GWAS-driven Pathway Analyses and Functional Validation Suggest *GLIS1* as a Susceptibility Gene for Mitral Valve Prolapse

Mengyao Yu[1,2], Adrien Georges[1,2], Nathan R. Tucker[3,4], Sergiy Kyryachenko[1,2], Katelyn Toomer[5], Jean-Jacques Schott[6,7,8], Francesca N. Delling[9], Patrick T. Ellinor[3,4], Robert A. Levine[10], Susan A. Slaugenhaupt[11], Albert A. Hagège[1,2,12], Christian Dina[6,7,8], Xavier Jeunemaitre[1,2,13], David J. Milan[3], Russell A. Norris[5], Nabila Bouatia-Naji[1,2*]

[1] INSERM, UMR970, Paris Cardiovascular Research Center, 75015 Paris, France

[2] University Paris Descartes, Sorbonne Paris Cité, Faculty of Medicine, 75006 Paris, France

[3] Cardiovascular Research Center, Cardiology Division, Massachusetts General Hospital, Harvard Medical School, 55 Fruit Street, Boston, Massachusetts 02114 USA,

[4] Precision Cardiology Laboratory, The Broad Institute, Cambridge, MA 02124 USA

[5] Cardiovascular Developmental Biology Center, Department of Regenerative Medicine and Cell Biology, College of Medicine, Children's Research Institute, Medical University of South Carolina, 171 Ashley Avenue, Charleston, SC 29425, USA

[6] Inserm U1087; institut du thorax; University Hospital Nantes, France

[7] CNRS, UMR 6291, Nantes, France

[8] Université de Nantes, Nantes, France.

[9] Department of Medicine, Division of Cardiology, University of California, San Francisco, 94143

[10] Cardiac Ultrasound Laboratory, Cardiology Division, Massachusetts General Hospital, Harvard Medical School, 55 Fruit Street, Boston, Massachusetts 02114 USA

[11] Center for Human Genetic Research, Massachusetts General Hospital and Department of Neurology, Harvard Medical School, 185 Cambridge St., Boston, MA 02114 USA.

[12] Assistance Publique – Hôpitaux de Paris, Department of Cardiology, Hôpital Européen Georges Pompidou, 75015, Paris, France

[13] Assistance Publique – Hôpitaux de Paris, Department of Genetics, Hôpital Européen Georges Pompidou, 75015, Paris, France

DJM, RAN and NBN are joint senior authors for this work.

*Correspondence to nabila.bouatia-naji@inserm.fr

Abstract (Abstract length: limited to 150 words)

Nonsyndromic Mitral valve prolapse (MVP) is a common degenerative valvular heart disease with severe health consequences, including arrhythmia, heart failure and sudden death. MVP is characterized by excess extracellular matrix secretion and cellular disorganization which leads to bulky valves that are unable to co-apt properly during ventricular systole. However, the triggering mechanisms of this process are mostly unknown. Using pathway enrichment tools applied to GWAS we show that genes at risk loci are involved in biological functions relevant to cell adhesion and migration during cardiac development and in response to shear stress. Through genetic, *in silico* and *in vivo* experiments we demonstrates the presence of several genes involved in gene regulation, including *GLIS1*, a transcription factor that regulates Hedgehog signaling. Our findings define genetic, molecular and cellular mechanisms underlying non-syndromic MVP and implicate disrupted endothelial to mesenchymal transition and cell migration as a potential common cause to this disease.

Keywords: heart valve disease, valve development, mitral valve prolapse, GWAS based enrichment analysis

Mitral valve prolapse (MVP) is a common heart valve disease with important health consequences and an estimated prevalence of 2.4% in general populations[1,2]. It is defined as an abnormal mitral leaflet displacement into the left atrium during systole[3] and is the most common and an increasing indication for surgical repair of mitral regurgitation (MR)[4,5]. MVP is a risk factor for heart failure and sudden death in community-based studies[1,4,6]. Prospective studies have showed that asymptomatic patients with low-risk presentation (e.g., moderate MR and ejection fraction ≥50%) can develop adverse MVP-related events, indicating wide heterogeneity in outcomes among individuals[4].

The mature valve structure represents the culmination of embryonic programs working together to create cellular and extracellular matrix (ECM) environments that can withstand the biomechanical stresses of repetitive cardiac motion [7]. These programs include early, growth factor-mediated endothelial to mesenchymal transition (EndoMT), followed by ECM organization and leaflet thinning. The mature postnatal leaflet structure is organized into stratified ECM layers of dense fibrous collagen in the ventricularis, a proteoglycan-rich central "spongiosa" and an elastrin-rich atrialis. Endothelial cells line the valve tissue whereas fibroblastic-like ECM-producing valve interstitial cells (VICs) cells make up the bulk of the valve[7]. In MVP, the normal layered structure is lost, with expansion of myxoid-like extracellular matrix (ECM) that includes fragmented collagen and elastin and overabundance of glycosaminoglycans. This renders the valve incompetent to withstand the normal physiological stresses of the beating heart and failure of normal leaflet apposition with prolapse into the left atrium. The process that leads to MVP and myxomatous degeneration is poorly understood. Previous reports have suggested that a potential mechanism is activation of quiescent VICs to myofibroblasts. These activated cells secrete matrix metalloproteinases that drive collagen and elastin fragmentation and release TGF-β that in turn

promotes further cell proliferation and myofibroblast differentiation[7]. However, the triggering mechanisms of myxomatous degeneration are still to be identified.

Previous familial and population genetic studies have contributed to the current understanding of MVP biology. The identification of mutations in the filamin A gene (*FLNA)*[8] followed by research in mice[9] confirmed the role of this actin-binding protein during fetal valve development, mainly by providing stability to F-actin networks and linking them to cell membranes, which protect cells against mechanical stress[10]. More recently, we found that loss of function mutations in *DCHS1*, coding a protein from the cadherin superfamily involved in cell adhesion, cause familial MVP[11]. *DCHS1* deficiency in VICs altered migration and cellular patterning from patients and provoked loss of cell polarity during valve development in mice and zebrafish models[11]. Additional clues to the etiology of MVP came through genome-wide association studies (GWAS) where we identified six risk loci that are robustly associated with genetic susceptibility to MVP[12]. Genetic and biological evidence supported the role of two genes; tensin 1 (*TNS1*), a focal adhesion protein involved in cytoskeleton organization, and a transcription factor called LIM and cysteine-rich domain 1 (*LMCD1*)[12]. However, biological functions and potential mechanisms at the remaining GWAS loci remain unknown.

In this study, we hypothesize that several loci, including those not prioritized according to the stringent GWAS statistical threshold (P-value$<5\times10^{-8}$), could contain biologically relevant genes and mechanisms to MVP. We applied several computational-based analytical methods to the GWAS findings to 1) highlight enriched biological mechanisms for MVP loci, 2) characterize their expression pattern in tissues and 3) identify biologically pertinent genes for follow-up. Our study provides evidence for the Hedgehog signaling component, *GLIS1* to be involved in MVP

46 susceptibility and provides supporting experimental validation of this gene in valve development
47 and the degenerative process.
48

49 # Results

50 ## Gene-set enrichment analyses for MVP loci

51 We first applied the SNP ratio test[13] (SRT) method that uses simulated datasets to estimate the
52 significance of a given pathway on the currently available GWAS data[12]. We analyzed ~2 million
53 SNPs that mapped within genes and found 42 nominally enriched pathways ($P_{empirical}$< 0.05)
54 (Supplementary Table 1). Among the top 10 enriched pathways we found several pathways that
55 are consistent with cardiac function, including viral myocarditis ($P_{empirical}$<0.001), hypertrophic
56 cardiomyopathy ($P_{empirical}$<0.002), dilated cardiomyopathy ($P_{empirical}$<0.002) and cardiac muscle
57 contraction ($P_{empirical}$<0.005). We also noted the enrichment in the phosphatidylinositol signaling
58 system ($P_{empirical}$<0.006), MAPK ($P_{empirical}$<0.05) and WNT ($P_{empirical}$<0.05) signaling that are
59 important for focal adhesion and the gap junction functions ($P_{empirical}$<0.05).
60 We next employed i-GSEA4GWAS, which highlighted 244 pathways as significantly enriched
61 (FDR < 0.05) for MVP (Supplementary Table 2) and uses a larger source of pathways and gene
62 sets compared to SRT (BioCarta, KEGG and Gene Ontology Biological processes). The most
63 enriched pathway was BioCarta EDG-1 pathways with 26 out of 27 genes harboring associated
64 variants to MVP. Globally, enriched pathways can be functionally classified into two main groups
65 related to valve biology. The first and largest group included 28 GO terms related to 'cytoskeleton',
66 'actin binding', 'focal adhesion', 'adhesion junction' and 'basolateral plasma membrane',
67 BioCarta gene set 'Integrin pathway' and KEGG pathway 'cell adhesion molecules'. The second

group was related to biological gene sets involved in the development of cardiovascular system and included 21 GO terms, notably 'heart development', 'vasculature development', 'regulation of heart contraction', 'cardiac muscle contraction', and 'angiogenesis'. Interestingly, many enriched gene sets covered functions related to gene regulation, specifically, 'transcription and regulation' (60 related pathways), e.g. 'transcription repressor activity', 'positive regulation of cell proliferation', 'positive regulation of transcription'. We note that this method also highlighted several consistent pathways with the enrichment obtained by the SRT method (e.g. dilated cardiomyopathy).

We also applied the integrative method *DEPICT* and used the recommended association cut-off (P_{GWAS}-value $<10^{-5}$) that identified 39 independent MVP loci harbouring 50 genes. We obtained 309 nominally enriched gene sets for MVP (P-value<0.05) and the Pearson distance matrix provided by DEPICT defined 36 clusters from these enriched gene sets (Supplementary Table 3). The second network is linked to developmental cardiovascular phenotypes which core node is cluster 3: vasculature development. This cluster contains mainly 'abnormal blood vessel morphology', 'abnormal outflow tract development', 'failure of heart looping', 'angiogenesis', 'dorso-ventral axis formation'. Other clusters in the second network were cluster 5: 'abnormal semilunar valve morphology', cluster 9: 'TGF-β receptor binding' and 'response to fluid shear stress', and cluster 13: 'abnormal fourth branchial arch artery morphology' and 'chromatin DNA binding'. Consistently, cluster 5, cluster 9 and cluster 13 were also significantly enriched for MVP genes and many gene sets included transcription factors as the most contributing genes (e.g *LMCD1*, *RUNX1*, *FOXL1*, *GLIS1*) (Supplementary Table 3).

Enrichment for expression in tissues and cell types

Based on reconstituted gene-tissue matrix generated by DEPICT, we analyzed the expression profile from arrays-based human transcriptomic data and found that 21 tissues or cell types were enriched for the expression of genes at MVP risk loci. The most enriched physiological system is the cardiovascular system with five significantly enriched tissues. Connective tissue, epithelial and stem cells, specifically mesenchymal stem cells and muscle were the most enriched cell types and tissues for the expression of MVP genes (Figure 1b, Supplementary Table 4).

Prioritization of *GLIS1* as a novel risk locus and gene for MVP

Prioritization from contribution to enrichment analyses

Among the loci that we included in the pathway analyses, we provide a focus on a sub-GWAS significant signal located on chromosome 1 that deserved prioritization in the light of gene function candidacy and clues from the results obtained in the enrichment analyses. We found that i-GSWA4GWAS revealed *GLIS1* as the best-ranked gene in six significantly (FDR<0.05) and two suggestively (FDR<0.25, P-value<0.05) enriched gene sets, all related to regulation of transcription (Supplementary Table 5). Interestingly, Hh signaling pathway (from KEGG) was also associated with MVP (P-value=0.015, FDR=0.054, Supplementary Table 2). Although *GLIS1* is not reported as part of this pathway according to KEGG, several other genes from this pathway contain significantly associated variants with MVP (e.g *WTN5B*, *BMP6*, *HHIP*, *SHH*, *GLI2* and *GLI3* Supplementary Table 6). According to the DEPICT analysis, we found that *GLIS1* significantly contributes to the enrichment of several tissues and cell types namely the connective tissue (*GLIS1* Z-score=2.3), connective tissue cells (*GLIS1* Z-score=2.8) and mesenchymal stem cells (Z-score=2.3) (Supplementary Table 4). Of note, *GLIS1* is the most contributing MVP gene

(Z-score=6.2) to suggestive enrichments (P-value=0.13, FDR>=0.20) of two cardiovascular system tissues (aortic valve and heart valves, Supplementary Table 4).

Genetic association at the GLIS1 *locus*

The association context at the *GLIS1* locus is presented in Figure 2a. The most associated SNP with MVP is rs1879734 and locates in the first intron (OR=1.30, P-value $=7.1\times10^{-6}$, risk allele: T, Freq: 0.28) (Supplementary Table 7). Follow-up of this SNP in four case control studies provided positive replication in the two largest studies, consistent direction of effect in all studies and no evidence for heterogeneity (OR_{all}=1.23, $P_{effect}=1.2\times10^{-7}$, $P_{Heterogeneity}=0.41$) (Supplementary Table 7).

We performed functional annotation for 111 SNPs in moderate LD with rs18797434 (r^2>0.5) that showed nominal association with MVP (P_{GWAS}-value<0.05) and all mapped to introns1 and 2 in *GLIS1*(Supplementary Table 8). We found that 47 SNPs locate in DNase Hypersensitive region (DNASEV) in diverse normal tissues and 16 locate in transcription factor binding sites (TFBS). rs1879734 is located in a TFBS for endothelial transcription factor GATA2. Interestingly, we observed that rs12091931, which locates in DNASEV in non-pigmented ciliary epithelial cells, is in TFBSs for JUND, MAFF, MAFK, UBTF, MAZ, CTCF, ZBTB7A, TBP (Supplementary Table 8). Hereafter, the examination of the GTEx eQTL portal indicates that 58 SNPs, including the lead SNP rs1879734 were potential nominal eQTLs (P<0.05) for *GLIS1* in heart atrial appendage tissue (N=264). The most significant eQTL for *GLIS1* is rs2950241 (P-value=9×10^{-3}), which is highly correlated to rs1879734 (r^2=0.96, Hapmap CEU).

The examination of histone marks (Figue 2b) in mesenchymal stem cell originated from H1-hESC shows the presence of robust enhancer mark (H3K4me1) in the vicinity of rs1879734. In H1-hESC,

strong signal of enhancer mark (H3K4me1) was presented at rs4927029, rs12097598, rs11206201, rs55786134, rs17109178. However, none of the 111 SNPs showed signals of histone marks in heart related tissue/cell types (pulmonary artery endothelial cell, cardiac mesoderm, heart left ventricle) (Figue 2b). *In silico* 4C experiment using four cell types including THP-1[14], HUVEC[15], H1-hESC[16] and NHEK[15] (Supplementary Figure 1) showed that the associated SNPs at this locus physically interact only with *GLIS1* regulatory sequences suggesting it is the potential target and causal gene at this locus.

Expression during heart development in mouse

To study the pattern of expression of *GLIS1* during valve development, we performed immunohistochemistry (IHC) experiments of the mouse ortholog, *Glis1* during embryonic, foetal and adult time points (Figure 3). During embryonic development, Glis1 is expressed predominantly in nuclei of endothelial cells of the valves as well as the VICs. As the valves mature during foetal gestation, Glis1 is retained in a subset of endocardial and interstitial cells. By 6-months of age, Glis1 is much weaker in the valve leaflet with only scant cytoplasmic staining in endocardial cells. Weaker expression of Glis1 could be detected in the myocytes, epicardium and endocardium of the ventricular myocardium. These data show that Glis1 is embryonically expressed and that levels of this protein are rarely detected in the postnatal mouse, suggesting an important role for Glis1 in regulating valve morphogenesis during early development.

Knockdown of *Glis-1* cause atrioventricular regurgitation in zebrafish

To analyse the potential effects of *GLIS-1* on valvular development and function, we chose to investigate its expression in the zebrafish. Due to genome duplication, zebrafish are predicted to have two orthologues of GLIS-1, *glis1a* and *glis1b*. We designed antisense morpholino oligonucleotides to target splice junctions in each with aims of rendering the transcript non-functional. Knockdown of *glis1a* was 65% efficient at 72 hours post fertilization, but had no discernible effect on atrioventricular valve function (Supplemental Figure 2). Knockdown of *glis1b* at 72 hours post fertilization was robust but slightly less efficient (37.3% and 36.5% reduction Figure 4a, b) and had minimal effects on the overt morphology of the developing embryo (Figure 4c). However, this knockdown resulted in a significant increase in the incidence of severe atrioventricular regurgitation when compared to controls. This increase was observed with two independently injected morpholinos with a combined fold increase of 1.6 (P=0.01) (Figure 4d).

Discussion

Our study supports that MVP genes are predominantly part of gene sets involved in cardiac development, the regulation of cell adherence and migration, and the regulation of cytoskeleton biology, focal adhesion and the interaction with the extracellular matrix. We also show that MVP genes are highly expressed in the cardiovascular system, connective and mesenchymal tissues. These genes are often regulatory genes expressed in the nucleus, as it is the case for *GLIS1*. Our findings suggest that this Hedgehog signalling related transcription factor is expressed in the

nucleus of developing mitral valves in mouse and is required for atrioventricular development and function in zebrafish.

The application of three gene set enrichment tools to MVP GWAS data, using overall different computational methods and diverse databases provided overall consistent enrichments and an unprecedented resource to understand mitral valve biology. One of the most significantly enriched gene sets for MVP genes according to the i-GSEA4GWAS method is the endothelial differentiation gene (EDG-1) pathway where 26 out of 27 genes are associated with MVP. EDG-1 is a key signalling pathway for actin assembly and lamellipodia formation through the activation of integrins alpha V (ITGAV) and beta 3 (ITGB3) via RAS homolog gene family member A (RHOA). Edg1 knockout mice show embryonic haemorrhage leading to intrauterine death due to incomplete vascular maturation and defects in SPP-induced migration response[17], which are important steps prior to valve sculpting at the embryonic stages. Cardiomyocyte specific knockout of EDG1 (alias S1PR1 for sphingosine 1-phosphate receptor-1) resulted in ventricular noncompaction and ventricular septal defects at 18.5 days post conception in mice, supporting a crucial role for this pathway in cardiac development[18]. Whether genes from EDG1 signalling are specifically important for valve development needs future examination. In support of this hypothesis is the reported interaction during lamellipodia formations of S1PR1 and Filamin A[19], which gene is reported to be mutated in familial forms of MVP[20]. Many additional gene sets enriched for MVP loci are also involved in cardiac development or disease. These enrichments suggest critical roles for genetic variability near genes acting during early heart development that require important cytoskeleton remodelling and organization for cell migration during the endothelial-to-mesenchymal transition (EndoMT).

Our enrichment analyses also pointed to various pathways and gene sets whose link to valve disease is currently in its infancy. For instance, our analyses indicate an enrichment for gene sets in the immune system and phagosome pathways in MVP disease. The contention that the immune system may be involved in MVP is further supported by work from our group as well as others[9,21]. Additionally, the phagosome pathway has recently been reported as dysregulated in human myxomatous mitral valves compared with healthy valves[22] and high activity of autophagosomes are present during cardiac development[23]. In addition, the knockdown of autophagy genes caused defects in cardiac looping and aberrant valve development as the results of ectopic expression of critical TF involved in heart development, mainly foxn4, tbx5 and tbx2[23].

