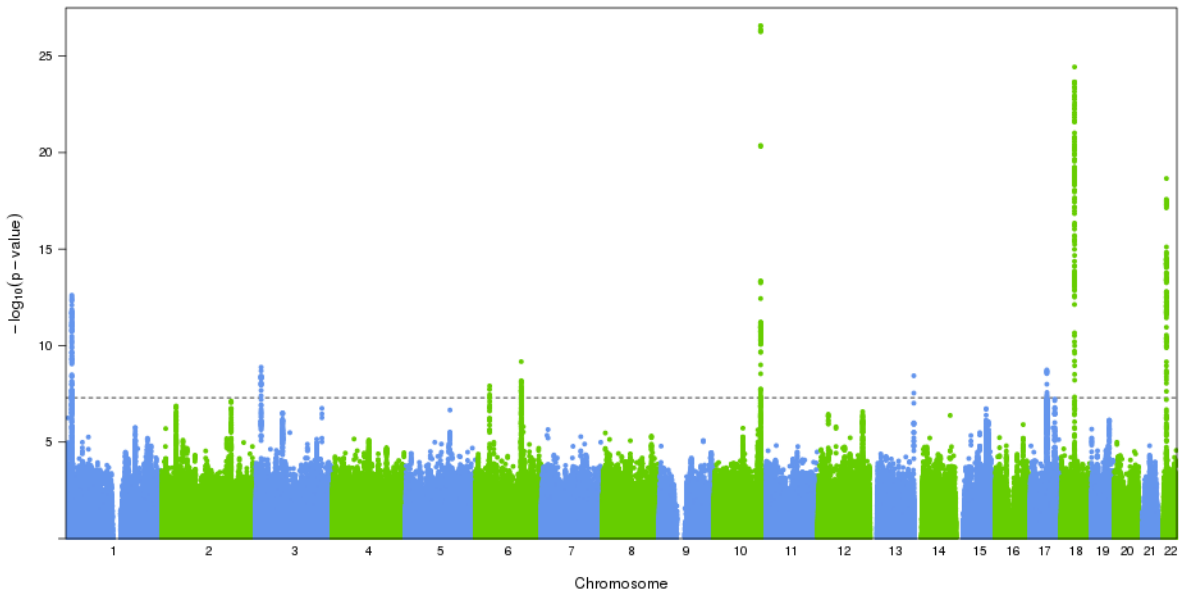



**Suppl Note Figure 13: Evidence of tissue enrichment in heart and muscle tissue.** Data derived from gene-level analysis performed in FUMA (v1.3.5e) (https://fuma.ctglab.nl/snp2gene) as presented in Suppl Note Figure 7. Tissue expression data obtained from GTEx (v8.0). *P* uncorrected for multiple testing.

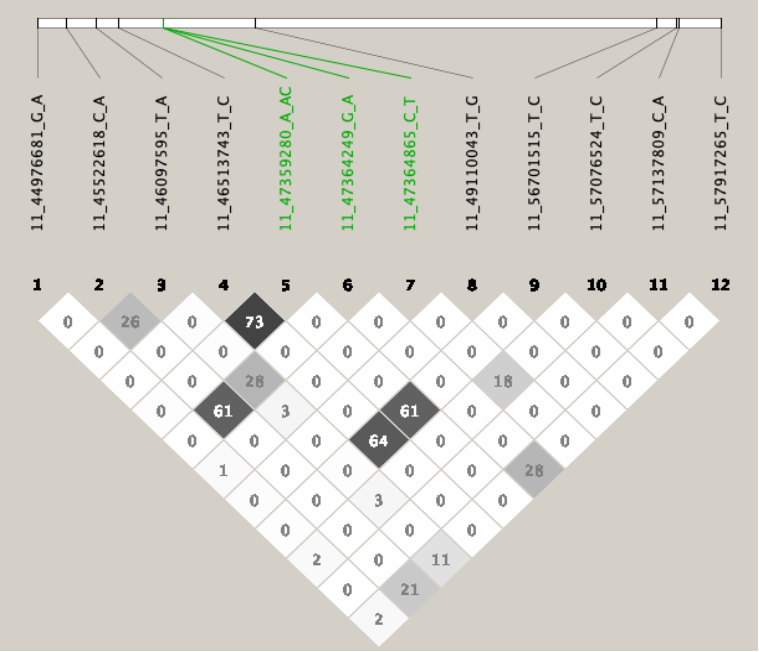
6. Sarcomere carrier status GWAS




**Suppl Note Figure 14: Sarcomere positive HCMR vs. UKBB GWAS results represented using a Manhattan plot.** Manhattan plot split by chromosome (x-axis) and -$\log_{10}$(p-value) (y-axis). Dotted line represents genome-wide significance threshold (p-value = $5\times10^{-8}$). Conditional and joint analysis performed using GCTA-cojo identified 12 loci, including a locus encompassing *MYBPC3* on chromosome 11 (Suppl Table S2). Excluding chromosome 11, due to long-range LD with pathogenic *MYBPC3* variants (Suppl Note Figure 16 and Suppl Table 4), results in 2 independent SNPs at genome-wide significance and 9 independent SNPs beneath the 5% FDR threshold (p-value = $1.59 \times 10^{-6}$). Analysis performed using logistic regression to fit an additive case-control association model using SNPTEST v.2.5.4-beta3. *P* uncorrected for multiple testing.