We report several top contributing genes at MVP loci to be transcription factors, especially to gene sets related to cardiovascular development. Consistently, we report enrichment for the protein localization to nuclear gene sets. Here, we followed-up specifically *GLIS1*, which was one of the top associated SNPs with MVP that we confirmed as a potential sub-GWAS risk locus for MVP. In addition to the expression in developing heart valves, *GLIS1* was the most significantly contributing gene to the enrichment of several gene sets related to the regulation of transcription. Little is known about GLIS1 function in connection with valve biology. We found that GLIS1 is expressed during developing valves in mouse valve endocardial and interstitial cells, suggesting for the first time its potential regulatory role during heart development. There is evidence for Glis1 to markedly increase the efficiency of generating induced pluripotent stem cells (iPSC) from mouse somatic cells in the presence of OCT3/4, SOX2 and KLF4[10]. This study demonstrated that GLIS1 directly interacts with KLF4 and induces the expression of Forkhead box genes, especially FOXA2 and several WNT genes to enhance mesenchymal to epithelial transition, a mechanism required for cell reprogramming[24]. The Hh signalling regulates both FOXA2 and WNT genes and

its suggestive enrichment in our MVP GWAS involves for the first time in valve disease this important pathway for cell migration and morphogenesis organisation during heart development[25]. The Hh signalling induces EndoMT, through the up-regulation of NOTCH and TGF-ß signalling[26] and is coordinated by primary cilia during cell migration and heart morphogenesis[25]. The contributing genes to the enrichment of the Hh signalling pathway included *BMP2* that harbours causative mutations for cardiac anomalies[27] and *BMP4* that is required for outflow-tract septation[28]. Interestingly, a common variant in the also associated *WNT8A*, member of the WNT genes family is associated with atrial fibrillation[29] and was previously shown to be regulated by GLIS1[10].

This work presents however several limitations. One critical step when performing enrichment analyses in GWAS loci is the attribution of variants to genes. The pathway analyses of i-GWSEA4GWAS and DEPICT rely on SNPs mapped to genes using physical distance and LD block information. This have limited the number of genes analysed at risk loci and excluded more distant genes of interest. There is established evidence in favour of the functional role of distant long-range regulatory variants in predisposition to complex diseases[30]. In most cases, associated variants in GWAS loci are less likely to regulate the closest genes and be hundreds of kilobases away from culprit genes. This limitation explains the absence from the prioritization list generated by DEPICT of *TNS1*, a focal adhesion protein-coding gene that we have previously incriminated in MVP through genetic and functional investigation[12].

In conclusion, our pathway investigation supports that genes near MVP associated loci are involved in biological functions relevant to cell adhesion and migration during cardiac development and in response to shear stress, and highlight the importance of regulatory mechanisms. Our study also provides genetic *in silico* and *in vivo* functional exploration of *GLIS1,* a transcription factor that regulate members of the Hh signalling and implicates this important

246 biological mechanism for EndoMT and cell migration during heart development in valve
247 myxomatous disease.

248 # Methods

249 ## GWAS Study populations

250 The clinical characteristics of MVP patients, controls and GWAS methods have been previously
251 described[12]. Briefly, we applied pathway-based methods to a meta-analysis of two GWAS
252 conducted in 953 patients from MVP-France and 1566 controls and 489 from MVP-Nantes and
253 873 controls[12]. Replication was performed in four case-control studies (totalling 1,422 European
254 MVP patients and 6,779 controls) as previously described[12]. Local ethics committees approved all
255 studies, and all patients and controls provided written informed consent.
256 The input data was 6.6 million (M) genotyped or imputed SNPs (MAF > 0.01) and used either
257 SNPTEST[31] (imputed) or PLINK[32] (directly genotyped) to perform the association test. Regional
258 association plot at *GLIS1* locus was created using LocusZoom[33].
259

260 ## Pathway analyses

261 *SNP ratio test (SRT)*
262 SRT is a tool that calculates an empirical P-value by comparing the proportion of significant SNPs
263 (P_{GWAS}-value <0.05) in the original GWAS to randomized GWAS for the phenotype of cases and
264 controls[13]. SRT tests enrichment using the KEGG pathways catalog and then allocates SNPs to
265 genes using physical position and does not consider intergenic SNPs. Approximately 2 million
266 GWAS SNPs were mapped to the reformed KEGG pathways and gene sets. We first generated
267 1000 simulated alternative phenotypes for individuals tested and ran the association analyses using

PLINK[32]. The empirical P-value for a given pathway is defined by $P_{empirical}=(s+1)/(N+1)$, where *s* is the number of times that a simulated ratio significant to non-significant was greater than or equal to the ratio obtained from the GWAS computed with real phenotypes, *N* is the total number of simulations (here *N*=1000) and enrichment was higher when P was small .We set significance to *P*-value<0.05 in simulated and real GWAS, and for the empirical P-value.

Gene set enrichment analyses using i-GSEA4GWAS

i-GSEA4GWAS v2 uses SNPs and their corresponding P-values from GWAS results as input[34,35]. All SNPs from the GWAS meta-analysis were used and were mapped to genes if they are exonic/intronic or to the closest genes if they are within 20 Kb upstream/downstream a gene. Of the 6.6M tested, 4.3M variants were mapped to 21,167 different genes and then genes were attributed to pathways and/or gene sets using BioCarta, KEGG, and GO terms from MSigDB v4.0. Only GO terms with experimental evidence (codes IDA IPI, IMP IGI, IEP), computational analysis evidence codes (ISS) and author evidence statement (TAS) were taken into account. In total, 936 gene sets contributed to calculate the significant proportion based enrichment score (SPES). We used the authors' recommendation to consider pathways/gene sets with FDR<0.25 as suggestively associated with disease and FDR<0.05 as high confidence or statistically significant enrichment.

Integrative analyses using DEPICT

Data-driven expression prioritized integration for complex traits (DEPICT) (version rel194) is an integrative tool that utilizes diverse sources to predict reconstituted gene sets specific to each meta-analysis of GWAS[36]. DEPICT uses predefined gene sets including protein-protein interactions database, Mouse Genetics Initiative database, Reactome, KEGG and GO terms. It also uses a co-regulation data frame downloaded from the Gene Expression Omnibus (GEO) database including human and rodent expression microarrays (total arrays 77,840, including 37,427 generated in

291 human tissues). This information was used to perform gene prioritization at loci, gene set
292 enrichment analyses, and provide genes expression profiles in 209 tissues or cell types defined by
293 Medical Subject Heading (MeSH). Input of DEPICT is a list of the most significantly associated
294 SNPs (P_{GWAS}-value$<1\times10^{-5}$) after LD pruning. Loci were mapped to genes using LD ($r^2>0.5$).
295 Genes where several SNPs map are counted once. To group overlapping gene sets, pairwise
296 Pearson's correlation coefficient of all enriched gene sets was calculated and the Affinity
297 Propagation (AP) method[36,37] was used to cluster gene sets that are highly correlated to
298 automatically define independent clusters based on the Pearson distance matrix. We set
299 significance for enrichment to P-value < 0.05, given that no gene sets reached an FDR<0.05.
300
301 **Variants Annotation**
302 We used the UCSC genome browser tool Variant Annotation Integrator (VAI) to annotate SNPs
303 at the *GLIS1* locus, and indicate if SNPs were located on DNase Hypersensitive region or
304 transcription factor binding site generated by ENCODE[38]. Tissue specific expression quantitative
305 trait loci (eQTL) annotation was extracted from GETx portal (https://www.gtexportal.org/home/,
306 Release V7, dbGaP Accession phs000424.v7.p2). Annotation of the 111 selected SNPs by Human
307 ChIP-Seq (markers: H3k4ME1 and H3k4ME3) and DNase-Seq were downloaded from ENCODE
308 (https://www.encodeproject.org/) and visualized by Integrated Genome Browser (IGB)[39]. In the
309 absence of valvular cells in this database, selected tissues/cells used in this analyses are
310 mesenchymal stem cell originated from H1-hESC[40], H1-hESC[40], pulmonary artery endothelial
311 cell[41], cardiac mesoderm[41] and heart left ventricle[40]. Accession numbers of those files are
312 (respectively): ENCFF152YQG (H3K4me1), ENCFF712CJP (H3K4me3), ENCFF623ZAW
313 (H3K4me3), ENCFF593OAZ (H3K4me1), ENCFF719ZEX (DNase-seq), ENCFF591TLE

(H3K4me3), ENCFF213FJN (DNase-seq), ENCFF094USN (DNase-seq), ENCFF254JZR (DNase-seq). The annotation of the *GLIS1* region used the Hi-C data from four cell types including THP-1[14], HUVEC[15], H1-hESC[16] and NHEK[15] from ENCODE, all visualized by IGB.

Protein detection in mouse embryos and adult hearts

In all IHC experiments, 5-min antigen retrieval was performed with VectaStain in a pressure cooker (Cuisinart). The antibodies used for IHC were: Glis1 (Novus, NB100-41087) and myosin heavy chain (Developmental Hybridoma Banks, MF20). Primary antibodies were used at a 1:100 dilution; Hoechst 33342 (nuclear stain) was used at a 1:10,000 dilution. Appropriate secondary antibodies were used for detection. Three time points were used in IHC: (i) completion of the endothelial-to-mesenchymal transition (EndoMT; embryonic day (E) 13.5), (ii) valve sculpting and elongation (E17.5) and (iii) achievement of the mature adult form (at 9 months).

Zebrafish experiments

Zebrafish experiments were performed in accordance with approved Institutional Animal Care and Use Committee (IACUC) protocols. Zebrafish of the Tu-AB strain were reared according to standard techniques. Morpholinos were designed against the two zebrafish orthologues of GLIS1: *glis1a* (5'-GAGAATGGTCGTACATACCGTGTCC) and *glis1b* (1: 5'-AAGTGCACTGAGGTCTCACCCTGTG 2: 5'-TTAAGGTCAGGTACTCACAGTGTCC). Empirically established effective doses of antisense morpholino oligonucleotides were injected at the single-cell stage, and morpholino-injected animals were compared to controls injected with a

335 non-targeting morpholino. Analysis of atrioventricular regurgitation was performed at 72 hours
336 post fertilization as described[12].

REFERENCES

1 Freed, L. A. *et al.* Prevalence and clinical outcome of mitral-valve prolapse. *New England Journal of Medicine* **341**, 1-7 (1999).

2 Delling, F. N. & Vasan, R. S. Epidemiology and pathophysiology of mitral valve prolapse: new insights into disease progression, genetics, and molecular basis. *Circulation* **129**, 2158-2170 (2014).

3 Guy, T. S. & Hill, A. C. Mitral valve prolapse. *Annual review of medicine* **63**, 277-292 (2012).

4 Avierinos, J. *et al.* Natural history of asymptomatic mitral valve prolapse in the community. *Circulation* **106**, 1355-1361 (2002).

5 Gammie, J. S. *et al.* Isolated Mitral Valve Surgery: The Society of Thoracic Surgeons Adult Cardiac Surgery Database Analysis. *Ann Thorac Surg* **106**, 716-727, doi:10.1016/j.athoracsur.2018.03.086 (2018).

6 Freed, L. A. *et al.* Mitral valve prolapse in the general population: the benign nature of echocardiographic features in the Framingham Heart Study. *J Am Coll Cardiol* **40**, 1298-1304 (2002).

7 Levine, R. A. *et al.* Mitral valve disease--morphology and mechanisms. *Nat Rev Cardiol* **12**, 689-710 (2015).

8 Kyndt, F. *et al.* Mutations in the gene encoding filamin A as a cause for familial cardiac valvular dystrophy. *Circulation* **115**, 40-49 (2007).

9 Sauls, K. *et al.* Developmental basis for filamin-A-associated myxomatous mitral valve disease. *Cardiovascular research* **96**, 109-119 (2012).

10 Maekawa, M. *et al.* Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. *Nature* **474**, 225 (2011).

11 Durst, R. *et al.* Mutations in DCHS1 cause mitral valve prolapse. *Nature* **525**, 109-113 (2015).

12 Dina, C. *et al.* Genetic association analyses highlight biological pathways underlying mitral valve prolapse. *Nat Genet* **47**, 1206-1211 (2015).

13 O'dushlaine, C. *et al.* The SNP ratio test: pathway analysis of genome-wide association datasets. *Bioinformatics* **25**, 2762-2763 (2009).

14 Phanstiel, D. H. *et al.* Static and Dynamic DNA Loops form AP-1-Bound Activation Hubs during Macrophage Development. *Molecular cell* **67**, 1037-1048 (2017).

15 Rao, S. S. P. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680 (2014).

16 Dixon, J. R. *et al.* Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376-380, doi:10.1038/nature11082 (2012).

17 Liu, Y. *et al.* Edg-1, the G protein-coupled receptor for sphingosine-1-phosphate, is essential for vascular maturation. *J Clin Invest* **106**, 951-961, doi:10.1172/JCI10905 (2000).

18 Clay, H. *et al.* Sphingosine 1-phosphate receptor-1 in cardiomyocytes is required for normal cardiac development. *Dev Biol* **418**, 157-165, doi:10.1016/j.ydbio.2016.06.024 (2016).

19 Maceyka, M., Alvarez, S. E., Milstien, S. & Spiegel, S. Filamin A links sphingosine kinase 1 and sphingosine-1-phosphate receptor 1 at lamellipodia to orchestrate cell migration. *Molecular and cellular biology* **28**, 5687-5697, doi:10.1128/MCB.00465-08 (2008).

20 Le Tourneau, T. *et al.* New insights into mitral valve dystrophy: a Filamin-A genotype-phenotype and outcome study. *European heart journal* **39**, 1269-1277 (2018).
21 Hulin, A. *et al.* Macrophage Transitions in Heart Valve Development and Myxomatous Valve Disease. *Arterioscler Thromb Vasc Biol* **38**, 636-644, doi:10.1161/ATVBAHA.117.310667 (2018).
22 Deroyer, C. *et al.* New biomarkers for primary mitral regurgitation. *Clinical proteomics* **12**, 25, doi:10.1186/s12014-015-9097-2 (2015).
23 Lee, E. *et al.* Autophagy is essential for cardiac morphogenesis during vertebrate development. *Autophagy* **10**, 572-587, doi:10.4161/auto.27649 (2014).
24 Scoville, D. W., Kang, H. S. & Jetten, A. M. GLIS1-3: emerging roles in reprogramming, stem and progenitor cell differentiation and maintenance. *Stem cell investigation* **4**, 80, doi:10.21037/sci.2017.09.01 (2017).
25 Koefoed, K., Veland, I. R., Pedersen, L. B., Larsen, L. A. & Christensen, S. T. Cilia and coordination of signaling networks during heart development. *Organogenesis* **10**, 108-125 (2014).
26 Katoh, Y. & Katoh, M. Hedgehog signaling, epithelial-to-mesenchymal transition and miRNA. *International journal of molecular medicine* **22**, 271-275 (2008).
27 Tan, T. Y. *et al.* Monoallelic BMP2 Variants Predicted to Result in Haploinsufficiency Cause Craniofacial, Skeletal, and Cardiac Features Overlapping Those of 20p12 Deletions. *Am J Hum Genet* **101**, 985-994, doi:10.1016/j.ajhg.2017.10.006 (2017).
28 Liu, W. *et al.* Bmp4 signaling is required for outflow-tract septation and branchial-arch artery remodeling. *Proc Natl Acad Sci U S A* **101**, 4489-4494, doi:10.1073/pnas.0308466101 (2004).
29 Ellinor, P. T. *et al.* Meta-analysis identifies six new susceptibility loci for atrial fibrillation. *Nat Genet* **44**, 670-675, doi:10.1038/ng.2261 (2012).
30 Kleinjan, D. A. & van Heyningen, V. Long-range control of gene expression: emerging mechanisms and disruption in disease. *Am J Hum Genet* **76**, 8-32, doi:10.1086/426833 (2005).
31 Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* **39**, 906-913 (2007).
32 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575 (2007).
33 Pruim, R. J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336-2337 (2010).
34 Zhang, K., Cui, S., Chang, S., Zhang, L. & Wang, J. i-GSEA4GWAS: a web server for identification of pathways/gene sets associated with traits by applying an improved gene set enrichment analysis to genome-wide association study. *Nucleic Acids Res* **38**, W90-W95 (2010).
35 Zhang, K., Chang, S., Guo, L. & Wang, J. I-GSEA4GWAS v2: a web server for functional analysis of SNPs in trait-associated pathways identified from genome-wide association study. *Protein cell* **6**, 221 (2015).
36 Pers, T. H. *et al.* Biological interpretation of genome-wide association studies using predicted gene functions. *Nat Commun* **6**, 5890 (2015).

37 Frey, B. J. & Dueck, D. Clustering by passing messages between data points. *science* **315**, 972-976 (2007).

38 Consortium, E. N. C. O. D. E. P. A user's guide to the encyclopedia of DNA elements (ENCODE). *PLoS biology* **9**, e1001046 (2011).

39 Freese, N. H., Norris, D. C. & Loraine, A. E. Integrated genome browser: visual analytics platform for genomics. *Bioinformatics* **32**, 2089-2095, doi:10.1093/bioinformatics/btw069 (2016).

40 Hawkins, R. D. *et al.* Distinct epigenomic landscapes of pluripotent and lineage-committed human cells. *Cell Stem Cell* **6**, 479-491, doi:10.1016/j.stem.2010.03.018 (2010).

41 Thurman, R. E. *et al.* The accessible chromatin landscape of the human genome. *Nature* **489**, 75-82, doi:10.1038/nature11232 (2012).

Acknowledgments

This study was supported by a PhD scholarship from the Chinese Scientific Council to MY, and partly funded by French Agency of Research (ANR-16-CE17-0015-02). The work at MUSC was performed in a facility constructed with support from the National Institutes of Health, Grant Number C06 RR018823 from the Extramural Research Facilities Program of the National Center for Research Resources. Other funding sources: National Heart Lung and Blood Institute: HL131546 (RAN), COBRE GM103342 (RAN), GM103444 (RAN), HL127692 (DJM, SAS, RAN, RAL), American Heart Association: 17CSA33590067 (RAN) and HL140187 (NRT). The recruitment of the MVP France cohort was supported by the French Society of Cardiology. We acknowledge the contribution of the Leducq Foundation, Paris for supporting a transatlantic consortium investigating the physiopathology of mitral valve disease, for which this genome-wide association study was a major project (coordinators: R.A.L. and A.A.H.). We acknowledge investigatrs who contributed access to validation in cohorts: Leticia Fernandez-Friera, Jorge Solis from Centro Nacional de Investigaciones Cardiovasculares (CNIC), Yohan Bossé and Philippe Pibarot from *Institut universitaire de cardiologie et de pneumologie de Québec-Université Laval*, Quebec, Canada, Ramachandran S. Vasan, Ming-Huei Chen and Emilia J. Benjamin from the Framingham Heart Study, USA, Thierry Le Tourneau, Richard Redon, Hervé Le Marec and Vincent Probst from Institut du Thorax, Nantes, France and Ronen Durst, Hassadah Hebrew University Medical Center, Jerusalem, IL.

Author contributions

Recruitment of patients: A.A.H, R.A.L, S.A.S, F.N.D, X.J.
Genotyping: X.J, J.-J.S, C.D, S.A.S.

Data analysis: M.Y, A.G, S.K
Animal experiments: D.J.M., R.A.N., N.T., P.T.E., K.T.
Manuscript writing: N.B.-N., M.Y, R.A.N., N.T., R.A.L.
Manuscript approval: all authors.

Competing interests

None

Materials & Correspondence

Figures

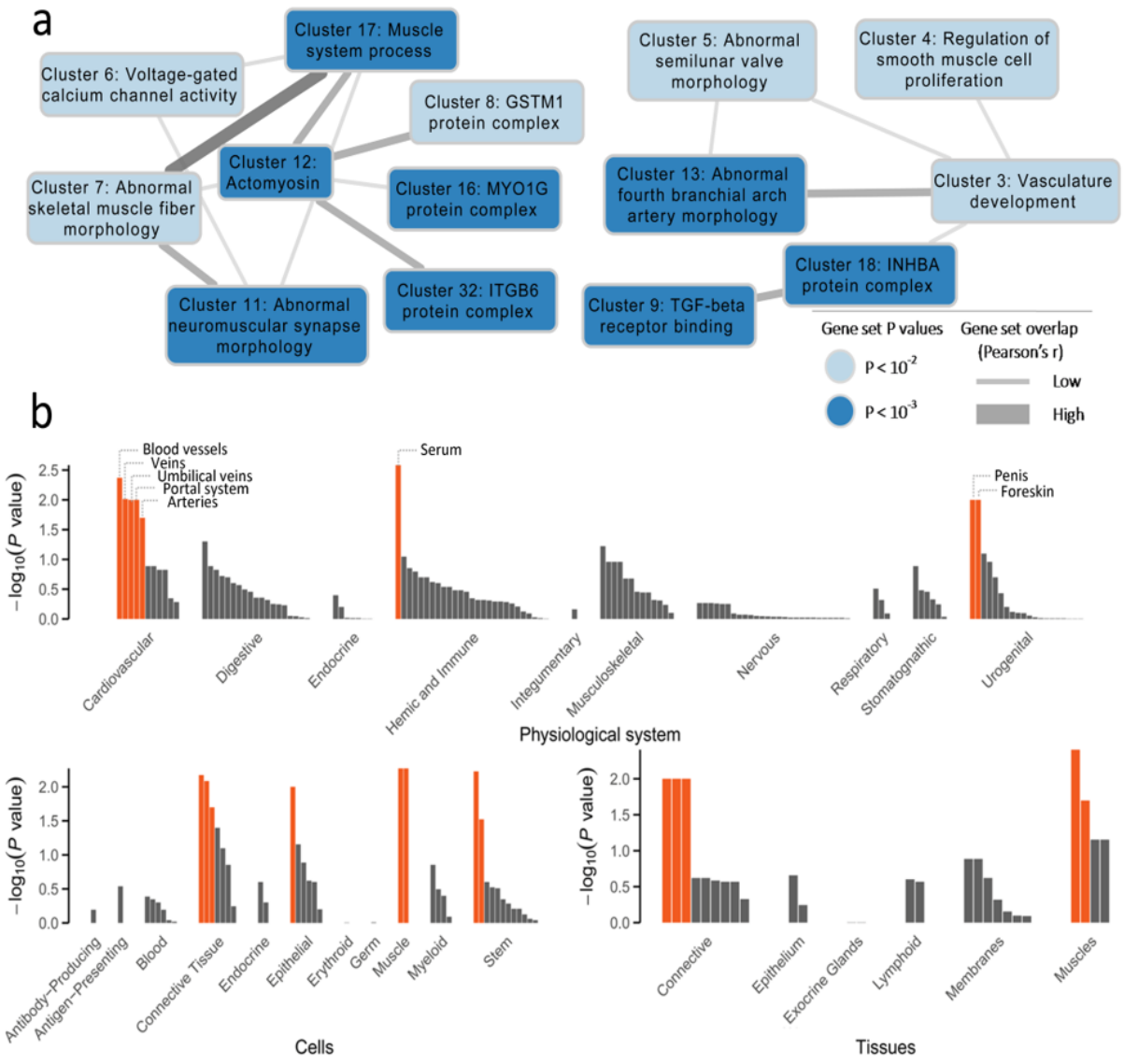


Figure 1. Networks of gene sets and tissue enrichment for the expression of genes at MVP associated loci. (a) Clustering networks formed from MVP enriched gene sets. The significance of the enrichment of gene sets are indicated (light blue: P-value < 10^{-2} and dark blue: P-value <10^{-3}). The thicker the edge, the more similar functions exist between connected gene sets (all Pearson's r > 0.3). **(b)** Tissue enriched for MVP genes. Enrichment is organised by physiological system, cells and tissues. Tissues/cells type with FDR < 0.20 are marked in orange.

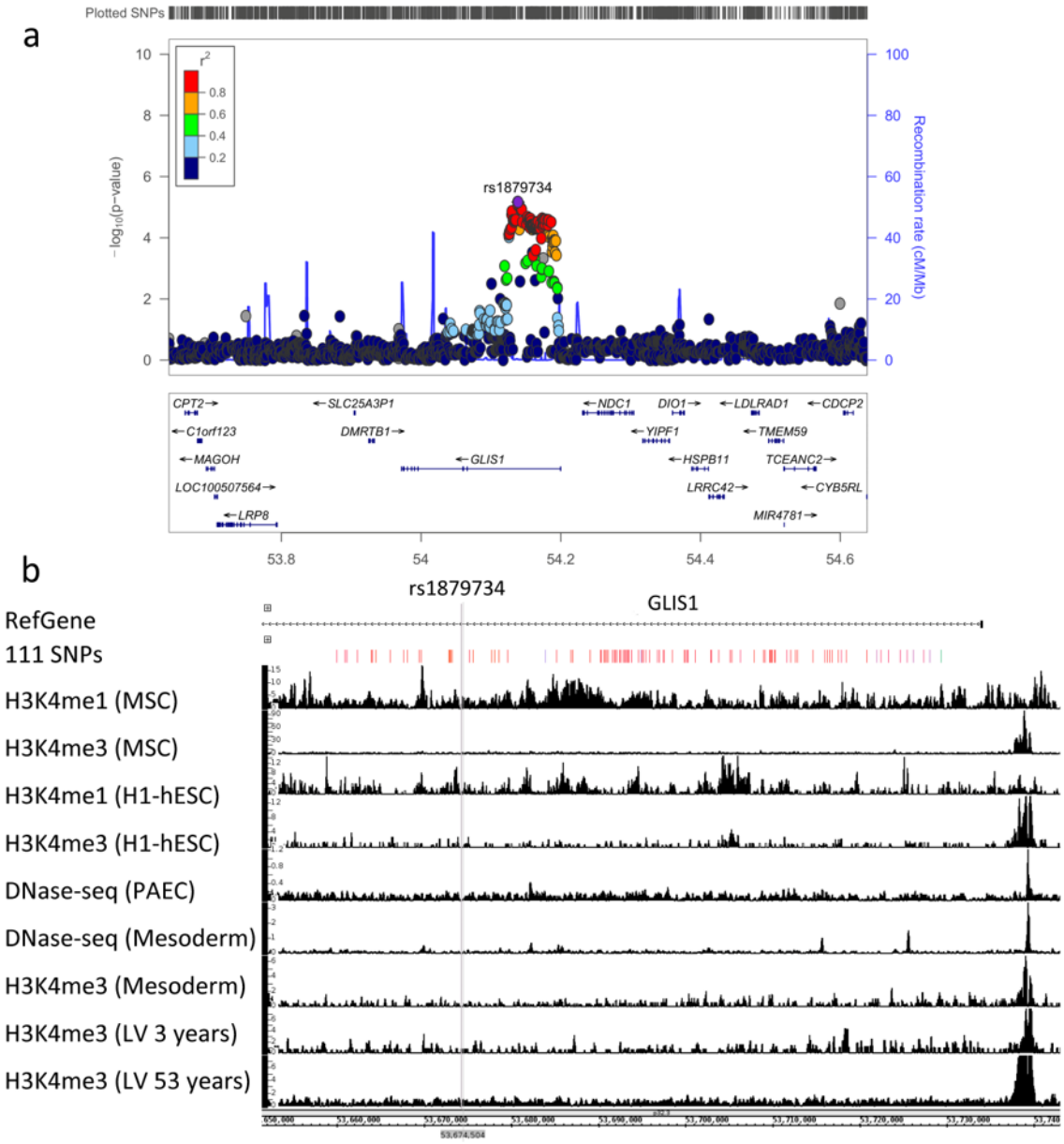


Figure 2. Genomic context and functional annotation of the association signal observed in the GWAS meta-analysis. **(a)** The regional association plot was generated using LocusZoom and displays surrounding genes (±500 Kb). The association signal is intronic to *GLIS1*. Round points represent SNPs in this region and purple point represent SNP rs1879734, the top associated SNP. **(b)** Visualization of histone marks and DNAse-seq density profiles in several tissues/cells based on ENCODE data. From top to bottom: reference gene; the selected 111 SNPs that is in high LD with rs1879734; H3K4me1 and H3K4me3 were from mesenchymal stem cell originated from H1-hESC; H3K4me1 and H3K4me3 from H1-hESC; DNase-seq from pulmonary artery endothelial cell; DNase-seq and H3K4me3 from cardiac mesoderm; H3K4me1 and H3K4me3 from heart left ventricle.

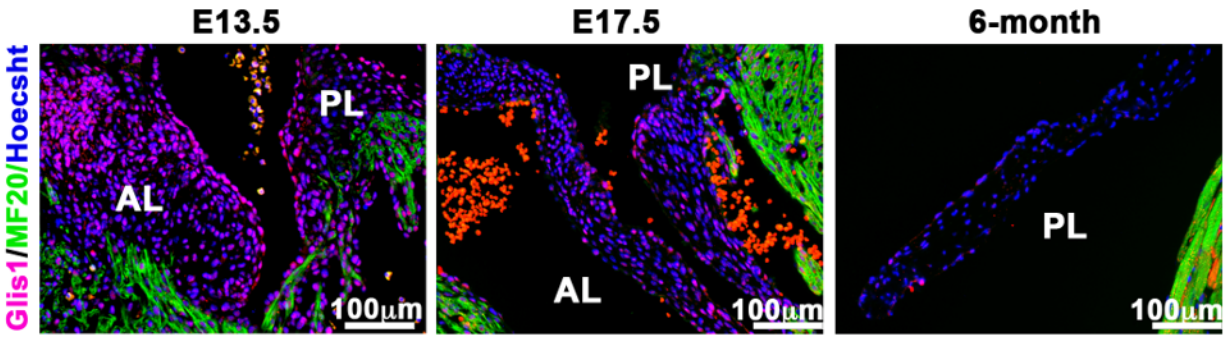


Figure 3. Glis1 expression in mouse developing and adult heart. Cardiac expression of Glis1 (red) was analysed at embryonic (E13.5), foetal (E17.5) and adult time points. Glis1 was detected during valve morphogenesis in mice (E.13 and E.17 stages), specifically during the completion of endothelial to mesenchymal transformation and valve sculpting and elongation and undetected in the adult valve (6 months). Glis1 is detected in nuclei from endothelial and valvular interstitial cells. Green tags are for MF20 marking sarcomeric myosin-myocytes, Blue is Hoescht coloration that indicates nuclei.

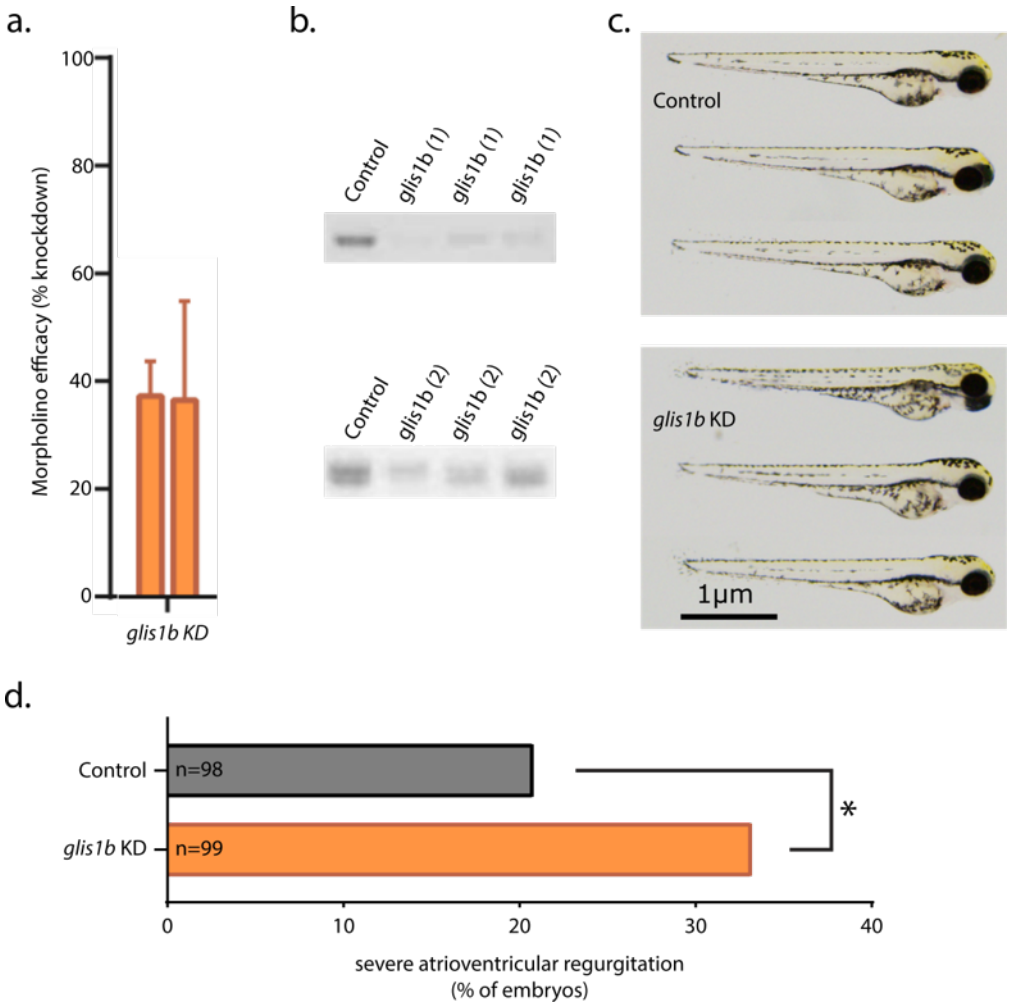
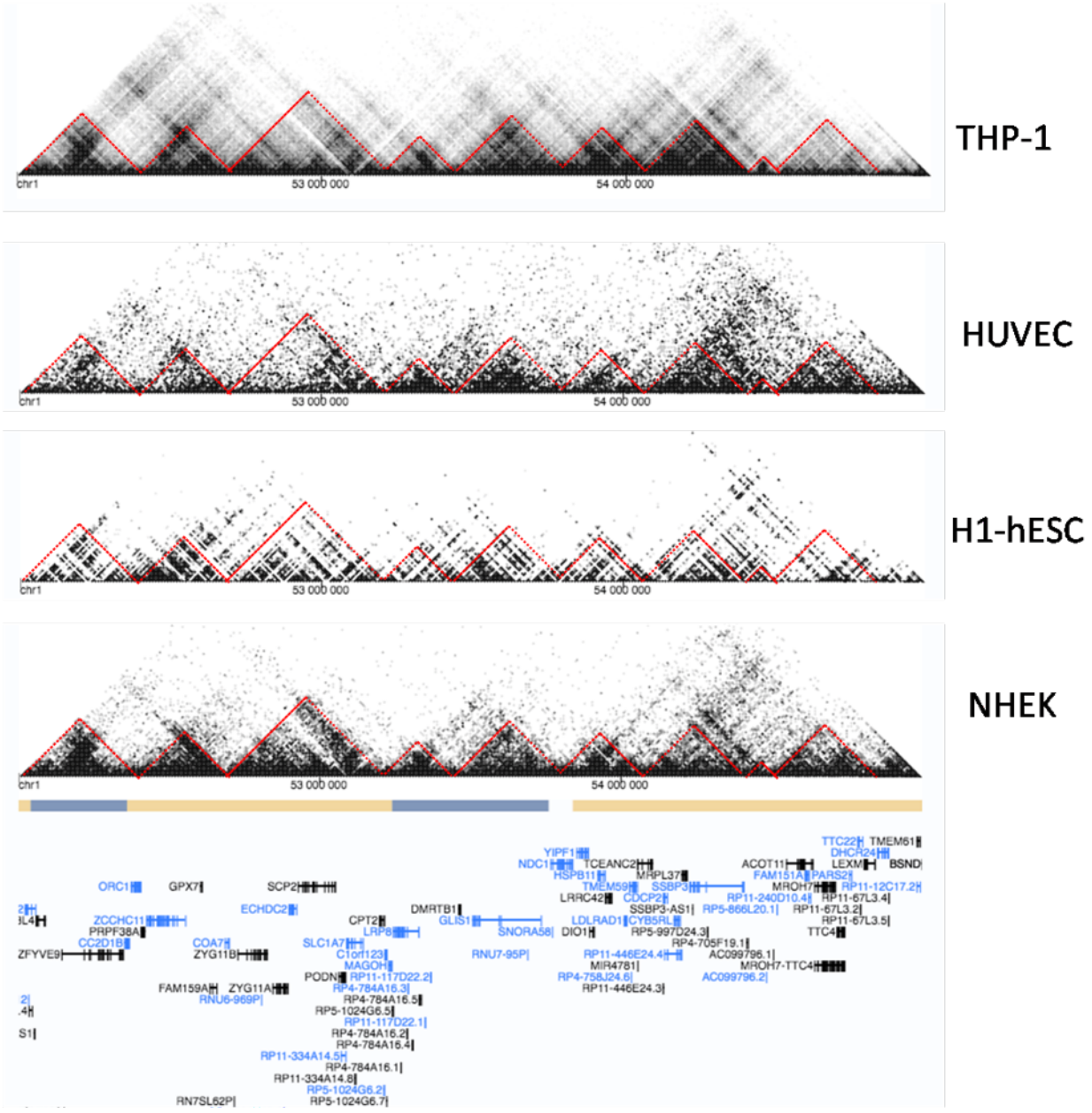
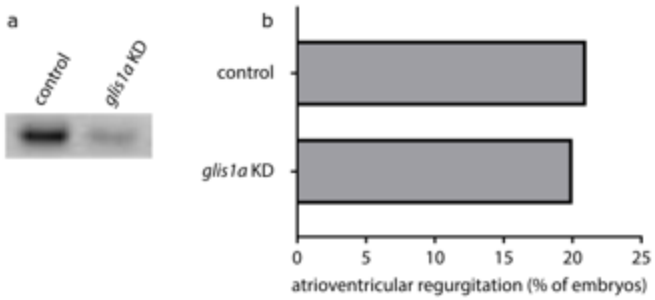


Figure 4. Assessment of cardiac regurgitation in zebrafish after morpholino knockdown of *glis1b*. **(a)** Morpholino-mediated knockdown efficacy. Efficacy for glis1b in embryonic zebrafish was measured by RT-PCR. **(b)** Representative gel images from analysis of morpholino efficacy. Control indicates samples amplified from control-injected embryos. All samples were obtained from 72-hpf embryos. **(c)**. Brightfield micrographs displaying gross morphology of 72-hpf embryos following glis1 knockdown. Scale bar represents 1 µm. **(d)** Fold change in observed atrioventricular regurgitation in 72-hpf zebrafish embryos after morpholino-mediated knockdown. All results are relative to clutchmate controls.



Supplementary Figure 1. Hi-C interaction of GLIS1 in several cell types. We selected Hi-C data from four cell types includes THP-1, HUVEC, H1-hESC, NHEK from ENCODE, gene interaction visualized by Integrated Genome Browser (IGB) was shown. By visualization of the topologically associating domains (TADs), the interaction in the GLIS1 region was found to be mainly exists inside of GLIS1.



Supplementary Figure 2. Assessment of cardiac regurgitation in zebrafish after morpholino knockdown of *glis1a*. **(a)** Representative gel images from analysis of morpholino efficacy. Control indicates samples amplified from control-injected embryos. All samples were obtained from 72-hpf embryos. (**b**) Fold change in observed atrioventricular regurgitation in 72-hpf zebrafish embryos after morpholino-mediated knockdown.

Supplementary Table 1) Enriched pathways for nominally associated loci in MVP GWAS according to the SNP-ratio test (SRT) method. Enriched KEGG pathways are those for which the 1000 simulated GWAS with permuted case control status provided smaller counts of ratios of significant to non significant SNPs, when compared to GWAS with real phenotyes. We provide the count of ratios and empirical p-values (<0.05) for pathways. E.g Compared to the real GWAS, SNPs in genes in the phagosome pathway presented more significant associations with MVP only in 1 out of 1000 simulated GWAS. On the other hand, SNPs in genes from the Wnt signaling pathway presented more significant associations with MVP 50 simulated GWAS.