**Suppl Note Figure 15: Sarcomere negative HCM meta-analysis results represented using a Manhattan plot.** Manhattan plot split by chromosome (x-axis) and -$\log_{10}$(p-value) (y-axis). Dotted line represents genome-wide significance threshold (p-value = $5\times10^{-8}$). The sarcomere negative HCMR vs. UKBB GWAS was meta-analysed with the BRRD vs. BRRD GWAS using GWAMA. The genomic control was calculated ($\lambda$= 1.047) and the 5% FDR (p-value of $1.56 \times 10^{-6}$) calculated. Using GCTA-cojo, 10 independent SNPs at genome-wide significance were identified alongside 15 independent SNPs beneath the 5% FDR threshold. Analysis performed using logistic regression to fit an additive case-control association model using SNPTEST v.2.5.4-beta3. *P* uncorrected for multiple testing.

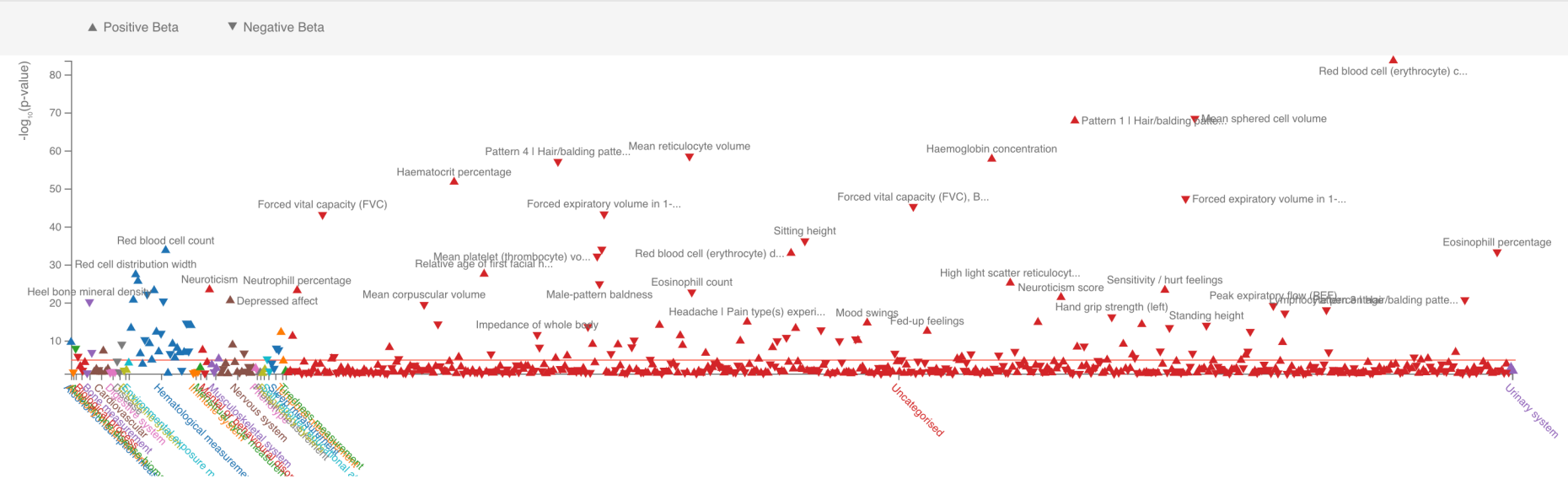



**Suppl Note Figure 16: Evaluation of long-range linkage disequilibrium in sarcomere-positive analysis.** Evidence of long-range linkage disequilibrium between common imputed genotypes (black) and rare pathogenic variants (green) in a causal HCM gene (*MYBPC3*). Represented here, 11_47359280_A_AC corresponds to MYBPC3 p.Trp792ValfsTer41; 11_47364249_G_A corresponds to *MYBPC3* p.R502W and 11_47364865_C_T corresponds to *MYBPC3* c.1224-52G>A. Image generated using Haploview (version 4.2), with grey/white numbers reflective of $r^2$ between markers.

| Pathogenic / imputed variant | $r^2$ | Pathogenic beta estimate (standard error) / imputed beta estimate (standard error) |
|---|---|---|
| *MYBPC3* p.R502W / 11_45522618_C_A | 0.61 | 7.56 (1.49) / -0.455 (0.398) |
| *MYBPC3* p.Trp792ValfsTer41 / 11_56701515_T_C | 0.61 | 6.82 (1.48) / 0.214 ( 0.402) |
| *MYBPC3* p.Trp792ValfsTer41 / 11_46513743_T_C | 0.73 | 7.66 (1.48) / -0.344 (0.331) |

**Suppl Note Table 4:** Modelling the impact of both rare and common variants with multiple logistic regression. HCM cases were limited to those individuals in whom any pathogenic or likely pathogenic variant in a core sarcomere gene had previously been identified (n= 851). Controls were drawn from the UK Biobank, (n=19,851). Analysis confirmed that HCM risk could be entirely attributed to the rare variants, indicating the presence of synthetic association.

7. Pleiotropy assessment




**Suppl Note Figure 17:** Extreme and diverse pleiotropy observed for rs28768976 with wide range of uncorrelated phenotypes, as curated from publicly available genome wide association studies. Data derived from https://genetics.opentargets.org/variant/17_45610951_A_G. *P* uncorrected for multiple testing.