Pathway ID in KEGG	Descriptive of the pathway	Count of ratios of significant to non-significant SNPs higher in simulated data	Empirical P value
hsa04145	Phagosome	1	0.001
hsa05416	**Viral myocarditis**	1	0.001
hsa05410	**Hypertrophic cardiomyopathy (HCM)**	2	0.002
hsa05414	**Dilated cardiomyopathy**	2	0.002
hsa00190	Oxidative phosphorylation	3	0.003
hsa05340	Primary immunodeficiency	3	0.003
hsa00010	Glycolysis / Gluconeogenesis	4	0.004
hsa04612	Antigen processing and presentation	4	0.004
hsa00051	Fructose and mannose metabolism	5	0.005
hsa04260	**Cardiac muscle contraction**	5	0.005
hsa04672	Intestinal immune network for IgA production	5	0.005
hsa05140	Leishmaniasis	5	0.005
hsa05310	Asthma	5	0.005
hsa05330	Allograft rejection	5	0.005
hsa05332	Graft-versus-host disease	5	0.005
hsa01100	Metabolic pathways	6	0.006
hsa04070	**Phosphatidylinositol signaling system**	6	0.006
hsa00562	Inositol phosphate metabolism	7	0.007
hsa05320	Autoimmune thyroid disease	7	0.007
hsa04662	B cell receptor signaling pathway	8	0.008
hsa04940	Type I diabetes mellitus	8	0.008
hsa05210	Colorectal cancer	10	0.010
hsa05144	Malaria	11	0.011
hsa05322	Systemic lupus erythematosus	11	0.011
hsa04660	T cell receptor signaling pathway	12	0.012
hsa00740	Riboflavin metabolism	15	0.015
hsa00750	Vitamin B6 metabolism	15	0.015
hsa04370	VEGF signaling pathway	16	0.016
hsa04930	Type II diabetes mellitus	18	0.018
hsa04962	Vasopressin-regulated water reabsorption	18	0.018
hsa00524	Butirosin and neomycin biosynthesis	25	0.025
hsa04010	**MAPK signaling pathway**	30	0.030
hsa00920	Sulfur metabolism	31	0.031
hsa00603	Glycosphingolipid biosynthesis - globo series	36	0.036
hsa04270	Vascular smooth muscle contraction	36	0.036
hsa00350	Tyrosine metabolism	37	0.037
hsa04540	**Gap junction**	41	0.041
hsa04650	Natural killer cell mediated cytotoxicity	45	0.045
hsa05215	Prostate cancer	45	0.045
hsa00970	Aminoacyl-tRNA biosynthesis	48	0.048
hsa04966	Collecting duct acid secretion	48	0.048
hsa04310	**Wnt signaling pathway**	50	0.050

Supplementary Table 2. Enriched gene sets identified by i-GSEA4GWAS. We display [illegible] significantly enriched pathways/gene sets (FDR < 0.05). We also added the Hedgehog Signaling Pathway result. The number of significantly contributing genes and the total number of genes analyzed are indicated. A short gene set description was provided when available, in addition to the catalog IDs (GO and KEGG). FDR: false discovery rate.

Gene Set Name	Significant genes	Total genes analyzed	P Value	FDR	Gene Set Description
Biocarta: EDG1 Pathway	26	27	0.001	1.00E-03	Phospholipids as signalling intermediaries
GO: Actin Binding	58	76	0.001	1.00E-03	GO:0003779. Interacting selectively with monomeric or multimeric forms of actin, including actin filaments.
GO: Adherens Junction	18	23	0.001	1.00E-03	GO:0005912. A cell junction at which the cytoplasmic face of the plasma membrane is attached to actin filaments.
GO: Anion Transport	24	31	0.001	1.00E-03	GO:0006820. The directed movement of anions, atoms or small molecules with a net negative charge, into, out of, within or between cells.
GO: Basolateral Plasma Membrane	28	35	0.001	1.00E-03	GO:0016323. Part of the plasma membrane that includes the basal end and sides of the cell. Often used in reference to animal polarized epithelial membranes, where the basal membrane is the part attached to the extracellular matrix, or in plant cells, where the basal membrane is defined with respect to the zygotic axis.
GO: Transcription Repressor Activity	94	148	0.001	1.00E-03	GO:0016564. Any transcription regulator activity that prevents or downregulates transcription.
GO: Vasculature Development	41	55	0.001	1.00E-03	GO:0001944. The process whose specific outcome is the progression of the vasculature over time, from its formation to the mature structure.
GO: Cytoskeletal Protein Binding	112	159	0.001	1.10E-03	GO:0008092. Interacting selectively with any protein component of any cytoskeleton (actin, microtubule, or intermediate filament cytoskeleton).
KEGG: Intestinal Immune Network for IGA Production	31	45	0.001	1.11E-03	hsa04672
GO: Ion Transport	131	180	0.001	1.13E-03	GO:0006811. The directed movement of charged atoms or small charged molecules into, out of, within or between cells.
Biocarta: Integrin Pathway	30	38	0.001	1.18E-03	Integrin Signaling Pathway
KEGG: Dilated Cardiomyopathy	68	90	0.001	1.31E-03	hsa05414
GO: Structural Constituent of Cytoskeleton	39	57	0.001	1.33E-03	GO:0005200. The action of a molecule that contributes to the structural integrity of a cytoskeletal structure.
GO: Actin Filament Binding	21	25	0.001	1.35E-03	GO:0051015. Interacting selectively with an actin filament, also known as F-actin, a helical filamentous polymer of globular G-actin subunits.
GO: Sulfur Metabolic Process	31	37	0.001	1.36E-03	GO:0006790. The chemical reactions and pathways involving the nonmetallic element sulfur or compounds that contain sulfur, such as the amino acids methionine and cysteine or the tripeptide glutathione.
GO: Angiogenesis	36	48	0.001	1.37E-03	GO:0001525. Blood vessel formation when new vessels emerge from the proliferation of pre-existing blood vessels.
GO: Cell Cortex	30	39	0.001	1.38E-03	GO:0005938. The region of a cell that lies just beneath the plasma membrane and often, but not always, contains a network of actin filaments and associated proteins.
KEGG: Cell Adhesion Molecules Cams	91	127	0.001	1.38E-03	hsa04514
GO: Positive Regulation of Cell Proliferation	97	145	0.001	1.39E-03	GO:0008284. Any process that activates or increases the rate or extent of cell proliferation.
GO: Anion Transmembrane Transporter Activity	43	59	0.001	1.39E-03	GO:0008509. Catalysis of the transfer of a negatively charged ion from one side of a membrane to the other.
GO: Anatomical Structure formation	42	56	0.001	1.40E-03	GO:0048646. The process pertaining to the initial formation of an anatomical structure from unspecified parts. This process begins with the specific processes that contribute to the appearance of the discrete structure and ends when the structural rudiment is recognizable. An anatomical structure is any biological entity that occupies space and is distinguished from its surroundings. Anatomical structures can be macroscopic such as a carpel, or microscopic such as an acrosome.
Biocarta: SPPA Pathway	20	22	0.001	1.40E-03	Aspirin Blocks Signaling Pathway Involved in Platelet Activation
KEGG: Focal Adhesion	145	198	0.001	1.41E-03	hsa04510
GO: Transcription Corepressor Activity	58	90	0.001	1.42E-03	GO:0003714. The function of a transcription cofactor that represses transcription from a RNA polymerase II promoter; does not bind DNA itself.
GO: DI TRI Valent Inorganic Cation Transmembrane Transporter Activity	18	22	0.001	1.43E-03	GO:0015082. Catalysis of the transfer of inorganic cations with a valency of two or three from one side of the membrane to the other. Inorganic cations are atoms or small molecules with a positive charge that do not contain carbon in covalent linkage.
GO: Positive Regulation of Transcription DNA Dependent	82	118	0.001	1.44E-03	GO:0045893. Any process that activates or increases the frequency, rate or extent of DNA-dependent transcription.
GO: Substrate Specific Channel Activity	108	154	0.001	1.45E-03	GO:0022838. Catalysis of energy-independent facilitated diffusion, mediated by passage of a specific solute through a transmembrane aqueous pore or channel. Stereospecificity is not exhibited but this transport may be specific for a particular molecular species or class of molecules.
GO: Ion Channel Activity	104	147	0.001	1.46E-03	GO:0005216. Catalysis of facilitated diffusion of an ion (by an energy-independent process) by passage through a transmembrane aqueous pore or channel without evidence for a carrier-mediated mechanism.
KEGG: Pentose Phosphate Pathway	20	25	0.001	1.55E-03	hsa00030
GO: Cell Cortex Part	19	24	0.001	1.55E-03	GO:0044448. Any constituent part of the cell cortex, the region of a cell that lies just beneath the plasma membrane and often, but not always, contains a network of actin filaments and associated proteins.
GO: Transcription Activator Activity	112	170	0.001	1.56E-03	GO:0016563. Any transcription regulator activity required for initiation or upregulation of transcription.
GO: Transmembrane Receptor Protein Kinase Activity	41	51	0.001	1.57E-03	GO:0019199.
GO: Positive Regulation of RNA Metabolic Process	83	120	0.001	1.57E-03	GO:0051254. Any process that activates or increases the frequency, rate or extent of the chemical reactions and pathways involving RNA.
GO: Metal Ion Transmembrane Transporter Activity	105	145	0.001	1.59E-03	GO:0046873. Catalysis of the transfer of metal ions from one side of a membrane to the other.
GO: mRNA Metabolic Process	43	82	0.001	1.59E-03	GO:0016071. The chemical reactions and pathways involving mRNA, messenger RNA, which is responsible for carrying the coded genetic 'message', transcribed from DNA, to sites of protein assembly at the ribosomes.
GO: Cytoskeleton Dependent Intracellular Transport	20	26	0.001	1.62E-03	GO:0030705. The directed movement of substances along cytoskeletal elements such as microfilaments or microtubules within a cell.
GO: Sulfotransferase Activity	23	28	0.001	1.64E-03	GO:0008146. Catalysis of the transfer of a sulfate group from 3'-phosphoadenosine 5'-phosphosulfate to the hydroxyl group of an acceptor, producing the sulfated derivative and 3'-phosphoadenosine 5'-phosphate.
GO: mRNA Processing GO 0006397	37	72	0.001	1.68E-03	GO:0006397. Any process involved in the conversion of a primary mRNA transcript into one or more mature mRNA(s) prior to translation into polypeptide.
GO: Enzyme Linked Receptor Protein Signaling Pathway	108	140	0.001	1.74E-03	GO:0007167. Any series of molecular signals initiated by the binding of an extracellular ligand to a receptor on the surface of the target cell, where the receptor possesses catalytic activity or is closely associated with an enzyme such as a protein kinase.

GO: Hemopoietic or Lymphoid Organ Development	53	73	0.001	1.83E-03	GO:0048534. The process whose specific outcome is the progression of any organ involved in hemopoiesis or lymphoid cell activation over time, from its formation to the mature structure. Such development includes differentiation of resident cell types (stromal cells) and of migratory cell types dependent on the unique microenvironment afforded by the organ for their proper differentiation.
GO: Lyase Activity	46	68	0.001	1.95E-03	GO:0016829. Catalysis of the cleavage of C-C, C-O, C-N and other bonds by other means than by hydrolysis or oxidation, or conversely adding a group to a double bond. They differ from other enzymes in that two substrates are involved in one reaction direction, but only one in the other direction. When acting on the single substrate, a molecule is eliminated and this generates either a new double bond or a new ring.
KEGG: Hypertrophic Cardiomyopathy HCM	60	83	0.001	1.98E-03	hsa05410
GO: Transmembrane Receptor Protein Tyrosine Kinase Activity	36	43	0.001	1.98E-03	GO:0004714. Catalysis of the reaction: ATP + a protein-L-tyrosine = ADP + a protein-L-tyrosine phosphate, to initiate a change in cell activity.
GO: Striated Muscle Development	26	40	0.001	2.00E-03	GO:0014706. The process whose specific outcome is the progression of a striated muscle over time, from its formation to the mature structure. Striated muscle contain fibers that are divided by transverse bands into striations, and cardiac and skeletal muscle are types of striated muscle. Skeletal muscle myoblasts fuse to form myotubes and eventually multinucleated muscle fibers. The fusion of cardiac cells is very rare and can only form binucleate cells.
GO: Cation Transport	102	143	0.001	2.00E-03	GO:0006812. The directed movement of cations, atoms or small molecules with a net positive charge, into, out of, within or between cells.
GO: Regulation of G Protein Coupled Receptor Protein Signaling Pathway	19	22	0.001	2.02E-03	GO:0008277. Any process that modulates the frequency, rate or extent of G-protein coupled receptor protein signaling pathway activity.
KEGG: Oxidative Phosphorylation	62	113	0.001	2.15E-03	hsa00190
GO: Regulation of Neurotransmitter Levels	17	24	0.001	2.19E-03	GO:0001505. Any process that modulates levels of neurotransmitter.
GO: Endosome Transport	17	23	0.001	2.61E-03	GO:0016197. The directed movement of substances into, out of or mediated by an endosome, a membrane-bound organelle that carries materials newly ingested by endocytosis. It passes many of the materials to lysosomes for degradation.
GO: Organ Morphogenesis	99	144	0.001	2.64E-03	GO:0009887. Morphogenesis of an organ. An organ is defined as a tissue or set of tissues that work together to perform a specific function or functions. Morphogenesis is the process by which anatomical structures are generated and organized. Organs are commonly observed as visibly distinct structures, but may also exist as loosely associated clusters of cells that work together to perform a specific function or functions.
GO: Regulation of Angiogenesis	20	26	0.001	2.75E-03	GO:0045765. Any process that modulates the frequency, rate or extent of angiogenesis.
GO: Deoxyribonuclease Activity	14	22	0.001	2.76E-03	GO:0004536. Catalysis of the hydrolysis of ester linkages within deoxyribonucleic acid.
GO: Immune System Development	55	79	0.001	2.77E-03	GO:0002520. The process whose specific outcome is the progression of an organismal system whose objective is to provide calibrated responses by an organism to a potential internal or invasive threat, over time, from its formation to the mature structure. A system is a regularly interacting or interdependent group of organs or tissues that work together to carry out a given biological process.
GO: DNA Catabolic Process	18	23	0.001	2.78E-03	GO:0006308. The chemical reactions and pathways resulting in the breakdown of DNA, deoxyribonucleic acid, one of the two main types of nucleic acid, consisting of a long unbranched macromolecule formed from one or two strands of linked deoxyribonucleotides, the 3'-phosphate group of each constituent deoxyribonucleotide being joined in 3',5'-phosphodiester linkage to the 5'-hydroxyl group of the deoxyribose moiety of the next one.
GO: Hemopoiesis	51	73	0.001	2.80E-03	GO:0030097. The process whose specific outcome is the progression of the myeloid and lymphoid derived organ/tissue systems of the blood and other parts of the body over time, from formation to the mature structure. The site of hemopoiesis is variable during development, but occurs primarily in bone marrow or kidney in many adult vertebrates.
GO: Structural Constituent of Muscle	22	33	0.002	2.84E-03	GO:0008307. The action of a molecule that contributes to the structural integrity of a muscle fiber.
GO: Detection of Stimulus	35	46	0.001	2.84E-03	GO:0051606. The series of events in which a stimulus is received by a cell and converted into a molecular signal.
GO: Actin Cytoskeleton	84	128	0.001	2.88E-03	GO:0015629. The part of the cytoskeleton (the internal framework of a cell) composed of actin and associated proteins. Includes actin cytoskeleton-associated complexes.
KEGG: Tight Junction	85	127	0.001	3.00E-03	hsa04530
Biocarta: MPR Pathway	25	34	0.001	3.02E-03	How Progesterone Initiates Oocyte Membrane
GO: Positive Regulation of Transcription	97	143	0.001	3.24E-03	GO:0045941. Any process that activates or increases the frequency, rate or extent of transcription.
GO: Di-, Tri-valent Inorganic Cation Transport	26	32	0.001	3.25E-03	GO:0015674. The directed movement of inorganic cations with a valency of two or three into, out of, within or between cells. Inorganic cations are atoms or small molecules with a positive charge which do not contain carbon in covalent linkage.
KEGG: Cardiac Muscle Contraction	48	72	0.001	3.33E-03	hsa04260
GO: RNA Processing	88	166	0.001	3.45E-03	GO:0006396. Any process involved in the conversion of one or more primary RNA transcripts into one or more mature RNA molecules.
GO: Sensory Perception	123	185	0.001	3.80E-03	GO:0007600. The series of events required for an organism to receive a sensory stimulus, convert it to a molecular signal, and recognize and characterize the signal.
Biocarta: METPathway	29	37	0.001	3.81E-03	Signaling of Hepatocyte Growth Factor Receptor
GO: Phosphoric Ester Hydrolase Activity	103	149	0.001	3.85E-03	GO:0042578. Catalysis of the reaction: RPO-R' + H2O = RPOOH + R'H. This reaction is the hydrolysis of any phosphoric ester bond, any ester formed from orthophosphoric acid, O=P(OH)3.
GO: Collagen	21	23	0.001	3.85E-03	GO:0005581. Any of the various assemblies in which collagen chains form a left-handed triple helix; may assemble into higher order structures.
GO: Calcium Ion Transport	22	27	0.001	3.88E-03	GO:0006816. The directed movement of calcium (Ca) ions into, out of, within or between cells.
Biocarta: PDINS Pathway	17	23	0.002	3.89E-03	Phosphoinositides and their downstream targets.
GO: Positive Regulation of Nucleobasenucleosidenucleotide and Nucleic Acid Metabolic Process	103	153	0.001	4.08E-03	GO:0045935. Any process that activates or increases the frequency, rate or extent of the chemical reactions and pathways involving nucleobases, nucleosides, nucleotides and nucleic acids.
GO: Microtubule Based Process	58	81	0.001	4.10E-03	GO:0007017. Any cellular process that depends upon or alters the microtubule cytoskeleton, that part of the cytoskeleton comprising microtubules and their associated proteins.
GO: Phosphoric Monoester Hydrolase Activity	76	108	0.001	4.53E-03	GO:0016791. Catalysis of the hydrolysis of phosphoric monoesters, releasing inorganic phosphate.
GO: Actin Filament Based Process	71	113	0.002	4.66E-03	GO:0030029. Any cellular process that depends upon or alters the actin cytoskeleton, that part of the cytoskeleton comprising actin filaments and their associated proteins.
GO: Regulation of Cell Differentiation	44	60	0.001	4.68E-03	GO:0045595. Any process that modulates the frequency, rate or extent of cell differentiation, the process whereby relatively unspecialized cells acquire specialized structural and functional features.
GO: Regulation of Cytokine Production	18	24	0.001	4.82E-03	GO:0001817. Any process that modulates the frequency, rate, or extent of production of a cytokine.
GO: Cell Migration	69	94	0.001	4.83E-03	GO:0016477. The orderly movement of cells from one site to another, often during the development of a multicellular organism.
GO: Kinase Inhibitor Activity	16	25	0.001	4.84E-03	GO:0019210. Stops, prevents or reduces the activity of a kinase, an enzyme which catalyzes of the transfer of a phosphate group, usually from ATP, to a substrate molecule.