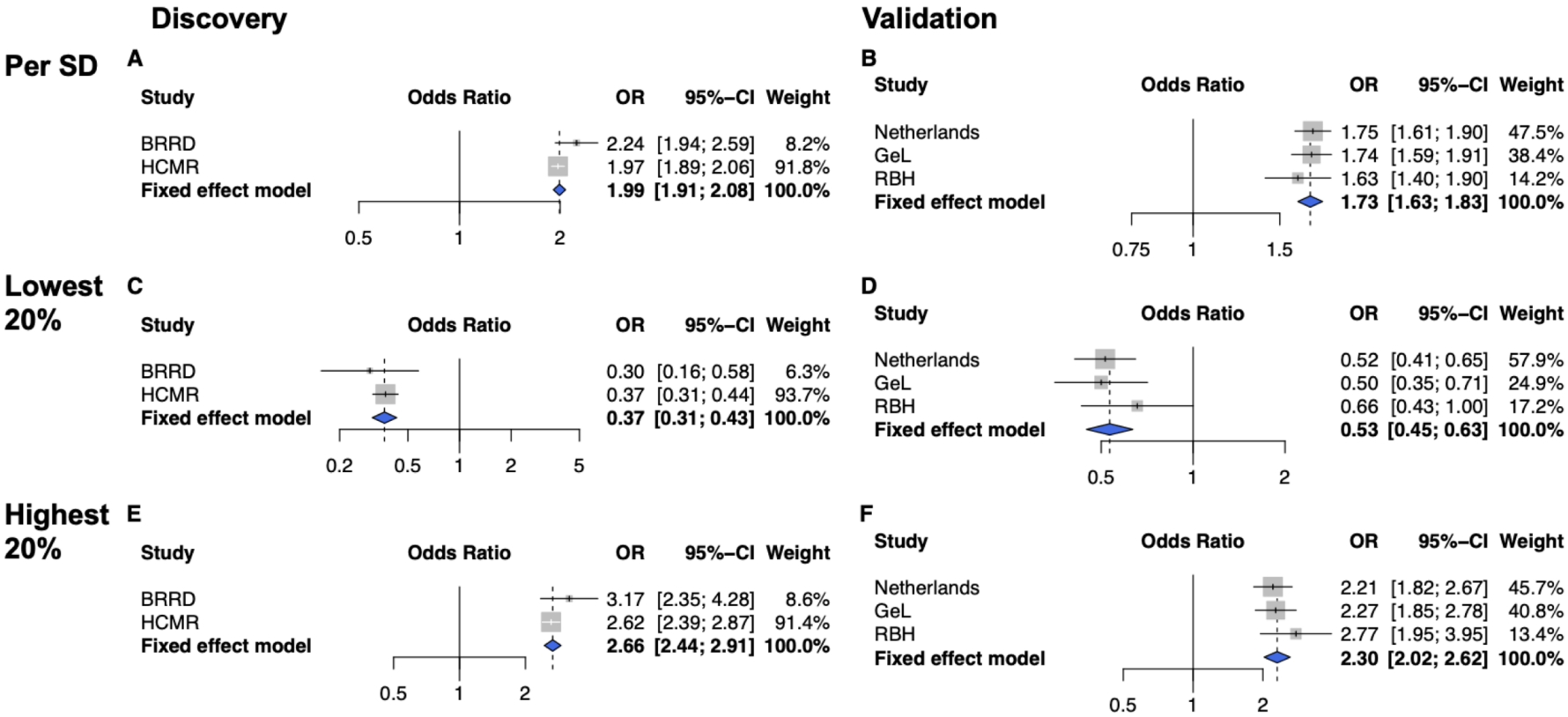
8. **Validation cohorts**

**Genomics England:** Access to GeL was provided via the Cardiovascular GeCIP, via Registry ID RR254. GeL sequenced 101,162 genomes from 90,643 individuals recruited via 13 Genomic Medicine Centres across the United Kingdom. Analyses were performed on a subset of the total cohort, specifically 38,344 distantly related individuals, with genome sequence data mapped to GRCh38 that passed quality control criteria (including: ≥ 250 bp insert size, ≥ 75% mapped reads, < 2% chimeric DNA fragments and <5% cross contamination). Phenotypic information was provided via hospital episode statistics and clinician entered human phenotype ontology terms. 435 HCM probands were available for analysis. 64 individuals were sarcomere-positive, based on the presence of a variant of uncertain significance, likely pathogenic or pathogenic variant.

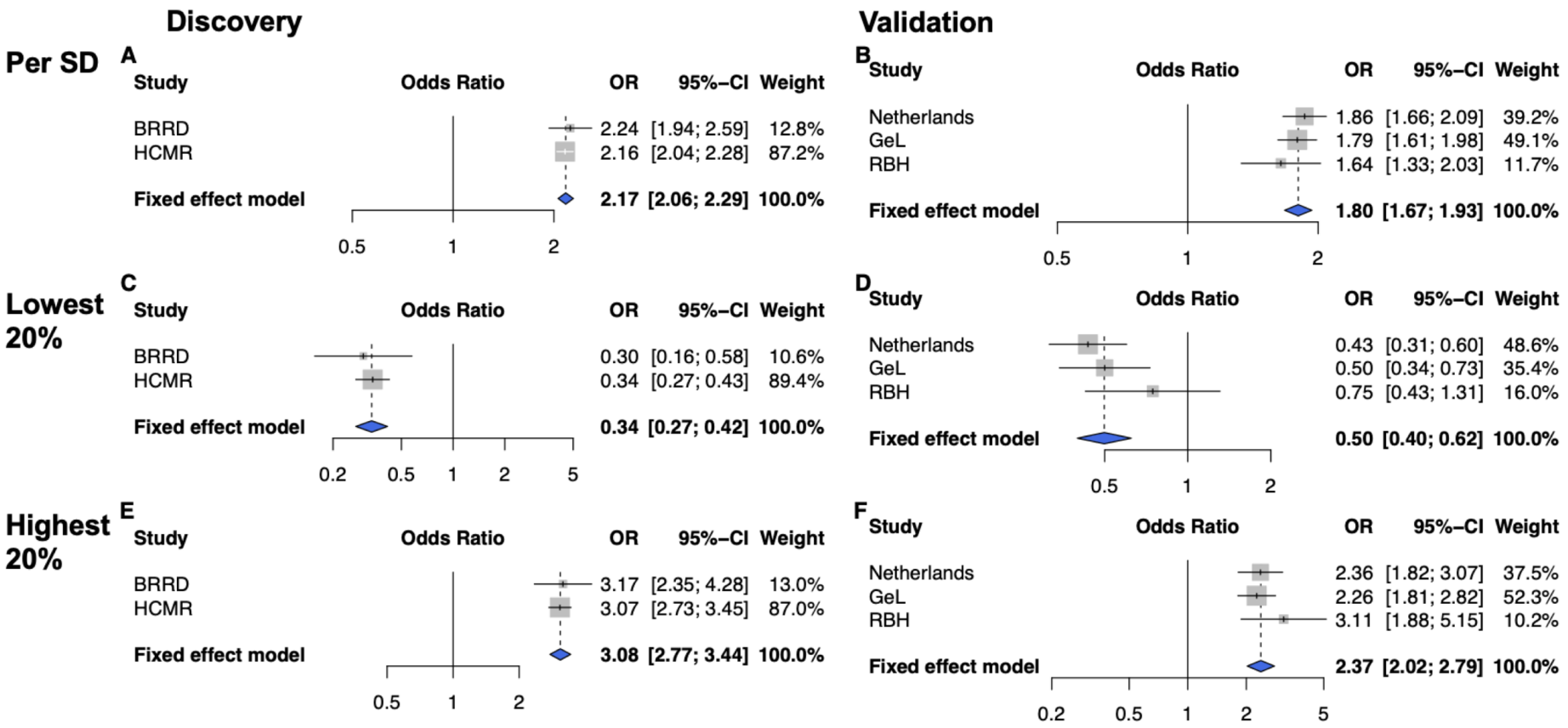
**Netherlands:** 999 cases were identified using current diagnostic criteria (left ventricular wall thickness ≥15mm or ≥13mm in presence of family history) from cardiovascular genetics referral centres in the Netherlands (Amsterdam University Medical Center, Erasmus Medical Center and the University Medical Center Groningen). Samples overlapping with the HCMR cohort were removed leaving 975 cases. 2117 controls were derived from a population cohort study from the Netherlands.[24] Genotyping was performed using on an Illumina Infinium BeadChip using an Illumina OmniExpress and Global Screening Array. SNPs were mapped to GRCh37, and removed if: missingness rate > 5%, Hardy-Weinberg equilibrium test $P<10^{-6}$ for controls or $P<10^{-10}$ for cases, or MAF <0.05. Individuals were excluded when: missingness exceeded 3%, inbreeding coefficient ≥0.1, genotype-phenotype sex mismatch existed, proportional identity by descent >0.05, or non-European ancestry was indicated by principal components analysis. Phasing (Eagle2) and imputation (Haplotype reference consortium (HRCr1.1) panel) was performed on the Michigan Imputation Server v.1.0.2. SNPs with MAF>0.01 and a Minimac $R^2$>0.5 were retained.

**Royal Brompton Hospital:** 410 HCM cases were recruited from the Royal Brompton & Harefield Hospitals NHS Trust Cardiovascular Research Biobank. Samples overlapping with either the HCMR or BRRD cohorts were removed leaving 359 cases. 1211 controls, screened for evidence of HCM using cardiac imaging, were recruited from the UK Digital Heart Project.[25] Genotyping was performed using the Illumina Human OmniExpress beadchip. SNPs were mapped to GRCh37 and excluded if MAF<0.01, Hardy-Weinberg equilibrium test $P < 10^{-7}$, or missingness rate >0.05. Sample QC excluded samples with a genotype-phenotype sex mismatch, heterozygosity rate >3 standard deviations from the mean, missingness rate >0.03 or evidence of non-European ancestry via principal components. Genotypes were phased using SHAPEIT (v2.r790) and imputed using IMPUTE2 (v2.3.2), against the UK10K and 1000 Genomes Project reference panel. SNPs with MAF>0.01 and INFO score >0.4 were retained.