GO: Sterol Metabolic Process	18	25	0.001	4.80E-03	GO:0008202. The chemical reactions and pathways involving steroids, compounds with a 1,2-cyclopentanoperhydrophenanthrene nucleus.
Biocarta: Mcalpain Pathway	19	25	0.001	4.90E-03	mCalpain and friends in Cell motility
GO: Inorganic Cation Transmembrane Transporter Activity	39	55	0.001	4.90E-03	GO:0022890. Catalysis of the transfer of inorganic cations from one side of a membrane to the other. Inorganic cations are atoms or small molecules with a positive charge that do not contain carbon in covalent linkage.
GO: DNA Damage Responsesignal Transduction	23	33	0.001	4.91E-03	GO:0042770. A cascade of processes induced by the detection of DNA damage within a cell.
KEGG: Leukocyte Transendothelial Migration	73	113	0.002	5.13E-03	hsa04670
KEGG: Glycine Serine and Threonine Metabolism	18	31	0.001	5.14E-03	hsa00260
GO: Transferase Activity Transferring Sulfur Containing Groups	25	32	0.002	5.26E-03	GO:0016782. Catalysis of the transfer of a sulfur-containing group from one compound (donor) to another (acceptor).
GO: Mitochondrial Part	69	139	0.003	5.41E-03	GO:0044429. Any constituent part of a mitochondrion, a semiautonomous, self replicating organelle that occurs in varying numbers, shapes, and sizes in the cytoplasm of virtually all eukaryotic cells. It is notably the site of tissue respiration.
Biocarta: ECM Pathway	19	24	0.001	5.44E-03	Erk and PI-3 Kinase Are Necessary for Collagen Binding in Corneal Epithelia
KEGG: Phosphatidylinositol Signaling System	50	76	0.001	5.58E-03	hsa04070
GO: Lipid Transporter Activity	21	27	0.001	5.60E-03	GO:0005319. Enables the directed movement of lipids into, out of, within or between cells.
GO: Sequence Specific DNA Binding	41	58	0.001	5.63E-03	GO:0043565. Interacting selectively with DNA of a specific nucleotide composition, e.g. GC-rich DNA binding, or with a specific sequence motif or type of DNA e.g. promotor binding or rDNA binding.
GO: Heart Development	26	37	0.001	5.66E-03	GO:0007507. The process whose specific outcome is the progression of the heart over time, from its formation to the mature structure. The heart is a hollow, muscular organ, which, by contracting rhythmically, keeps up the circulation of the blood.
KEGG: Glycerophospholipid Metabolism	50	76	0.002	5.66E-03	hsa00564
GO: Secondary Metabolic Process	16	26	0.001	5.67E-03	GO:0019748. The chemical reactions and pathways resulting in many of the chemical changes of compounds that are not necessarily required for growth and maintenance of cells, and are often unique to a taxon. In multicellular organisms secondary metabolism is generally carried out in specific cell types, and may be useful for the organism as a whole. In unicellular organisms, secondary metabolism is often used for the production of antibiotics or for the utilization and acquisition of unusual nutrients.
GO: Positive Regulation of Cell Differentiation	18	25	0.001	5.71E-03	GO:0045597. Any process that activates or increases the frequency, rate or extent of cell differentiation.
GO: Alcohol Metabolic Process	56	87	0.001	5.95E-03	GO:0006066. The chemical reactions and pathways involving alcohols, any of a class of alkyl compounds containing a hydroxyl group.
GO: Response to UV	17	25	0.001	6.00E-03	GO:0009411. A change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of an ultraviolet radiation (UV light) stimulus. Ultraviolet radiation is electromagnetic radiation with a wavelength in the range of 10 to 380 nanometers.
GO: Gated Channel Activity	87	121	0.001	6.01E-03	GO:0022836. Catalysis of the transmembrane transfer of a solute by a channel that opens in response to a specific stimulus.
KEGG: Adherens Junction	56	72	0.001	6.04E-03	hsa04520
KEGG: Mtor Signaling Pathway	36	52	0.001	6.08E-03	hsa04150
GO: Metal Ion Transport	81	115	0.001	6.10E-03	GO:0030001. The directed movement of metal ions, any metal ion with an electric charge, into, out of, within or between cells.
Biocarta: Insulin Pathway	15	22	0.001	6.15E-03	Insulin Signaling Pathway
GO: Calcium Channel Activity	26	33	0.001	6.19E-03	GO:0005262. Catalysis of facilitated diffusion of an calcium (by an energy-independent process) involving passage through a transmembrane aqueous pore or channel without evidence for a carrier-mediated mechanism.
GO: Fatty Acid Metabolic Process	41	63	0.001	6.27E-03	GO:0006631. The chemical reactions and pathways involving fatty acids, aliphatic monocarboxylic acids liberated from naturally occurring fats and oils by hydrolysis.
GO: G Protein Coupled Receptor Activity	131	184	0.001	6.30E-03	GO:0004930. A receptor that binds an extracellular ligand and transmits the signal to a heterotrimeric G-protein complex. These receptors are characteristically seven-transmembrane receptors and are made up of hetero- or homodimers.
KEGG: TGF-Beta Signaling Pathway	58	84	0.001	6.38E-03	hsa04350
GO: Protein C Terminus Binding	48	73	0.001	6.40E-03	GO:0008022. Interacting selectively with a protein C-terminus, the end of any peptide chain at which the 1-carboxy function of a constituent amino acid is not attached in peptide linkage to another amino-acid residue.
GO: Transmembrane Receptor Protein Tyrosine Kinase Signaling Pathway	65	83	0.001	7.21E-03	GO:0007169. The series of molecular signals generated as a consequence of a transmembrane receptor tyrosine kinase binding to its physiological ligand.
Biocarta: NOS1 Pathway	15	21	0.001	7.24E-03	Nitric Oxide Signaling Pathway
GO: Cell Cycle Checkpoint	27	47	0.001	7.37E-03	GO:0000075. A point in the eukaryotic cell cycle where progress through the cycle can be halted until conditions are suitable for the cell to proceed to the next stage.
Biocarta: Inflam Pathway	17	29	0.002	7.44E-03	Cytokines and Inflammatory Response
GO: G Protein Coupled Receptor Binding	31	54	0.004	7.50E-03	GO:0001664. Interacting selectively with a G-protein-coupled receptor.
KEGG: Glycerolipid Metabolism	36	49	0.001	8.01E-03	hsa00561
GO: Activation of Protein Kinase Activity	19	26	0.002	8.04E-03	GO:0032147. Any process that initiates the activity of an inactive protein kinase.
GO: Glucose Metabolic Process	17	28	0.002	8.28E-03	GO:0006006. The chemical reactions and pathways involving glucose, the aldohexose gluco-hexose. D-glucose is dextrorotatory and is sometimes known as dextrose; it is an important source of energy for living organisms and is found free as well as combined in homo- and hetero-oligosaccharides and polysaccharides.
KEGG: Glycosaminoglycan Biosynthesis Heparan Sulfate	21	26	0.001	8.33E-03	hsa00534
Biocarta: CSKPathway	17	22	0.002	8.38E-03	Activation of Csk by cAMP-dependent Protein Kinase Inhibits Signaling through the T Cell Receptor
KEGG: Ether Lipid Metabolism	22	33	0.001	8.47E-03	hsa00565
GO: Regulation of Binding	39	57	0.002	8.53E-03	GO:0051098. Any process that modulates the frequency, rate or extent of binding, the selective interaction of a molecule with one or more specific sites on another molecule.
GO: Protein Kinase Inhibitor Activity	15	24	0.001	8.87E-03	GO:0004860. Stops, prevents or reduces the activity of a protein kinase, an enzyme which phosphorylates a protein.
KEGG: Valine Leucine and Isoleucine Degradation	30	44	0.001	8.98E-03	hsa00280
GO: Muscle Development	55	93	0.001	9.01E-03	GO:0007517. The process whose specific outcome is the progression of the muscle over time, from its formation to the mature structure. The muscle is an organ consisting of a tissue made up of various elongated cells that are specialized to contract and thus to produce movement and mechanical work.
GO: Glycosaminoglycan Binding	24	33	0.002	9.19E-03	GO:0005539. Interacting selectively with any glycan (polysaccharide) containing a substantial proportion of aminomonosaccharide residues.
GO: Positive Regulation of Transcription From RNA Polymerase ii Promoter	43	65	0.001	9.24E-03	GO:0045944. Any process that activates or increases the frequency, rate or extent of transcription from the RNA polymerase II promoter.
GO: Envelope	91	164	0.004	9.24E-03	GO:0031975. A multilayered structure surrounding all or part of a cell; encompasses one or more lipid bilayers, and may include a cell wall layer, also includes the space between layers.
GO: Organelle Envelope	91	164	0.004	9.24E-03	GO:0031967. A double membrane structure enclosing an organelle, including two lipid bilayers and the region between them. In some cases, an organelle envelope may have more than two membranes.

GO: Protein Serine/Threonine Phosphatase Activity	15	22	0.002	9.15E-03	GO:0004722. Catalysis of the reaction: protein serine/threonine phosphate + H2O = protein serine/threonine + phosphate.
KEGG: Insulin Signaling Pathway	84	135	0.001	9.26E-03	hsa04910
Biocarta: PPATA Pathway	41	57	0.002	9.28E-03	Mechanism of Gene Regulation by Peroxisome Proliferators via PPARa(alpha)
GO: Negative Regulation of Cell Differentiation	22	28	0.002	9.56E-03	GO:0045596. Any process that stops, prevents or reduces the frequency, rate or extent of cell differentiation.
Biocarta: NO1 Pathway	22	30	0.002	9.81E-03	Actions of Nitric Oxide in the Heart
GO: DNA Integrity Checkpoint	14	23	0.001	9.92E-03	GO:0031570. Any cell cycle checkpoint that delays or arrests cell cycle progression in response to changes in DNA structure.
GO: Response to Hormone Stimulus	22	32	0.003	1.05E-02	GO:0009725. A change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a hormone stimulus.
GO: Membrane Organization and Biogenesis	90	133	0.003	1.10E-02	GO:0016044. A process that is carried out at the cellular level which results in the formation, arrangement of constituent parts, or disassembly of membranes inside and surrounding the cell.
GO: Negative Regulation of Transcription	115	184	0.002	1.14E-02	GO:0016481. Any process that stops, prevents or reduces the frequency, rate or extent of transcription.
GO: Contractile Fiber	15	25	0.005	1.14E-02	GO:0043292. Fibers, composed of actin, myosin, and associated proteins, found in cells of smooth or striated muscle.
GO: Polysaccharide Binding	25	35	0.001	1.15E-02	GO:0030247. Interacting selectively with any polysaccharide.
GO: Actin Filament Organization	13	23	0.003	1.25E-02	GO:0007015. Control of the spatial distribution of actin filaments; includes organizing filaments into meshworks, bundles, or other structures, as by cross-linking.
GO: Cellular Protein Complex Assembly	22	32	0.004	1.26E-02	GO:0043623. The aggregation, arrangement and bonding together of a set of components to form a protein complex, occurring at the level of an individual cell.
GO: Actin Cytoskeleton Organization and Biogenesis	64	103	0.004	1.26E-02	GO:0030036. A process that is carried out at the cellular level which results in the formation, arrangement of constituent parts, or disassembly of cytoskeletal structures comprising actin filaments and their associated proteins.
GO: Transforming Growth Factor Beta Receptor Signaling Pathway	27	36	0.002	1.27E-02	GO:0007179. The series of molecular signals generated as a consequence of a transforming growth factor beta receptor binding to one of its physiological ligands.
GO: Response to Abiotic Stimulus	52	83	0.004	1.27E-02	GO:0009628. A change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of an abiotic (non-living) stimulus.
GO: Regulation of Heart Contraction	18	25	0.002	1.27E-02	GO:0008016. Any process that modulates the frequency, rate or extent of heart contraction. Heart contraction is the process by which the heart decreases in volume in a characteristic way to propel blood through the body.
GO: Transcription Coactivator Activity	75	120	0.003	1.27E-02	GO:0003713. The function of a transcription cofactor that activates transcription from a RNA polymerase II promoter; does not bind DNA itself.
GO: Regulation of Cell Migration	22	28	0.003	1.28E-02	GO:0030334. Any process that modulates the frequency, rate or extent of cell migration.
GO: Kinase Regulator Activity	29	45	0.002	1.28E-02	GO:0019207. Modulates the activity of a kinase, an enzyme which catalyzes of the transfer of a phosphate group, usually from ATP, to a substrate molecule.
GO: Response to Light Stimulus	28	44	0.005	1.28E-02	GO:0009416. A change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of a light stimulus, electromagnetic radiation of wavelengths classified as infrared, visible or ultraviolet light.
GO: Small Protein Conjugating Enzyme Activity	36	52	0.002	1.29E-02	GO:0008639. Catalysis of the covalent attachment of small proteins, such as ubiquitin or ubiquitin-like proteins, to lysine residues on a target protein. This function may be performed alone or in conjunction with an E3, ubiquitin-like protein ligase.
GO: Voltage Gated Channel Activity	52	73	0.002	1.39E-02	GO:0022832. Catalysis of the transmembrane transfer of a solute by a channel whose open state is dependent on the voltage across the membrane in which it is embedded.
GO: Mitosis	48	80	0.004	1.39E-02	GO:0007067. Progression through mitosis, the division of the eukaryotic cell nucleus to produce two daughter nuclei that, usually, contain the identical chromosome complement to their mother.
GO: Electron Carrier Activity	49	78	0.001	1.40E-02	GO:0009055. Any molecular entity that serves as an electron acceptor and electron donor in an electron transport system.
KEGG: Purine Metabolism	95	154	0.004	1.40E-02	hsa00230
KEGG: Fructose and Mannose Metabolism	21	33	0.006	1.47E-02	hsa00051
GO: Rhodopsin Like Receptor Activity	90	128	0.002	1.47E-02	GO:0001584. A G-protein coupled receptor that is structurally/functionally related to the rhodopsin receptor.
GO: Contractile Fiber Part	14	23	0.006	1.52E-02	GO:0044449. Any constituent part of a contractile fiber, a fiber composed of actin, myosin, and associated proteins, found in cells of smooth or striated muscle.
GO: Carbohydrate Biosynthetic Process	39	49	0.002	1.53E-02	GO:0016051. The chemical reactions and pathways resulting in the formation of carbohydrates, any of a group of organic compounds based of the general formula Cx(H2O)y.
GO: Cation Channel Activity	83	118	0.002	1.55E-02	GO:0005261. Catalysis of the energy-independent passage of cations across a lipid bilayer down a concentration gradient.
GO: Phosphoprotein Phosphatase Activity	56	79	0.002	1.59E-02	GO:0004721. Catalysis of the reaction: a phosphoprotein + H2O = a protein + phosphate. Together with protein kinases, these enzymes control the state of phosphorylation of cell proteins and thereby provide an important mechanism for regulating cellular activity.
KEGG: Arrhythmogenic Right Ventricular Cardiomyopathy Arvc	55	74	0.001	1.76E-02	hsa05412
GO: Nuclease Activity	30	55	0.004	1.77E-02	GO:0004518. Catalysis of the hydrolysis of ester linkages within nucleic acids.
KEGG: Vascular Smooth Muscle Contraction	76	115	0.003	1.78E-02	hsa04270
GO: Regulation of Organelle Organization and Biogenesis	27	39	0.003	1.83E-02	GO:0033043. Any process that modulates the frequency, rate or extent of the processes involved in the formation, arrangement of constituent parts, or disassembly of an organelle.
Biocarta: Ceramide Pathway	17	22	0.005	1.85E-02	Ceramide Signaling Pathway
GO: Mitochondrial Envelope	46	94	0.005	1.86E-02	GO:0005740. The double lipid bilayer enclosing the mitochondrion and separating its contents from the cell cytoplasm; includes the intermembrane space.
GO: Mitochondrial Lumen	24	46	0.006	1.89E-02	GO:0031980. The volume enclosed by the mitochondrial inner membrane.
GO: Mitochondrial Matrix	24	46	0.006	1.89E-02	GO:0005759. The gel-like material, with considerable fine structure, that lies in the matrix space, or lumen, of a mitochondrion. It contains the enzymes of the tricarboxylic acid cycle and, in some organisms, the enzymes concerned with fatty-acid oxidation.
KEGG: Sphingolipid Metabolism	27	40	0.008	2.02E-02	hsa00600
GO: Excretion	23	36	0.006	2.06E-02	GO:0007588. The elimination by an organism of the waste products that arise as a result of metabolic activity. These products include water, carbon dioxide (CO2), and nitrogenous compounds.
GO: Positive Regulation of Multicellular Organismal Process	43	64	0.003	2.06E-02	GO:0051240. Any process that activates or increases the frequency, rate or extent of an organismal process, the processes pertinent to the function of an organism above the cellular level; includes the integrated processes of tissues and organs.
GO: Regulation of Mitosis	25	40	0.006	2.08E-02	GO:0007088. Any process that modulates the frequency, rate or extent of mitosis.
KEGG: Progesterone Mediated Oocyte Maturation	53	84	0.004	2.11E-02	hsa04914
KEGG: Long Term Depression	52	70	0.002	2.21E-02	hsa04730
GO: Monovalent Inorganic Cation Transport	62	90	0.004	2.21E-02	GO:0015672. The directed movement of inorganic cations with a valency of one into, out of, within or between cells. Inorganic cations are atoms or small molecules with a positive charge which do not contain carbon in covalent linkage.
KEGG: Dorso Ventral Axis formation	20	24	0.003	2.22E-02	hsa04320
GO: Growth	50	76	0.004	2.22E-02	GO:0040007. The increase in size or mass of an entire organism, a part of an organism or a cell.

GO: Oxidoreductase Activity Acting on CH-OH Group of Donors	36	62	0.007	2.45E-02	GO:0016614. Catalysis of an oxidation-reduction (redox) reaction in which a CH-OH group act as a hydrogen or electron donor and reduces a hydrogen or electron acceptor.
GO: Regulation of Cell Growth	28	45	0.005	2.45E-02	GO:0001558. Any process that modulates the frequency, rate or extent of cell growth.
GO: Hydro Lyase Activity	17	27	0.006	2.35E-02	GO:0016836. Catalysis of the cleavage of a carbon-oxygen bond by elimination of water.
GO: Lipase Activity	33	50	0.003	2.35E-02	GO:0016298. Catalysis of the hydrolysis of a lipid or phospholipid.
KEGG: Neurotrophin Signaling Pathway	81	126	0.005	2.36E-02	hsa04722
GO: Transmembrane Receptor Protein Serine Threonine Kinase Signaling Pathway	34	47	0.007	2.36E-02	GO:0007178. The series of molecular signals generated as a consequence of a transmembrane receptor serine/threonine kinase binding to its physiological ligand.
Biocarta: Gleevec Pathway	17	23	0.006	2.37E-02	Inhibition of Cellular Proliferation by Gleevec
KEGG: Notch Signaling Pathway	32	47	0.008	2.37E-02	hsa04330
GO: Dephosphorylation	47	69	0.004	2.38E-02	GO:0016311. The process of removing one or more phosphoric (ester or anhydride) residues from a molecule.
Biocarta: NFAT Pathway	35	53	0.005	2.39E-02	NFAT and Hypertrophy of the heart (Transcription in the broken heart)
Biocarta: IGF1 Pathway	15	21	0.004	2.49E-02	IGF-1 Signaling Pathway
Biocarta: TPO Pathway	18	24	0.004	2.49E-02	TPO Signaling Pathway
Biocarta: Chrebp2 Pathway	28	41	0.005	2.50E-02	Regulation And Function Of ChREBP in Liver
Biocarta: BAD Pathway	20	25	0.005	2.55E-02	Regulation of BAD phosphorylation
GO: Positive Regulation of Transport	13	22	0.005	2.56E-02	GO:0051050. Any process that activates or increases the frequency, rate or extent of the directed movement of substances (such as macromolecules, small molecules, ions) into, out of, within or between cells.
GO: Positive Regulation of Protein Modification Process	21	28	0.005	2.56E-02	GO:0031401. Any process that activates or increases the frequency, rate or extent of the covalent alteration of one or more amino acid residues within a protein.
GO: Skeletal Muscle Development	19	31	0.003	2.57E-02	GO:0007519. The developmental sequence of events leading to the formation of adult muscle that occurs in the anima. In vertebrate skeletal muscle the main events are: the fusion of myoblasts to form myotubes that increase in size by further fusion to them of myoblasts, the formation of myofibrils within their cytoplasm and the establishment of functional neuromuscular junctions with motor neurons. At this stage they can be regarded as mature muscle fibers.
GO: Establishment of Protein Localization	108	185	0.007	2.71E-02	GO:0045184. The directed movement of a protein to a specific location.
GO: Transcription Factor Complex	50	88	0.009	2.73E-02	GO:0005667. Any complex, distinct from RNA polymerase, including one or more polypeptides capable of binding DNA at promoters or at cis-acting regulatory sequences, and regulating transcription.
KEGG: Long Term Potentiation	52	69	0.004	2.76E-02	hsa04720
GO: One Carbon Compound Metabolic Process	19	26	0.007	2.79E-02	GO:0006730. The chemical reactions and pathways involving compounds containing a single carbon atom.
GO: Oxidoreductase Activity GO 0016616	31	56	0.01	2.86E-02	GO:0016616. Catalysis of an oxidation-reduction (redox) reaction in which a CH-OH group acts as a hydrogen or electron donor and reduces NAD+ or NADP.
Biocarta: HER2 Pathway	19	22	0.006	2.87E-02	Role of ERBB2 in Signal Transduction and Oncology
GO: Protein Kinase Binding	37	61	0.008	2.93E-02	GO:0019901. Interacting selectively with a protein kinase, any enzyme that catalyzes the transfer of a phosphate group, usually from ATP, to a protein substrate.
Biocarta: DC Pathway	13	22	0.009	2.94E-02	Dendritic cells in regulating TH1 and TH2 Development
GO: Response to Endogenous Stimulus	113	193	0.006	2.95E-02	GO:0009719. A change in state or activity of a cell or an organism (in terms of movement, secretion, enzyme production, gene expression, etc.) as a result of an endogenous stimulus.
GO: Regulation of Transport	40	66	0.005	2.98E-02	GO:0051049. Any process that modulates the frequency, rate or extent of the directed movement of substances (such as macromolecules, small molecules, ions) into, out of, within or between cells.
Biocarta: VEGF Pathway	21	28	0.003	2.99E-02	VEGF, Hypoxia, and Angiogenesis
GO: Oxidoreductase Activity Acting on The CH-CH Group of Donors	15	23	0.008	2.99E-02	GO:0016627. Catalysis of an oxidation-reduction (redox) reaction in which a CH-CH group acts as a hydrogen or electron donor and reduces a hydrogen or electron acceptor.
Biocarta: CTCF Pathway	15	23	0.004	3.00E-02	CTCF: First Multivalent Nuclear Factor
Biocarta: CCR3 Pathway	20	23	0.005	3.02E-02	CCR3 signaling in Eosinophils
GO: Negative Regulation of Developmental Process	120	194	0.004	3.05E-02	GO:0051093. Any process that stops, prevents or reduces the rate or extent of development, the biological process whose specific outcome is the progression of an organism over time from an initial condition (e.g. a zygote, or a young adult) to a later condition (e.g. a multicellular animal or an aged adult).
KEGG: Gnrh Signaling Pathway	66	101	0.009	3.18E-02	hsa04912
GO: Muscle Cell Differentiation	16	22	0.003	3.27E-02	GO:0042692. The process whereby a relatively unspecialized cell acquires specialized features of a muscle cell.
GO: Regulation of DNA Metabolic Process	26	45	0.007	3.28E-02	GO:0051052. Any process that modulates the frequency, rate or extent of the chemical reactions and pathways involving DNA.
GO: Receptor Complex	35	54	0.007	3.28E-02	GO:0043235. Any protein group composed of two or more subunits, which may or may not be identical, which undergoes combination with a hormone, neurotransmitter, drug or intracellular messenger to initiate a change in cell function.
GO: Digestion	28	41	0.009	3.58E-02	GO:0007586. The whole of the physical, chemical, and biochemical processes carried out by multicellular organisms to break down ingested nutrients into components that may be easily absorbed and directed into metabolism.
GO: M Phase	65	112	0.01	3.59E-02	GO:0000279. Progression through M phase, the part of the cell cycle comprising nuclear division.
GO: Carbon Oxygen Lyase Activity	18	31	0.011	3.59E-02	GO:0016835. Catalysis of the breakage of a carbon-oxygen bond.
GO: Phosphatase Regulator Activity	17	25	0.006	3.66E-02	GO:0019208. Modulates the activity of a phosphatase, an enzyme which catalyzes of the removal of a phosphate group from a substrate molecule.
KEGG: Axon Guidance	91	127	0.006	3.79E-02	hsa04360
Biocarta: TCR Pathway	31	44	0.008	3.88E-02	T Cell Receptor Signaling Pathway
GO: Regulation of Nucleocytoplasmic Transport	15	22	0.008	3.90E-02	GO:0046822. Any process that modulates the frequency, rate or extent of the directed movement of substances between the nucleus and the cytoplasm.
GO: Intracellular Protein Transport	83	142	0.01	3.95E-02	GO:0006886. The directed movement of proteins in a cell, including the movement of proteins between specific compartments or structures within a cell, such as organelles of a eukaryotic cell.
GO: Membrane Fusion	20	28	0.009	4.00E-02	GO:0006944. The joining of two lipid bilayers to form a single membrane.
GO: Chemokine Receptor Binding	26	43	0.017	4.00E-02	GO:0042379. Interacting selectively with any chemokine receptor.
Biocarta: GPCR Pathway	25	34	0.008	4.07E-02	Signaling Pathway from G-Protein Families
GO: Vesicle	71	121	0.008	4.08E-02	GO:0031982. Any small, fluid-filled, spherical organelle enclosed by membrane or protein.
GO: Protein Tyrosine Kinase Activity	49	63	0.01	4.19E-02	GO:0004713. Catalysis of the reaction: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
GO: Protein Targeting	64	108	0.015	4.24E-02	GO:0006605. The process of targeting specific proteins to particular membrane-bound subcellular organelles. Usually requires an organelle specific protein sequence motif.
Biocarta: PDGF Pathway	23	32	0.01	4.30E-02	PDGF Signaling Pathway
Biocarta: Il2rb Pathway	23	38	0.011	4.42E-02	IL-2 Receptor Beta Chain in T cell Activation
Biocarta: IGF1r Pathway	18	23	0.011	4.50E-02	Multiple antiapoptotic pathways from IGF-1R signaling lead to BAD phosphorylation
KEGG: Taste Transduction	28	48	0.008	4.56E-02	hsa04742
GO: Gtpase Regulator Activity	78	123	0.012	4.56E-02	GO:0030695. Modulates the rate of GTP hydrolysis by a GTPase.
GO: Nuclear Organization and Biogenesis	19	29	0.01	4.65E-02	GO:0006997. A process that is carried out at the cellular level which results in the formation, arrangement of constituent parts, or disassembly of the nucleus.

GO: Nuclear Membrane	33	49	0.015	4.74E-02	GO:0031965. Either of the lipid bilayers that surround the nucleus and form the nuclear envelope; excludes the intermembrane space.
GO: Cell Junction	51	84	0.014	4.74E-02	GO:0030054. A specialized region of connection between two cells or between a cell and the extracellular matrix.
GO: Epidermis Development	41	71	0.014	4.74E-02	GO:0008544. The process whose specific outcome is the progression of the epidermis over time, from its formation to the mature structure. The epidermis is the outer epithelial layer of a plant or animal, it may be a single layer that produces an extracellular material (e.g. the cuticle of arthropods) or a complex stratified squamous epithelium, as in the case of many vertebrate species.
GO: Cell Cycle Phase	95	168	0.014	4.74E-02	GO:0022403. A cell cycle process comprising the steps by which a cell progresses through one of the biochemical and morphological phases and events that occur during successive cell replication or nuclear replication events.
GO: Chemokine Activity	25	42	0.015	4.75E-02	GO:0008009. The function of a family of chemotactic pro-inflammatory activation-inducible cytokines acting primarily upon hemopoietic cells in immunoregulatory processes; all chemokines possess a number of conserved cysteine residues involved in intramolecular disulfide bond formation.
GO: Generation of A Signal Involved In Cell Cell Signaling	19	29	0.014	4.75E-02	GO:0003001. The cellular process by which a physical entity or change in state, a signal, is created that originates in one cell and is used to transfer information to another cell. This process begins with the initial formation of the signal and ends with the mature form and placement of the signal.
GO: Regulation of Growth	34	57	0.01	4.92E-02	GO:0040008. Any process that modulates the frequency, rate or extent of the growth of all or part of an organism so that it occurs at its proper speed, either globally or in a specific part of the organism's development.
GO: Regulation of DNA Binding	29	46	0.017	4.93E-02	GO:0051101. Any process that modulates the frequency, rate or extent of DNA binding, selective interaction with deoxyribonucleic acid.
KEGG: Melanogenesis	66	100	0.013	4.93E-02	hsa04916
GO: Protein Complex Binding	37	54	0.011	4.96E-02	GO:0032403. Interacting selectively with any protein complex (a complex of two or more proteins that may include other nonprotein molecules).
KEGG: T Cell Receptor Signaling Pathway	67	107	0.014	4.96E-02	hsa04660
GO: Neuron Differentiation	57	75	0.011	4.98E-02	GO:0030182. The process whereby a relatively unspecialized cell acquires specialized features of a neuron.
GO: Regulation of Protein Modification Process	30	43	0.017	5.00E-02	GO:0031399. Any process that modulates the frequency, rate or extent of the covalent alteration of one or more amino acid residues within a protein.
GO: Protein Tyrosine Phosphatase Activity	38	53	0.013	5.00E-02	GO:0004725. Catalysis of the reaction: protein tyrosine phosphate + H2O = protein tyrosine + phosphate.
KEGG: Hedgehog Signaling Pathway		**55**	**0.015**	**5.76E-02**	**hsa04340**

Supplementary Table 3: Enriched reconstituted gene sets identified by DEPICT. DEPICT was used to assess whether the 99 genes located in associated regions (GWAS P-value < 1×10^{-5}) were enriched for any of the reconstituted gene sets defined by DEPICT. We display 36 reconstituted clusters with at least one gene set that reaches suggestive enrichment (P-value < 0.01). For each cluster we provide all gene sets that contribute to this cluster and were nominally enriched (P-value<0.05). Associated regions were constructed by mapping genes to a given region if the genes resided within, or overlapped with linkage disequilibrium (r2 > 0.5) flanks of a given lead SNP. Ex: Cluster ID and cluster center was defined by Affinity Propagation method which was used to cluster gene sets that were highly correlated and automatically defines independent clusters based on calculated Pearson's correlation matrix (See methods for dertails). The central gene set of a given cluster is the most significant gene set in that cluster (bold). We provide the list of genes near associated loci allocated to each gene set and their respective likelihood to be part of each gene set. We only provide genes with significant likelihood |z-score| > 1.96 corresponding to P-value <0.05. NA: not available, is indicated for gene sets for which no gene presented a significant z-score.

Cluster ID	Gene set ID	Gene set Name	Central gene set of the cluster	Most significant gene set	P value	Genes at associated loci contributing to gene sets
1	MP:0010300	Increased skin tumor incidence	Intestinal adenocarcinoma	Increased skin tumor incidence	0.0001	XPC (3.7), SRR (3.1), SLC7A11 (3.0), BPNT1 (2.9), TMEM40 (2.3), AC068353.34 (2.3), TACR3 (-2.1)
1	MP:0002957	Intestinal adenocarcinoma			0.001	FOXL1 (2.5), SLFN11 (2.5), AC068353.34 (2.2), UFL1 (-2.6)
1	ENSG00000138180	CEP55 PPI subnetwork			0.02	AL450226.1 (2.2)
1	MP:0002404	Increased intestinal adenoma incidence			0.02	SRR (3.2), CAND2 (2.8), XPC (2.7), RASSF3 (2.2), AC068353.34 (2.1), FOXL1 (2.0), LINC00973 (2.0), AL450226.1 (2.0), AKR1C3 (-2.4)
2	GO:0030684	Preribosome	FLG2 PPI subnetwork	Preribosome	0.0004	TSR1 (3.6), SETD4 (3.2), ERCC4 (2.2), SRR (-2.3)
2	ENSG00000065268	WDR18 PPI subnetwork			0.004	SMG6 (3.1), TSR1 (2.9), SLFN11 (2.4), RUNX1 (-2.3)
2	ENSG00000143520	FLG2 PPI subnetwork			0.03	TSR1 (3.0), SMG6 (3.0)
2	ENSG00000136021	SCYL2 PPI subnetwork			0.04	LINC02064 (2.6), RP1L1 (2.2), PEAK1 (-2.4)
3	MP:0001689	Incomplete somite formation	Vasculature development	Incomplete somite formation	0.0004	PTPRM (3.3), LINC00312 (-2.6)
3	GO:0045766	Positive regulation of angiogenesis			0.0006	LMCD1 (2.3), LINC00312 (2.2)
3	MP:0001614	Abnormal blood vessel morphology			0.0009	LMCD1 (2.5), PEAK1 (2.5), PTPRM (2.4), FHL5 (2.0)
3	MP:0004787	Abnormal dorsal aorta morphology			0.007	PTPRM (2.9), RASSF3 (2.8), LMCD1 (2.0), SEZ6L (-2.3)
3	ENSG00000087303	NID2 PPI subnetwork			0.009	PEAK1 (2.7), PTPRM (2.3), LMCD1 (2.1)
3	MP:0006126	Abnormal outflow tract development			0.01	LINC01013 (2.0), ERCC4 (-2.1)
3	GO:0001570	Vasculogenesis			0.01	FOXL1 (4.0), PTPRM (2.9), RASSF3 (2.9), UBTD1 (2.4), FHL5 (2.4), LINC01648 (2.0), AC004944.1 (2.0)
3	MP:0008803	Abnormal placental labyrinth vasculature			0.01	RASSF3 (2.2), FOXL1 (2.1), PTPRM (2.0), LINC00299 (2.0)
3	MP:0004251	Failure of heart looping			0.02	SMG6 (2.8), PBX1 (2.7), RASSF3 (2.5), CAND2 (2.3), PIEZO1 (2.3), PEAK1 (2.1), RAB3GAP2 (2.1), FHL5 (-2.0), LINC00312 (-2.7)
3	ENSG00000142798	HSPG2 PPI subnetwork			0.02	PTPRM (2.8), PEAK1 (2.6), LMCD1 (2.5), MN1 (2.5), TACR3 (2.2)
3	MP:0010392	Prolonged qrs complex duration			0.02	PEAK1 (2.6)
3	MP:0000272	Abnormal aorta morphology			0.02	LMCD1 (3.9), LINC01013 (2.6), PEAK1 (2.5), LINC00312 (2.5), LINC00299 (2.2), SETD4 (2.1), HMG20A (2.0), UBTD1 (-2.1), TMEM40
3	ENSG00000109072	SEBOX PPI subnetwork			0.02	LMCD1 (2.6), FOXL1 (2.3), LINC00973 (2.2), TACR3 (2.0)
3	MP:0005592	Abnormal vascular smooth muscle morphology			0.02	LMCD1 (4.1), RASSF3 (3.9), FHL5 (3.4), PTPRM (3.2), UBTD1 (2.0), FOXL1 (2.0)
3	MP:0010856	Dilated respiratory conducting tubes			0.03	RUNX1 (2.3), FOXL1 (2.3), PEAK1 (2.0)
3	GO:0045785	Positive regulation of cell adhesion			0.03	AC004944.1 (3.1), PEAK1 (3.0), LMCD1 (2.6), LINC01648 (2.5), AL450226.1 (2.3), LINC00973 (2.0), CAND2 (-2.0)
3	ENSG00000128052	KDR PPI subnetwork			0.03	PTPRM (3.1), RAB3GAP2 (2.9), FHL5 (2.6), UBTD1 (2.3), LINC00973 (2.1)
3	ENSG00000182492	BGN PPI subnetwork			0.03	LMCD1 (2.8), AL450226.1 (2.4)
3	GO:0048514	Blood vessel morphogenesis			0.03	LMCD1 (3.6), UBTD1 (3.2), PTPRM (3.0), LINC00312 (2.8), PEAK1 (2.8), FHL5 (2.7)
3	MP:0003071	Decreased vascular permeability			0.04	PTPRM (3.6), FHL5 (3.4), LINC01013 (2.8), SLFN11 (2.2)
3	GO:0001525	Angiogenesis			0.04	LMCD1 (3.5), UBTD1 (3.2), PEAK1 (2.8), FHL5 (2.7), LINC00312 (2.6), PTPRM (2.4), AC004944.1 (-2.1)
3	KEGG DORSO VENTRAL AXIS FORMATION	Kegg dorso ventral axis formation			0.04	PEAK1 (3.3), FOXL1 (3.1), PTPRM (2.4), MN1 (2.3), RASSF3 (2.1), FHL5 (2.0)
3	GO:0044319	Wound healing, spreading of cells			0.04	LINC02064 (2.4), LINC01013 (2.3), UBTD1 (2.1)
3	GO:0001944	Vasculature development			0.04	LMCD1 (3.6), PEAK1 (3.2), UBTD1 (2.9), PTPRM (2.8), LINC00312 (2.6), FHL5 (2.3), FOXL1 (2.0)
3	GO:0001945	Lymph vessel development			0.04	LINC01921 (3.1), FHL5 (3.0), UBTD1 (2.5), FOXL1 (2.5), AC004944.1 (2.2)
3	GO:0032403	Protein complex binding			0.04	RUNX1 (2.9), PTPRM (2.8)
3	GO:0001568	Blood vessel development			0.04	LMCD1 (3.5), PEAK1 (3.2), PTPRM (2.9), UBTD1 (2.9), LINC00312 (2.8), FHL5 (2.2), FOXL1 (2.0)
3	MP:0004921	Decreased placenta weight			0.04	PIEZO1 (3.0), TMEM40 (2.4), TRIB1 (2.1)
4	GO:0048661	Positive regulation of smooth muscle cell	Regulation of smooth muscle cell proliferation	Positive regulation of smooth muscle cell proliferation	0.0006	LINC00312 (3.7), FHL5 (3.6), LINC01648 (2.6), LMCD1 (2.2), PTPRM (2.1), ANKRD20A11P (2.1), TRIB1 (2.0)
4	GO:0033002	Muscle cell proliferation			0.004	FHL5 (3.1), LMCD1 (2.8), LINC00312 (2.7), FOXL1 (2.6), PTPRM (2.3), AC068353.34 (2.0), ANKRD20A11P (2.0), CBR1 (-2.2)
4	GO:0048660	Regulation of smooth muscle cell proliferation			0.01	FHL5 (3.8), LINC00312 (3.5), PTPRM (2.4), LMCD1 (2.3), ANKRD20A11P (2.2)
4	GO:0048659	Smooth muscle cell proliferation			0.01	FHL5 (4.0), LINC00312 (3.1), ANKRD20A11P (2.3), PTPRM (2.3), LMCD1 (2.0)