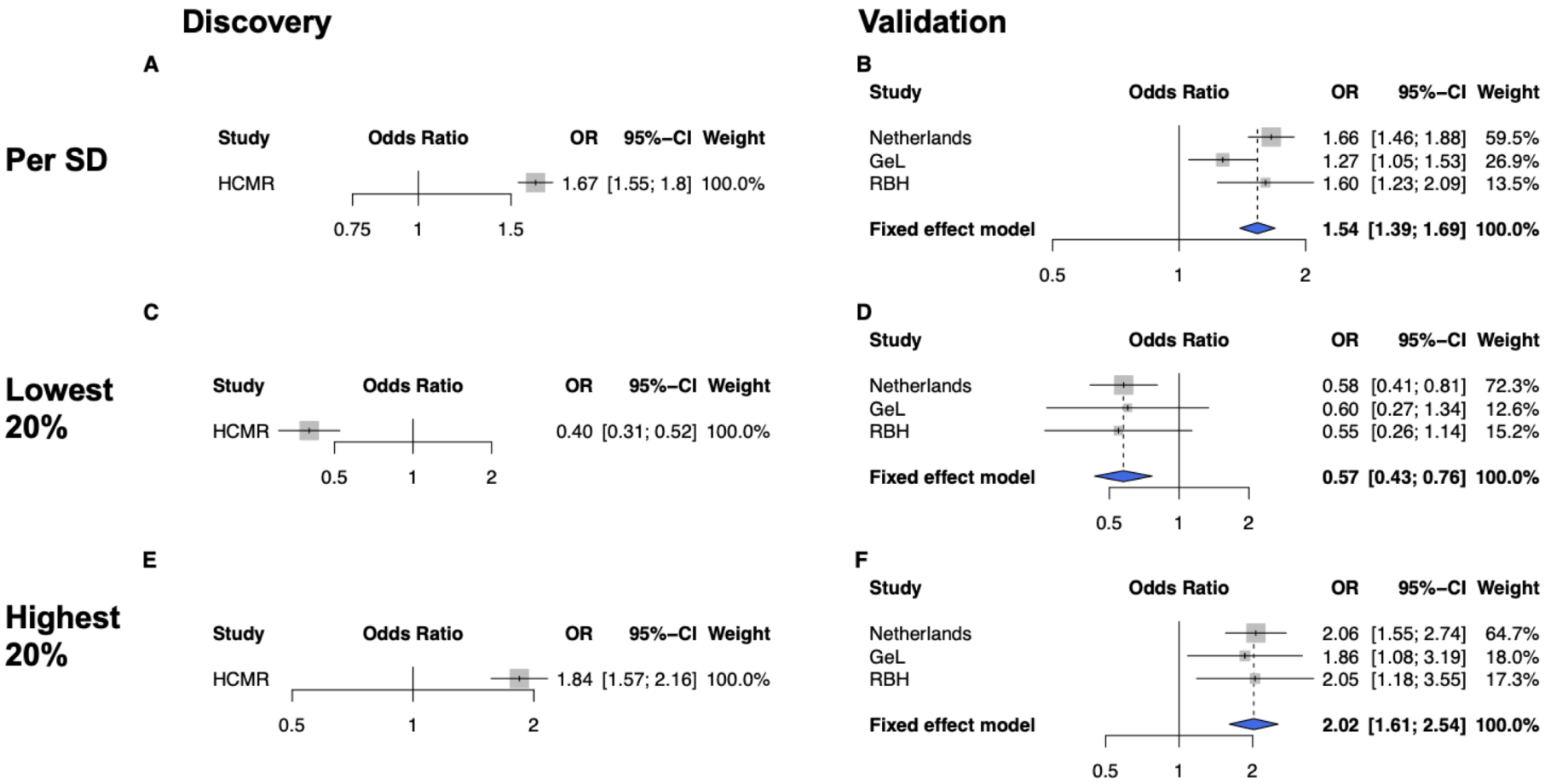
**9.** Genetic risk score analysis




**Suppl. Note Figure 18: All comer HCM analysis.** HCMR: HCM registry; GeL: Genomics England 100,000 Genomes; RBH: Royal Brompton Hospital cohort; Netherlands: cases recruited from Amsterdam Medical Center, Rotterdam and Groningen. Panels A, C and E present analysis specific to discovery cohorts. Panels B, D and F present analysis specific to validation cohorts. Panels A and B present per standard deviation analysis. Panels C and D present effect sizes, corresponding to the bottom 20% of the population compared with the middle 60% of the population using a quintile-based analysis. Panels E and F present effect sizes, corresponding to the top 20% of the population compared with the middle 60% of the population using a quintile-based analysis.

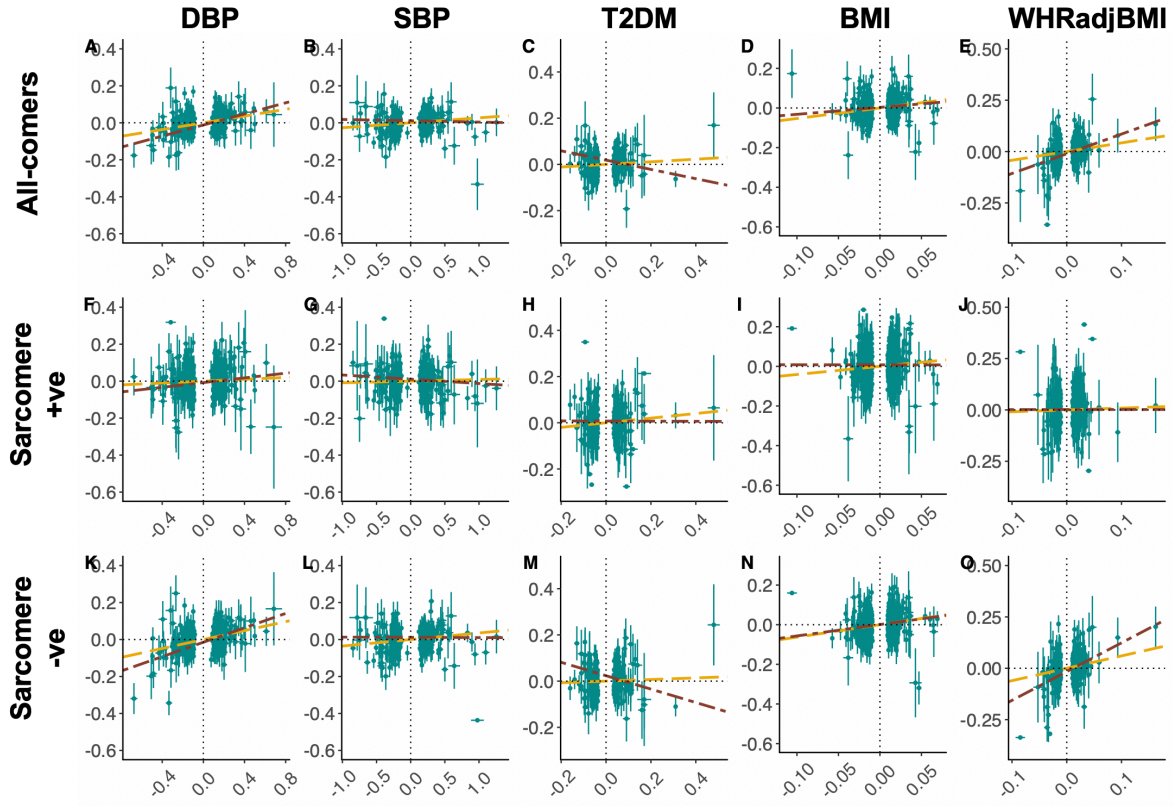



**Suppl. Note Figure 19: Sarcomere negative HCM.** HCMR: HCM registry; GeL: Genomics England 100,000 Genomes; RBH: Royal Brompton Hospital cohort; Netherlands: cases recruited from Amsterdam Medical Center, Rotterdam and Groningen. Panels A, C and E present analysis specific to discovery cohorts. Panels B, D and F present analysis specific to validation cohorts. Panels A and B present per standard deviation analysis. Panels C and D present effect sizes, corresponding to the bottom 20% of the population compared with the middle 60% of the population using a quintile-based analysis. Panels E and F present effect sizes, corresponding to the top 20% of the population compared with the middle 60% of the population using a quintile-based analysis.