4	GO:0045787	Positive regulation of cell cycle			0.02	HDGFL1 (2.6), LINC01921 (2.4), SLC7A11 (2.2)
4	ENSG00000071539	TRIP13 PPI subnetwork			0.02	AKR1C3 (3.3), LINC00312 (2.0), SIPA1L1 (-2.0)
5	MP:0002747	Abnormal aortic valve morphology	Abnormal semilunar valve morphology	Abnormal aortic valve morphology	0.0006	LINC00299 (2.2), SLFN11 (2.1), TRIB1 (2.1), LINC00312 (2.1), AC068353.34 (2.0), RUNX1 (2.0), CBR1 (-2.1)
5	MP:0002746	Abnormal semilunar valve morphology			0.007	LMCD1 (3.5), PTPRM (3.0), RUNX1 (2.6), LINC00312 (2.3), GLIS1 (2.0)
5	MP:0005269	Abnormal occipital bone morphology			0.02	GLIS1 (2.9), SIPA1L1 (2.8), FOXL1 (2.0), RAB3GAP2 (-2.2)
5	MP:0002754	Dilated heart right ventricle			0.02	LMCD1 (5.9), LINC00312 (4.1), FHL5 (2.5), PTPRM (2.4), PEAK1 (2.0)
6	MP:0008840	Abnormal spike wave discharge	Voltage-gated calcium channel activity	Abnormal spike wave discharge	0.0006	ANKRD20A11P (2.6), SEZ6L (2.3)
6	MP:0003990	Decreased neurotransmitter release			0.005	NRXN3 (3.0), ABLIM2 (2.7), SEZ6L (2.6), PTPRM (2.1), MN1 (2.1), FHL5 (2.0)
6	MP:0003063	Increased coping response			0.02	ANKRD20A11P (3.1), GLIS1 (2.8), CAND2 (2.5), NRXN3 (2.5), SIPA1L1 (2.1), MN1 (2.1), LINC00312 (-2.0)
6	GO:0005245	Voltage-gated calcium channel activity			0.03	ABLIM2 (3.2), SEZ6L (2.8), MN1 (2.3), ANKRD20A11P (2.2), NRXN3 (2.2), XPC (-2.0)
6	GO:0007628	Adult walking behavior			0.04	NRXN3 (2.7), TACR3 (2.4), LINC01648 (2.2), AC004944.1 (2.0)
6	GO:0051924	Regulation of calcium ion transport			0.04	LINC00299 (2.3), ANKRD20A11P (2.2), ABLIM2 (2.1)
7	MP:0009409	Abnormal skeletal muscle fiber type ratio	Abnormal skeletal muscle fiber morphology	Abnormal skeletal muscle fiber type ratio	0.0006	CAND2 (3.5), ABLIM2 (2.6), LINC01648 (2.3)
7	MP:0000005	Increased brown adipose tissue amount			0.01	FOXL1 (3.1), FHL5 (2.7)
7	MP:0009404	Centrally nucleated skeletal muscle fibers			0.01	CAND2 (4.1), LMCD1 (2.5), UBTD1 (2.0)
7	MP:0002841	Impaired skeletal muscle contractility			0.01	CAND2 (4.8), ABLIM2 (4.4), LMCD1 (2.1), RP1L1 (2.0)
7	MP:0004091	Abnormal z lines			0.01	CAND2 (3.2), UBTD1 (2.2), ABLIM2 (2.1), LMCD1 (2.0)
7	MP:0000747	Muscle weakness			0.01	ABLIM2 (2.7), CAND2 (2.4)
7	MP:0004087	Abnormal muscle fiber morphology			0.02	RUNX1 (3.7), MN1 (3.4), LMCD1 (3.2), LINC00312 (2.7), CAND2 (2.4), AKR1C3 (2.0), SRR (-2.0), SETD4 (-2.2)
7	MP:0002279	Abnormal diaphragm morphology			0.02	LMCD1 (4.5), CAND2 (3.2), MN1 (2.4), RUNX1 (2.3), PIEZO1 (2.2), UBTD1 (2.1)
7	MP:0000748	Progressive muscle weakness			0.02	LMCD1 (2.8), PEAK1 (2.3), IARS2 (2.2), UBTD1 (2.0)
7	MP:0003084	Abnormal skeletal muscle fiber morphology			0.03	CAND2 (3.7), LMCD1 (3.5), UBTD1 (2.9), PIEZO1 (2.5), ABLIM2 (2.4), LINC00312 (2.0)
7	MP:0009400	Decreased skeletal muscle fiber size			0.04	TRIB1 (3.0), LMCD1 (2.8), ABLIM2 (2.6), CAND2 (2.6), LINC00312 (2.0)
7	MP:0002332	Abnormal exercise endurance			0.04	LMCD1 (3.8), LINC00312 (3.5), TRIB1 (2.8), ABLIM2 (2.2), CAND2 (2.1), CBR1 (2.0), SRR (-2.2)
8	ENSG00000093167	LRRFIP2 PPI subnetwork	GSTM1 PPI subnetwork	LRRFIP2 PPI subnetwork	0.0009	SMG6 (3.5), AC068353.34 (2.7), LINC00973 (2.4), TMEM40 (2.3)
8	ENSG00000166033	HTRA1 PPI subnetwork			0.007	LMCD1 (4.5), AC068353.34 (3.1), PEAK1 (2.2), GLIS1 (2.0), FOXL1 (2.0)
8	ENSG00000171992	SYNPO PPI subnetwork			0.01	AC068353.34 (2.8), SMG6 (2.6), LMCD1 (2.1), ANKRD20A11P (2.1)
8	ENSG00000197321	SVIL PPI subnetwork			0.01	AC068353.34 (3.4), AKR1C3 (2.2), ANKRD20A11P (2.1), TMEM40 (2.1)
8	ENSG00000106028	SSBP1 PPI subnetwork			0.01	SMG6 (3.5), LMCD1 (3.1), AC068353.34 (2.4), PEAK1 (2.3)
8	ENSG00000167657	DAPK3 PPI subnetwork			0.02	LMCD1 (3.0)
8	ENSG00000196586	MYO6 PPI subnetwork			0.02	PEAK1 (2.7), SMG6 (2.6), LMCD1 (2.4), SEZ6L (2.3), TMEM40 (2.1)
8	ENSG00000134184	GSTM1 PPI subnetwork			0.02	AC068353.34 (2.9), ANKRD20A11P (2.5), SMG6 (2.3), LMCD1 (2.2), AKR1C3 (2.2)
8	ENSG00000071909	MYO3B PPI subnetwork			0.02	AC068353.34 (2.8), LMCD1 (2.4), ANKRD20A11P (2.2), AKR1C3 (2.0)
8	ENSG00000105357	MYH14 PPI subnetwork			0.02	LMCD1 (3.0), ANKRD20A11P (2.7), AC068353.34 (2.6), SMG6 (2.0)
8	ENSG00000198467	TPM2 PPI subnetwork			0.02	LMCD1 (4.4), AC068353.34 (2.5), SMG6 (2.2), LINC01013 (-2.2)
8	ENSG00000134202	GSTM3 PPI subnetwork			0.02	SMG6 (3.1), AC068353.34 (3.0), FOXL1 (2.6), LMCD1 (2.6), RASSF3 (2.6), PEAK1 (2.4), ABLIM2 (2.1), LINC01013 (-2.5)
8	ENSG00000100014	SPECC1L PPI subnetwork			0.02	ANKRD20A11P (2.7), AKR1C3 (2.6), AC068353.34 (2.1), SMG6 (2.1)
8	ENSG00000110880	CORO1C PPI subnetwork			0.03	SMG6 (2.4), AC068353.34 (2.1), LMCD1 (2.0), AKR1C3 (2.0), ANKRD20A11P (2.0)
8	ENSG00000101331	C20ORF160 PPI subnetwork			0.03	LMCD1 (2.3), AC068353.34 (2.2), SMG6 (2.2), RASSF3 (2.0)
8	ENSG00000120265	PCMT1 PPI subnetwork			0.03	GLIS1 (2.5), PEAK1 (2.0)
8	ENSG00000006747	SCIN PPI subnetwork			0.04	SLC7A11 (2.3), LMCD1 (2.3), SMG6 (2.3), AC068353.34 (2.1)
8	ENSG00000128487	SPECC1 PPI subnetwork			0.04	LMCD1 (2.5), RASSF3 (2.2), SMG6 (2.1), AC068353.34 (2.1), PTPRM (2.0)
8	ENSG00000204361	FAM55B PPI subnetwork			0.04	AC068353.34 (2.1), LMCD1 (2.1), LINC01648 (-2.0)
9	GO:0034405	Response to fluid shear stress			0.001	PIEZO1 (3.1), ERCC4 (2.6), AL450226.1 (2.5), LINC02064 (2.5), LINC01921 (2.1), LMCD1 (2.1), RASSF3 (2.0), LINC01648 (2.0)
9	ENSG00000139567	ACVRL1 PPI subnetwork			0.003	RASSF3 (2.1), RUNX1 (2.0)
9	ENSG00000198742	SMURF1 PPI subnetwork			0.008	HMG20A (3.7), PBX1 (2.7), LINC00312 (2.2)
9	GO:0060389	Pathway-restricted smad protein phosphorylation			0.009	LINC02064 (3.2), LMCD1 (2.3), GLIS1 (2.0), SRR (-2.0), TMEM40 (-2.2)

9	GO:0005160	TGF-beta receptor binding	TGF-beta receptor binding	Response to fluid shear stress	0.02	LMCD1 (2.6), PTPRM (2.0), RASSF3 (2.1), AKR1C3 (2.1), LINC00312 (2.0)
9	ENSG00000120693	SMAD9 PPI subnetwork			0.03	UFL1 (2.3), TRIB1 (2.1)
9	ENSG00000163513	TGFBR2 PPI subnetwork			0.03	AC068353.34 (2.7), LMCD1 (2.5), RUNX1 (2.0), SLFN11 (-2.0)
9	ENSG00000123612	ACVR1C PPI subnetwork			0.03	LINC00973 (3.5), RUNX1 (3.2), LMCD1 (3.0), BPNT1 (-2.1)
9	GO:0017015	Regulation of transforming growth factor beta			0.04	RASSF3 (2.5), LINC00299 (-2.0)
10	MP:0000223	Decreased monocyte cell number			0.001	BPNT1 (2.0)
10	MP:0002465	Abnormal eosinophil physiology			0.002	LINC01013 (3.3)
10	GO:0005544	Calcium-dependent phospholipid binding			0.005	PIEZO1 (2.5), AKR1C3 (2.3)
10	REACTOME PROSTANOID	Reactome prostanoid metabolism			0.01	AKR1C3 (3.1), SLFN11 (2.4), RUNX1 (2.2), PBX1 (2.2), LINC01013 (2.1)
10	ENSG00000196415	PRTN3 PPI subnetwork			0.02	NA
10	ENSG00000101336	HCK PPI subnetwork	Impaired neutrophil chemotaxis	Decreased monocyte cell number	0.02	RUNX1 (2.6), RAB3GAP2 (2.3), UBTD1 (2.1), HMG20A (2.0)
10	MP:0000219	Increased neutrophil cell number			0.03	SLFN11 (2.2), PEAK1 (2.0)
10	ENSG00000138032	PPM1B PPI subnetwork			0.03	PIEZO1 (2.6), RP1L1 (2.3)
10	MP:0008720	Impaired neutrophil chemotaxis			0.04	SLFN11 (2.1)
10	MP:0008657	Increased interleukin-1 beta secretion			0.04	LINC00299 (2.2), TRIB1 (2.0)
10	MP:0000511	Abnormal intestinal mucosa morphology			0.04	LINC01013 (2.6)
11	MP:0000761	Thin diaphragm muscle			0.001	LINC00973 (2.5), ABLIM2 (2.1)
11	MP:0004835	Abnormal miniature endplate potential			0.01	ABLIM2 (4.7), UBTD1 (3.4), LINC00973 (2.4), TACR3 (2.4), PEAK1 (2.3), RUNX1 (2.2), CAND2 (2.0)
11	MP:0002826	Tonic seizures			0.01	NRXN3 (3.2), LINC00312 (2.2), SRR (-2.0)
11	ENSG00000125968	ID1 PPI subnetwork			0.02	ABLIM2 (2.2)
11	MP:0001053	Abnormal neuromuscular synapse morphology	Abnormal neuromuscular synapse morphology	Thin diaphragm muscle	0.02	ABLIM2 (5.2), UBTD1 (3.6), PTPRM (3.0), TACR3 (2.6), NRXN3 (2.5), RUNX1 (2.3)
11	ENSG00000030304	MUSK PPI subnetwork			0.02	UBTD1 (3.6), RUNX1 (2.9), ABLIM2 (2.7), LMCD1 (2.2), TACR3 (2.1)
11	MP:0000753	Paralysis			0.03	RAB3GAP2 (3.8), CBR1 (3.6), NRXN3 (2.7), SLC7A11 (2.4), CAND2 (2.2), SLFN11 (2.1), RASSF3 (-2.7)
11	ENSG00000174469	CNTNAP2 PPI subnetwork			0.03	NRXN3 (2.1)
11	GO:0031594	Neuromuscular junction			0.04	ABLIM2 (3.6), CAND2 (3.0), RUNX1 (2.4)
11	MP:0002914	Abnormal endplate potential			0.04	ABLIM2 (3.9), LINC01013 (3.4), RUNX1 (2.9), UBTD1 (2.9), LINC00973 (2.6), CAND2 (2.2)
12	ENSG00000165702	GFI1B PPI subnetwork			0.001	SMG6 (2.2)
12	GO:0051015	Actin filament binding			0.002	AC068353.34 (2.3), ANKRD20A11P (2.3), PIEZO1 (2.0)
12	GO:0032432	Actin filament bundle			0.009	LMCD1 (4.0), CAND2 (3.2), LINC00312 (2.9), PIEZO1 (2.6), AC068353.34 (2.5), RASSF3 (2.1), SIPA1L1 (2.1)
12	GO:0001725	Stress fiber			0.009	LMCD1 (3.7), CAND2 (3.2), LINC00312 (2.9), AC068353.34 (2.3), RASSF3 (2.3), PIEZO1 (2.1), FHL5 (2.0)
12	ENSG00000163466	ARPC2 PPI subnetwork			0.01	AC068353.34 (2.7), SMG6 (2.5), PEAK1 (2.1), SRR (2.1)
12	GO:0042641	Actomyosin			0.01	LMCD1 (3.6), CAND2 (3.2), LINC00312 (2.6), RASSF3 (2.5), SIPA1L1 (2.1), PIEZO1 (2.1), SMG6 (2.1), AC068353.34 (2.0)
12	ENSG00000135269	TES PPI subnetwork			0.03	RASSF3 (2.9), SIPA1L1 (2.1), CAND2 (2.0), LMCD1 (2.0), PBX1 (2.0), LINC00973 (-2.5)
12	ENSG00000145349	CAMK2D PPI subnetwork	Actomyosin	GFI1B PPI subnetwork	0.03	RAB3GAP2 (2.6), CAND2 (2.4), SIPA1L1 (2.2), RASSF3 (2.0)
12	REACTOME SMOOTH MUSCLE CONTRACTION	Reactome smooth muscle contraction			0.03	FHL5 (4.7), RASSF3 (3.8), LMCD1 (3.6), NRXN3 (2.5), LINC00312 (2.2), LINC01648 (-2.1)
12	ENSG00000150867	PIP4K2A PPI subnetwork			0.03	SIPA1L1 (2.7), AC068353.34 (2.7), PEAK1 (2.3), ABLIM2 (2.2), RAB3GAP2 (2.1), ERCC4 (2.1), CAND2 (2.0), SEZ6L (-2.1)
12	ENSG00000154556	SORBS2 PPI subnetwork			0.03	PEAK1 (2.2)
12	ENSG00000144061	NPHP1 PPI subnetwork			0.04	MN1 (3.0), PIEZO1 (2.8), ERCC4 (-2.0)
12	GO:0005925	Focal adhesion			0.04	AC068353.34 (2.6), LMCD1 (2.6), PIEZO1 (2.6), RASSF3 (2.5), XPC (2.4), LINC00312 (2.0), LINC01648 (-3.0)
12	ENSG00000106976	DNM1 PPI subnetwork			0.04	RAB3GAP2 (2.0)
12	GO:0015629	Actin cytoskeleton			0.04	LMCD1 (3.5), CAND2 (3.3), RASSF3 (2.1)
13	MP:0006354	Abnormal fourth branchial arch artery morphology			0.002	FOXL1 (5.3), FHL5 (3.2), PEAK1 (2.8), PIEZO1 (2.1), PTPRM (2.0)
13	MP:0010465	Aberrant origin of the right subclavian artery			0.004	PIEZO1 (2.7), PBX1 (2.3), TACR3 (2.2), CAND2 (2.2), SIPA1L1 (-2.2)
13	MP:0004157	Interrupted aortic arch			0.004	FOXL1 (3.0), ERCC4 (2.6), PIEZO1 (2.3), TACR3 (2.2), RUNX1 (2.2)
13	GO:0031490	Chromatin DNA binding	Abnormal fourth branchial arch artery morphology	Abnormal fourth branchial arch artery morphology	0.01	FOXL1 (5.4), MN1 (3.2), LINC00973 (2.7), RUNX1 (2.5), FHL5 (2.3), UBTD1 (2.0), SEZ6L (2.0), IARS2 (-2.0)
13	GO:0048844	Artery morphogenesis			0.02	FOXL1 (4.4), GLIS1 (2.3), LMCD1 (2.0)
13	GO:0001569	Patterning of blood vessels			0.04	LINC01648 (3.2), FHL5 (2.6), FOXL1 (2.3), PEAK1 (2.3)

13	MP:0004136	Right aortic arch			0.04	TRIB1 (2.8), DAB1 (2.4), CREC4 (2.1), RUNX1 (2.0)
14	GO:0043073	Germ cell nucleus	Germ cell nucleus	Germ cell nucleus	0.00238	LINC01921 (4.8), FHL5 (2.9), LINC01013 (2.3), SRR (2.3)
14	GO:0001673	Male germ cell nucleus			0.00313	LINC01921 (4.9), FHL5 (3.3), SRR (2.6), LINC02064 (2.1), LINC01013 (2.1)
15	MP:0009230	Abnormal sperm head morphology	Teratozoospermia	Abnormal sperm head morphology	0.00244	LINC01921 (3.4), FHL5 (3.3), LINC00312 (2.6), HDGFL1 (2.4)
15	MP:0009238	Coiled sperm flagellum			0.01	LINC01921 (6.0), FHL5 (2.3), AL450226.1 (2.0)
15	MP:0004542	Impaired acrosome reaction			0.01	LINC01921 (3.8), FHL5 (3.7), UBTD1 (3.4), AC004944.1 (2.7)
15	MP:0005578	Teratozoospermia			0.02	LINC01921 (5.7), FHL5 (3.7), HDGFL1 (3.6), UBTD1 (2.0)
15	MP:0000242	Impaired fertilization			0.03	LINC01921 (4.5), HDGFL1 (2.7), FHL5 (2.6), AC004944.1 (2.4), XPC (-2.7)
15	MP:0009232	Abnormal sperm nucleus morphology			0.03	LINC01921 (5.6), FHL5 (2.7), HDGFL1 (2.0), AL450226.1 (2.0)
15	GO:0007281	Germ cell development			0.04	HDGFL1 (5.3), LINC01921 (5.2), FHL5 (2.7)
15	MP:0010769	Abnormal survival			0.04	TACR3 (2.4), RASSF3 (2.4)
16	ENSG00000125503	PPP1R12C PPI subnetwork	MYO1G PPI subnetwork	PPP1R12C PPI subnetwork	0.003	PIEZO1 (2.2), AC068353.34 (2.2), ANKRD20A11P (2.1)
16	ENSG00000106992	AK1 PPI subnetwork			0.004	RASSF3 (2.1)
16	ENSG00000136286	MYO1G PPI subnetwork			0.005	RASSF3 (2.2), ANKRD20A11P (2.0)
16	ENSG00000013297	CLDN11 PPI subnetwork			0.005	RASSF3 (2.4), ANKRD20A11P (2.4), PIEZO1 (2.0)
16	ENSG00000166866	MYO1A PPI subnetwork			0.01	AC068353.34 (3.8), RASSF3 (2.7), LMCD1 (2.4), PEAK1 (2.3), LINC01648 (-2.2)
16	ENSG00000118260	CREB1 PPI subnetwork			0.02	UBTD1 (3.2), SMG6 (2.2), LINC01921 (2.1)
16	ENSG00000169220	RGS14 PPI subnetwork			0.02	RASSF3 (2.7)
16	ENSG00000136156	ITM2B PPI subnetwork			0.02	RASSF3 (2.4), AC068353.34 (2.2), ANKRD20A11P (2.0)
16	ENSG00000136068	FLNB PPI subnetwork			0.02	SRR (3.2), AC068353.34 (2.6), SMG6 (2.4), PIEZO1 (2.1), RASSF3 (2.0), UBTD1 (2.0)
16	ENSG00000105376	ICAM5 PPI subnetwork			0.03	RASSF3 (2.1), ANKRD20A11P (2.0)
16	ENSG00000139514	SLC7A1 PPI subnetwork			0.03	ANKRD20A11P (2.3), RASSF3 (2.2), GLIS1 (-2.0)
17	GO:0043500	Muscle adaptation	Muscle system process	Muscle adaptation	0.003	LINC00312 (3.9), RASSF3 (2.5), CAND2 (2.2), AL450226.1 (-2.5)
17	GO:0090257	Regulation of muscle system process			0.01	CAND2 (3.6), TACR3 (2.1)
17	GO:0031143	Pseudopodium			0.01	CAND2 (4.2), LMCD1 (3.7), ABLIM2 (2.1)
17	GO:0003012	Muscle system process			0.02	CAND2 (4.7), ABLIM2 (2.7), RASSF3 (2.5), LMCD1 (2.1), ANKRD20A11P (2.0)
17	GO:0006937	Regulation of muscle contraction			0.03	CAND2 (3.7)
17	ENSG00000109846	CRYAB PPI subnetwork			0.03	CAND2 (2.8), SLC7A11 (2.5), FHL5 (-2.6)
17	ENSG00000135424	ITGA7 PPI subnetwork			0.03	FHL5 (2.9), ABLIM2 (2.5), TACR3 (2.4), UBTD1 (2.4)
17	GO:0055117	Regulation of cardiac muscle contraction			0.04	LINC01013 (3.9), RASSF3 (2.1), AC004944.1 (2.0), LINC00312 (-2.4)
17	GO:0048747	Muscle fiber development			0.04	AL450226.1 (3.2), CAND2 (3.2), ABLIM2 (3.0), RASSF3 (2.2), TACR3 (2.0), LINC01921 (2.0), SLFN11 (-2.4)
17	ENSG00000177791	MYOZ1 PPI subnetwork			0.04	CAND2 (3.8), LMCD1 (2.2)
17	GO:0016460	Myosin ii complex			0.04	CAND2 (4.2), LMCD1 (3.0)
17	MP:0005140	Decreased cardiac muscle contractility			0.04	LMCD1 (3.4), CAND2 (3.2)
17	GO:0000146	Microfilament motor activity			0.04	CAND2 (2.8), LMCD1 (2.5)
17	MP:0005620	Abnormal muscle contractility			0.04	FHL5 (3.0), CAND2 (3.0), NRXN3 (2.5), ABLIM2 (2.2), LINC00312 (2.1), LMCD1 (2.0)
17	GO:0006936	Muscle contraction			0.04	CAND2 (4.8), ABLIM2 (2.7), RASSF3 (2.5), LMCD1 (2.0)
17	GO:0005516	Calmodulin binding			0.04	SIPA1L1 (3.1), ABLIM2 (3.1), LMCD1 (2.2), CAND2 (2.1)
18	ENSG00000122641	INHBA PPI subnetwork	INHBA PPI subnetwork	INHBA PPI subnetwork	0.004	RUNX1 (4.5), LMCD1 (4.0)
18	ENSG00000106991	ENG PPI subnetwork			0.02	LMCD1 (4.2), RUNX1 (2.6), LINC01013 (2.1), SRR (-2.2)
18	MP:0003645	Increased pancreatic beta cell number			0.03	AC068353.34 (2.5), LINC00312 (2.0)
18	ENSG00000175189	INHBC PPI subnetwork			0.03	RUNX1 (5.3), LMCD1 (3.7), LINC00973 (2.6), MN1 (2.2)
18	GO:0001558	Regulation of cell growth			0.04	LMCD1 (3.0), ANKRD20A11P (2.7), AC004944.1 (2.4), GLIS1 (2.2)
18	ENSG00000163083	INHBB PPI subnetwork			0.04	RUNX1 (4.6), LMCD1 (2.8), MN1 (2.6)
19	GO:0034504	Protein localization to nucleus	Protein localization to nucleus	Protein localization to nucleus	0.004	MN1 (2.9), XPC (2.0), PBX1 (2.0)
19	GO:0006606	Protein import into nucleus			0.03	MN1 (2.8), PBX1 (2.1), ANKRD20A11P (-2.0)
19	GO:0051170	Nuclear import			0.04	MN1 (2.8), PBX1 (2.0), RUNX1 (2.0)