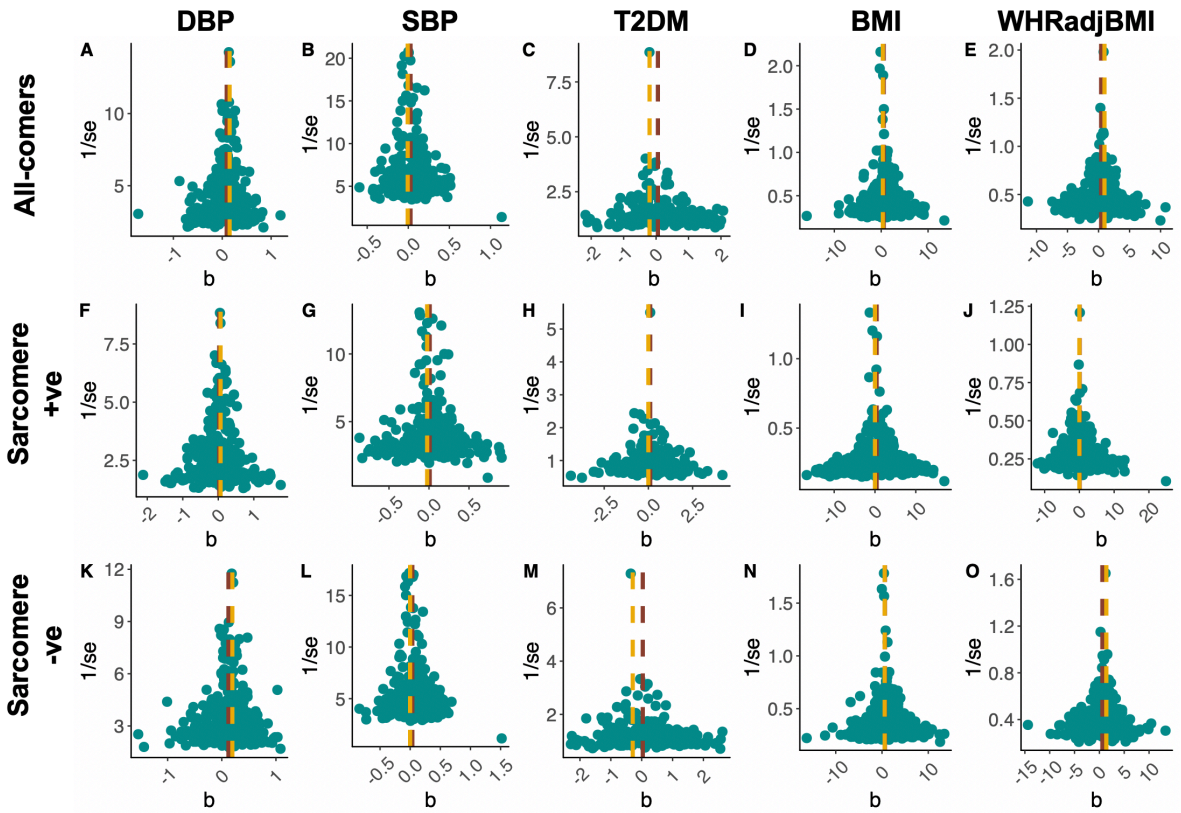



**Suppl. Note Figure 20: Sarcomere positive HCM.** HCMR: HCM registry; GeL: Genomics England 100,000 Genomes; RBH: Royal Brompton Hospital cohort; Netherlands: cases recruited from Amsterdam Medical Center, Rotterdam and Groningen. Panels A, C and E present analysis specific to discovery cohorts. Panels B, D and F present analysis specific to validation cohorts. Panels A and B present per standard deviation analysis. Panels C and D present effect sizes, corresponding to the bottom 20% of the population compared with the middle 60% of the population using a quintile-based analysis. Panels E and F present effect sizes, corresponding to the top 20% of the population compared with the middle 60% of the population using a quintile-based analysis.

## 10. Two sample Mendelian randomisation




**Suppl. Note Figure 21: Scatter plots for two-sample Mendelian randomisation studies.** Graphs are grouped according the exposure and outcome phenotypes. For each graph, the x-axis reports beta estimate for exposure phenotype (DBP: diastolic blood pressure; SBP: systolic blood pressure; T2DM: Type 2 diabetes; BMI: body mass index; WHRadjBMI: Waist-hip ratio adjusted for BMI) and y-axis reports beta estimates for the outcome phenotype (all-comer HCM, sarcomere-positive HCM, or sarcomere-negative HCM). The yellow dashed line is representative of inverse variance weighted (IVW) regression line and red dash-dot line represents the Egger regression line. The turquoise data points reflect the instrumental variables used for each analysis. A full summary of the Mendelian randomisation studies is available in Supplementary Table S19.




**Suppl. Note Figure 22: Funnel plot for two sample Mendelian randomisation analysis.** Graphs are grouped according the exposure (DBP: diastolic blood pressure; SBP: systolic blood pressure; T2DM: Type 2 diabetes; BMI: body mass index; WHRadjBMI: Waist-hip ratio adjusted for BMI) and outcome phenotypes (all-comer HCM, sarcomere-positive HCM, or sarcomere-negative HCM). For each graph, the x-axis reports beta estimate for SNPs included as an instrumental variable. Asymmetry may be indicative of horizontal pleiotropy. The y-axis reports the SNP's precision (1/standard error). The yellow dashed line is representative of inverse variance weighted (IVW) regression line and red dash-dot line represents the Egger regression line. The turquoise data points reflect the instrumental variables used for each analysis. A full summary of the Mendelian randomisation studies is available in Supplementary Table S19.

11. Consortia author list:

**Genomics England Research Consortium** (https://public.huddle.com/a/kzWvRoL/index.html)**:**
Ambrose J. C.[1] , Arumugam P.[1] , Baple E. L.[1] , Bleda M.[1] , Boardman-Pretty F.[1,2] , Boissiere J. M.[1], Boustred C. R.[1] , Brittain H. [1], Caulfield M. J. [1,2], Chan G. C. [1], Craig C. E. H. [1], Daugherty L. C. [1], de Burca A. [1], Devereau, A. [1], Elgar G. [1,2], Foulger R. E. [1], Fowler T. [1], Furió-Tarí P. [1], Hackett J. M. [1], Halai D. [1], Hamblin A. [1], Henderson S. [1,2], Holman J. E. [1], Hubbard T. J. P. [1], Ibáñez K. [1,2], Jackson R. [1], Jones L. J. [1,2], Kasperaviciute D. [1,2], Kayikci M. [1], Lahnstein L. [1], Lawson K. [1], Leigh S. E. A. [1], Leong I. U. S. [1], Lopez F. J. [1], Maleady-Crowe F. [1], Mason J. [1], McDonagh E. M. [1,2], Moutsianas L. [1,2], Mueller M. [1,2], Murugaesu N. [1], Need A. C. [1,2], Odhams C. A. [1], Patch C. [1,2], Perez-Gil D. [1], Polychronopoulos D. [1], Pullinger J. [1], Rahim T. [1], Rendon A. [1], Riesgo-Ferreiro P. [1], Rogers T. [1], Ryten M. [1], Savage K. [1], Sawant K.[1], Scott R. H. [1], Siddiq A. [1], Sieghart A. [1], Smedley D. [1,2], Smith K. R. [1,2], Sosinsky A. [1,2], Spooner W. [1], Stevens H. E. [1], Stuckey A. [1], Sultana R. [1], Thomas E. R. A. [1,2], Thompson S. R. [1], Tregidgo C. [1], Tucci A. [1,2], Walsh E. [1], Watters, S. A. [1], Welland M. J. [1], Williams E. [1], Witkowska K. [1,2], Wood S. M. [1,2], Zarowiecki M.[1].

1. Genomics England, London, UK
2. William Harvey Research Institute, Queen Mary University of London, London, EC1M 6BQ, UK.

**Supplementary references**


1. Grace, C., Farrall, M., Watkins, H. & Goel, A. Manhattan++: Displaying genome-wide association summary statistics with multiple annotation layers. *BMC Bioinformatics* (2019) doi:10.1186/s12859-019-3201-y.
2. Odgerel, Z. *et al.* Inheritance patterns and phenotypic features of myofibrillar myopathy associated with a BAG3 mutation. *Neuromuscul. Disord.* **20**, 438–442 (2010).
3. Selcen, D. *et al.* Mutation in BAG3 causes severe dominant childhood muscular dystrophy. *Ann. Neurol.* **65**, 83–89 (2009).
4. Sun, B. B. *et al.* Genomic atlas of the human plasma proteome. *Nature* **558**, 73–79 (2018).
5. 10_119656173_G_A | Open Targets Genetics. https://genetics.opentargets.org/variant/10_119656173_G_A.
6. Knezevic, T. *et al.* BAG3: a new player in the heart failure paradigm. *Heart Fail. Rev.* **20**, 423–434 (2015).
7. Shah, S. *et al.* Genome-wide association study provides new insights into the genetic architecture and pathogenesis of heart failure. *bioRxiv* (2019) doi:10.1101/682013.
8. Villard, E. *et al.* A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. *Eur. Heart J.* (2011) doi:10.1093/eurheartj/ehr105.
9. Aung, N. *et al.* Genome-Wide Analysis of Left Ventricular Image-Derived Phenotypes Identifies Fourteen Loci Associated with Cardiac Morphogenesis and Heart Failure Development. *Circulation* (2019) doi:10.1161/circulationaha.119.041161.
10. Wooten, E. C. *et al.* Formin homology 2 domain containing 3 variants associated with hypertrophic cardiomyopathy. *Circulation: Cardiovascular Genetics* vol. 6 10–18 (2013).
11. Ochoa, J. P. *et al.* Formin Homology 2 Domain Containing 3 (FHOD3) Is a Genetic Basis for Hypertrophic Cardiomyopathy. *J. Am. Coll. Cardiol.* **72**, 2457–2467 (2018).
12. van der Harst, P. *et al.* 52 Genetic Loci Influencing Myocardial Mass. *J. Am. Coll. Cardiol.* **68**, 1435–1448 (2016).
13. Ushijima, T. *et al.* The actin-organizing formin protein Fhod3 is required for postnatal development and functional maintenance of the adult heart in mice. *J. Biol. Chem.* **293**, 148–162 (2018).
14. Matsuyama, S. *et al.* Interaction between cardiac myosin-binding protein C and formin Fhod3. *Proc. Natl. Acad. Sci. U. S. A.* **115**, E4386–E4395 (2018).
15. Cappola, T. P. *et al.* Loss-of-function DNA sequence variant in the CLCNKA chloride channel implicates the cardio-renal axis in interindividual heart failure risk variation. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 2456–2461 (2011).
16. Juo, L. Y. *et al.* HSPB7 interacts with dimerized FLNC and its absence results in progressive myopathy in skeletal muscles. *J. Cell Sci.* **129**, 1661–1670 (2016).
17. Wu, T. *et al.* HSPB7 is indispensable for heart development by modulating actin filament assembly. *Proc. Natl. Acad. Sci. U. S. A.* **114**, 11956–11961 (2017).
18. Lin, H. *et al.* Common and Rare Coding Genetic Variation Underlying the Electrocardiographic PR Interval. *Circ. Genomic Precis. Med.* **11**, e002037 (2018).
19. Smith, S. J. *et al.* The cardiac-restricted protein ADP-ribosylhydrolase-like 1 is essential for heart chamber outgrowth and acts on muscle actin filament assembly. *Dev. Biol.* **416**, 373–388 (2016).
20. Pion, P. D., Kittleson, M. D., Rogers, Q. R. & Morris, J. G. Myocardial failure in cats associated with low plasma taurine: A reversible cardiomyopathy. *Science (80-. ).* (1987) doi:10.1126/science.3616607.
21. Moise, N. S. *et al.* Dietary taurine deficiency and dilated cardiomyopathy in the fox. *Am. Heart J.* (1991) doi:10.1016/0002-8703(91)90724-V.
22. Takihara, K. *et al.* Beneficial effect of taurine in rabbits with chronic congestive heart failure. *Am. Heart J.* (1986) doi:10.1016/0002-8703(86)90360-1.
23. Kaplan, J. L. *et al.* Taurine deficiency and dilated cardiomyopathy in golden retrievers

fed commercial diets. *PLoS One* (2018) doi:10.1371/journal.pone.0209112.
24. Van Rheenen, W. *et al.* Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nat. Genet.* (2016) doi:10.1038/ng.3622.
25. Schafer, S. *et al.* Titin-truncating variants affect heart function in disease cohorts and the general population. *Nat. Genet.* **49**, 46–53 (2017).