19	GO:0033365	Protein localization to organelle			0.04	XPC (2.0)
20	ENSG00000158290	CUL4B PPI subnetwork			0.004	HMG20A (4.5), RAB3GAP2 (3.5), IARS2 (2.7), XPC (2.6), UFL1 (2.5), AC068353.34 (2.3), CDC5L (2.3), LINC00299 (2.2)
20	ENSG00000003096	KLHL13 PPI subnetwork			0.008	CDC5L (3.3), RAB3GAP2 (2.9), SIPA1L1 (2.5), IARS2 (2.4)
20	ENSG00000108094	CUL2 PPI subnetwork	VPRBP PPI subnetwork	CUL4B PPI subnetwork	0.01	HMG20A (2.6)
20	ENSG00000139842	CUL4A PPI subnetwork			0.01	HMG20A (4.0), IARS2 (2.8), XPC (2.5), LINC00299 (2.0), RAB3GAP2 (2.0), SLC7A11 (-2.5)
20	ENSG00000145041	VPRBP PPI subnetwork			0.01	CDC5L (3.2), RAB3GAP2 (3.1), SIPA1L1 (2.5), IARS2 (2.3), CAND2 (2.1)
20	ENSG00000120696	KBTBD7 PPI subnetwork			0.03	CDC5L (3.1), RAB3GAP2 (2.9), UFL1 (2.3), CAND2 (2.0)
21	ENSG00000112559	MDFI PPI subnetwork			0.004	LINC01921 (2.2), LINC00973 (2.1), UBTD1 (2.1), SETD4 (2.0)
21	ENSG00000127191	TRAF2 PPI subnetwork			0.004	RAB3GAP2 (2.7)
21	ENSG00000064300	NGFR PPI subnetwork	TRAF2 PPI subnetwork	MDFI PPI subnetwork	0.02	RUNX1 (2.0)
21	ENSG00000183763	TRAIP PPI subnetwork			0.03	SETD4 (2.3), TRIB1 (-2.1)
21	ENSG00000056972	TRAF3IP2 PPI subnetwork			0.04	SLC7A11 (2.5)
22	ENSG00000164091	WDR82 PPI subnetwork			0.004	HMG20A (5.0), XPC (3.4), LINC00299 (2.8), RAB3GAP2 (2.1)
22	ENSG00000136997	MYC PPI subnetwork			0.008	HMG20A (2.7), RAB3GAP2 (2.5)
22	ENSG00000139496	NUPL1 PPI subnetwork			0.0095	CDC5L (2.9), RAB3GAP2 (2.6), IARS2 (2.5), TSR1 (2.0)
22	REACTOME PURINE RIBONUCLEOSIDE	Reactome purine ribonucleoside			0.01	IARS2 (4.5), TSR1 (2.3), XPC (2.2), LINC01921 (2.0), ANKRD20A11P (-2.8)
22	REACTOME GLUCOSE TRANSPORT	Reactome glucose transport			0.02	RAB3GAP2 (2.2)
22	ENSG00000093000	NUP50 PPI subnetwork			0.03	RAB3GAP2 (2.5), CDC5L (2.4)
22	REACTOME ISG15 ANTIVIRAL MECHANISM	Reactome isg15 antiviral mechanism			0.03	NA
22	REACTOME ANTIVIRAL MECHANISM BY	Reactome antiviral mechanism by	Reactome transport of mature mrna derived from an intronless transcript	WDR82 PPI subnetwork	0.03	NA
22	ENSG00000124789	NUP153 PPI subnetwork			0.03	CDC5L (3.4), IARS2 (3.0), TSR1 (2.1), RAB3GAP2 (2.0)
22	REACTOME HEXOSE TRANSPORT	Reactome hexose transport			0.03	HDGFL1 (2.3), RAB3GAP2 (2.3)
22	ENSG00000132341	RAN PPI subnetwork			0.04	CDC5L (3.2), TSR1 (2.5)
22	REACTOME TRANSPORT OF MATURE MRNA	Reactome transport of mature mRNA derived			0.04	RAB3GAP2 (2.3),CDC5L (2.3),IARS2 (2.3),UBTD1 (-2.3)
22	ENSG00000126883	NUP214 PPI subnetwork			0.04	CDC5L (2.5),IARS2 (2.2),RAB3GAP2 (2.0)
22	ENSG00000124571	XPO5 PPI subnetwork			0.04	CDC5L (2.6),RAB3GAP2 (2.0),ANKRD20A11P (-2.1)
22	ENSG00000101146	RAE1 PPI subnetwork			0.04	CDC5L (3.7),TSR1 (3.4),RAB3GAP2 (2.4)
22	REACTOME TRANSPORT OF MATURE MRNAS	Reactome transport of mature mRNAs derived			0.04	CDC5L (2.4),IARS2 (2.2),RAB3GAP2 (2.1),UBTD1 (-2.2)
23	ENSG00000165156	ZHX1 PPI subnetwork			0.004	AKR1C3 (2.5),SLC7A11 (2.1),SLFN11 (2.1),ABLIM2 (-2.3),CAND2 (-3.2)
23	ENSG00000180370	PAK2 PPI subnetwork			0.01	SMG6 (2.3),PIEZO1 (2.2),TACR3 (-2.0)
23	ENSG00000198959	TGM2 PPI subnetwork			0.02	PIEZO1 (3.8),TMEM40 (2.7),RASSF3 (2.4),UFL1 (-2.0)
23	ENSG00000068654	POLR1A PPI subnetwork			0.02	TSR1 (3.6),CDC5L (2.5)
23	ENSG00000177889	UBE2N PPI subnetwork			0.02	XPC (2.6),BPNT1 (2.2)
23	ENSG00000112592	TBP PPI subnetwork			0.03	CDC5L (2.5),ERCC4 (2.4)
23	ENSG00000168274	HIST1H2AE PPI subnetwork			0.03	CDC5L (2.7),TSR1 (2.4)
23	ENSG00000137259	HIST1H2AB PPI subnetwork			0.03	CDC5L (2.7),TSR1 (2.4)
23	ENSG00000143190	POU2F1 PPI subnetwork			0.03	AC068353.34 (2.0)
23	ENSG00000196866	HIST1H2AD PPI subnetwork	HIST1H2AG PPI subnetwork	ZHX1 PPI subnetwork	0.03	CDC5L (2.9),TSR1 (2.1),PBX1 (2.0)
23	ENSG00000198374	HIST1H2AL PPI subnetwork			0.03	CDC5L (2.9),TSR1 (2.1),PBX1 (2.0)
23	ENSG00000196747	HIST1H2AI PPI subnetwork			0.03	CDC5L (2.9),TSR1 (2.1),PBX1 (2.0)
23	ENSG00000196787	HIST1H2AG PPI subnetwork			0.03	CDC5L (2.9),TSR1 (2.1),PBX1 (2.0)
23	ENSG00000184348	HIST1H2AK PPI subnetwork			0.03	CDC5L (2.9),TSR1 (2.1),PBX1 (2.0)
23	ENSG00000137055	PLAA PPI subnetwork			0.03	XPC (2.2)
23	ENSG00000145216	FIP1L1 PPI subnetwork			0.03	ERCC4 (2.6),HMG20A (2.5),LINC02064 (2.5)
23	ENSG00000153147	SMARCA5 PPI subnetwork			0.04	CDC5L (2.7),RUNX1 (2.4),HMG20A (2.3)
23	ENSG00000183558	HIST2H2AA3 PPI subnetwork			0.04	CDC5L (2.6),TSR1 (2.1),HMG20A (2.0)
23	ENSG00000203812	HIST2H2AA4 PPI subnetwork			0.04	CDC5L (2.6),TSR1 (2.1),HMG20A (2.0)

23	ENSG00000184486	C12ORF75 PPI subnetwork			0.004	CDC5L (3.1)
24	GO:0045624	Positive regulation of t-helper cell differentiation	Positive regulation of cd4-positive, alpha-beta t cell activation	Positive regulation of t-helper cell differentiation	0.005	RUNX1 (2.0)
24	GO:2000516	Positive regulation of cd4-positive, alpha-beta t cell			0.007	NA
24	GO:0043372	Positive regulation of cd4-positive, alpha-beta t cell			0.007	NA
24	GO:0046637	Regulation of alpha-beta t cell differentiation			0.04	RUNX1 (2.2),LINC00299 (2.1)
24	GO:0045086	Positive regulation of interleukin-2 biosynthetic			0.04	AC004944.1 (3.7),ERCC4 (2.3)
25	GO:0042744	Hydrogen peroxide catabolic process	Reticulocytosis	Hydrogen peroxide catabolic process	0.005	LINC01921 (3.5),LINC02064 (3.4),AKR1C3 (3.0),ANKRD20A11P (2.3),BPNT1 (2.2)
25	MP:0002640	Reticulocytosis			0.009	AKR1C3 (2.7),BPNT1 (-2.3)
25	GO:0015923	Mannosidase activity			0.01	UFL1 (4.7),PIEZO1 (2.3)
25	GO:0048872	Homeostasis of number of cells			0.03	AKR1C3 (3.2),ERCC4 (2.9)
25	MP:0001585	Hemolytic anemia			0.04	AKR1C3 (3.1)
26	ENSG00000088832	FKBP1A PPI subnetwork	ARF6 PPI subnetwork	FKBP1A PPI subnetwork	0.006	SLC7A11 (2.6)
26	ENSG00000165527	ARF6 PPI subnetwork			0.008	IARS2 (3.9),SLC7A11 (2.8),HDGFL1 (2.8),UFL1 (2.5)
26	ENSG00000184489	PTP4A3 PPI subnetwork			0.01	IARS2 (3.8),UFL1 (3.8),RAB3GAP2 (2.3),TSR1 (2.2),HDGFL1 (2.2),BPNT1 (2.0)
26	REACTOME ASSOCIATION OF	Reactome association of tricct with target proteins			0.02	IARS2 (2.8),HDGFL1 (2.5),UFL1 (2.4)
26	ENSG00000198824	CHAMP1 PPI subnetwork			0.02	UFL1 (2.8),TSR1 (2.5),RUNX1 (2.4),CDC5L (2.2),FOXL1 (-2.4)
26	ENSG00000139637	C12ORF10 PPI subnetwork			0.02	SLC7A11 (2.4),BPNT1 (2.4),FHL5 (-2.6)
26	ENSG00000028137	TNFRSF1B PPI subnetwork			0.03	IARS2 (2.9),SLC7A11 (2.4),UFL1 (2.3),TSR1 (2.1)
26	ENSG00000180104	EXOC3 PPI subnetwork			0.03	RAB3GAP2 (3.6),UFL1 (2.5),CDC5L (2.5),SEZ6L (2.0)
26	ENSG00000129465	RIPK3 PPI subnetwork			0.03	IARS2 (2.9)
26	ENSG00000132432	SEC61G PPI subnetwork			0.03	UFL1 (5.5)
26	ENSG00000173530	TNFRSF10D PPI subnetwork			0.03	IARS2 (3.3),XPC (2.5),SLC7A11 (2.5),TSR1 (2.2),CDC5L (2.1)
26	ENSG00000141480	ARRB2 PPI subnetwork			0.04	RASSF3 (2.3),SMG6 (2.3),UBTD1 (2.3),AC004944.1 (-2.1)
26	ENSG00000068323	TFE3 PPI subnetwork			0.04	IARS2 (3.7),SLC7A11 (2.8),CDC5L (2.0)
27	GO:0045933	Positive regulation of muscle contraction	Positive regulation of smooth muscle contraction	Positive regulation of muscle contraction	0.007	TACR3 (2.7),LINC01013 (2.6),LINC00312 (2.1),AC004944.1 (2.0)
27	GO:0045987	Positive regulation of smooth muscle			0.03	TACR3 (3.2),LINC00299 (3.0),LINC00312 (2.4),LINC01013 (2.0)
28	ENSG00000078061	ARAF PPI subnetwork	SHCBP1 PPI subnetwork	ARAF PPI subnetwork	0.007	CAND2 (2.4),SIPA1L1 (2.1),AL450226.1 (-2.3)
28	ENSG00000120509	PDZD11 PPI subnetwork			0.008	UBTD1 (2.0),AL450226.1 (-2.7)
28	ENSG00000136153	LMO7 PPI subnetwork			0.01	SIPA1L1 (2.9),TMEM40 (2.9),CAND2 (2.7),RASSF3 (2.6),GLIS1 (2.0)
28	ENSG00000131941	RHPN2 PPI subnetwork			0.02	AC068353.34 (3.0),TMEM40 (2.5),LINC01648 (2.1),PTPRM (2.0)
28	ENSG00000161800	RACGAP1 PPI subnetwork			0.02	AC068353.34 (3.4),LINC00312 (2.2)
28	ENSG00000143375	CGN PPI subnetwork			0.02	TMEM40 (2.3),AC068353.34 (2.3)
28	ENSG00000171241	SHCBP1 PPI subnetwork			0.02	AC068353.34 (2.8),TMEM40 (2.6)
28	ENSG00000168476	REEP4 PPI subnetwork			0.03	AC068353.34 (2.4),TMEM40 (2.4),LINC01648 (2.0),PTPRM (2.0)
28	ENSG00000125731	SH2D3A PPI subnetwork			0.03	TMEM40 (2.7),AC068353.34 (2.6),SETD4 (-2.0)
28	ENSG00000115963	RND3 PPI subnetwork			0.04	TMEM40 (2.8),AC068353.34 (2.7),PTPRM (2.0),SRR (-2.7)
28	ENSG00000189319	FAM53B PPI subnetwork			0.04	TMEM40 (2.5),AC068353.34 (2.5)
28	ENSG00000163362	C1ORF106 PPI subnetwork			0.04	TMEM40 (2.5),AC068353.34 (2.3),PTPRM (2.0)
28	ENSG00000138771	SHROOM3 PPI subnetwork			0.04	AC068353.34 (2.5),TMEM40 (2.5),LINC01648 (2.0),SETD4 (-2.0)
28	ENSG00000197555	SIPA1L1 PPI subnetwork			0.04	AC068353.34 (2.5),TMEM40 (2.4),LINC01648 (2.0)
28	GO:0030216	Keratinocyte differentiation			0.04	TMEM40 (4.9)
28	ENSG00000075413	MARK3 PPI subnetwork			0.04	FHL5 (3.1),AC068353.34 (2.2)
28	ENSG00000171219	CDC42BPG PPI subnetwork			0.04	AC068353.34 (2.8),TMEM40 (2.3),LINC01648 (2.3)
29	GO:0017119	Golgi transport complex	Golgi transport complex	Golgi transport complex	0.008	UFL1 (4.4),RAB3GAP2 (4.2),IARS2 (2.8),BPNT1 (2.3)
29	ENSG00000135775	COG2 PPI subnetwork			0.02	RAB3GAP2 (2.2),LINC01013 (2.1),XPC (2.0)
29	ENSG00000164597	COG5 PPI subnetwork			0.02	RAB3GAP2 (2.6)
29	GO:0007030	Golgi organization			0.04	UFL1 (4.7),RAB3GAP2 (3.6)

29	ENSG00000082108	SCFD1 PPI subnetwork			0.04	UFL1 (3.9),FAM1 (2.5)
29	ENSG00000103051	COG4 PPI subnetwork			0.04	UFL1 (5.0),RAB3GAP2 (3.2),GLIS1 (2.8)
30	ENSG00000184575	XPOT PPI subnetwork			0.008	CDC5L (2.8),PTPRM (2.0)
30	ENSG00000090054	SPTLC1 PPI subnetwork	SPTLC1 PPI subnetwork	XPOT PPI subnetwork	0.03	IARS2 (3.0),CDC5L (2.7),TSR1 (2.4)
30	ENSG00000137767	SQRDL PPI subnetwork			0.04	PIEZO1 (2.6),UFL1 (2.0)
31	MP:0001147	Small testis			0.008	HDGFL1 (5.1),LINC01921 (2.6),ANKRD20A11P (2.3)
31	MP:0008261	Arrest of male meiosis	Azoospermia	Small testis	0.04	HDGFL1 (5.0)
31	MP:0005159	Azoospermia			0.04	HDGFL1 (5.9),LINC01921 (2.0)
32	ENSG00000082781	ITGB5 PPI subnetwork			0.009	PIEZO1 (2.9),LMCD1 (2.9),LINC00973 (2.6),UBTD1 (2.3),PEAK1 (2.1),PTPRM (2.0),LINC01648 (-2.4),BPNT1 (-2.6)
32	ENSG00000115221	ITGB6 PPI subnetwork			0.01	PIEZO1 (2.9),LINC01648 (-2.1)
32	ENSG00000105855	ITGB8 PPI subnetwork	ITGB6 PPI subnetwork	ITGB5 PPI subnetwork	0.02	PIEZO1 (2.8),FOXL1 (2.3)
32	GO:0008305	Integrin complex			0.04	TMEM40 (3.4),PIEZO1 (2.8),RASSF3 (2.0),LINC01013 (-2.7)
32	MP:0000069	Kyphoscoliosis			0.04	PIEZO1 (2.6),PTPRM (2.0)
33	GO:0050708	Regulation of protein secretion			0.009	NA
33	GO:0050707	Regulation of cytokine secretion			0.02	LINC00299 (3.8),LINC01013 (2.8),SLFN11 (2.3)
33	GO:0051048	Negative regulation of secretion	Regulation of protein secretion	Regulation of protein secretion	0.04	LINC01648 (2.7),SEZ6L (2.1),HMG20A (-2.3)
33	GO:0009306	Protein secretion			0.04	NA
33	GO:0070201	Regulation of establishment of protein			0.04	NA
34	ENSG00000197905	TEAD4 PPI subnetwork			0.01	FOXL1 (3.2),SLFN11 (2.9),PTPRM (2.2),GLIS1 (2.1),UBTD1 (2.0),TMEM40 (2.0),ABLIM2 (-2.0),ANKRD20A11P (-2.2)
34	MP:0005205	Abnormal eye anterior chamber morphology			0.01	MN1 (3.4),LINC00973 (3.0),LMCD1 (2.1),LINC02064 (2.0)
34	MP:0001322	Abnormal iris morphology	Abnormal eye anterior chamber morphology	TEAD4 PPI subnetwork	0.02	NA
34	MP:0005542	Corneal vascularization			0.02	FASLG (2.3),HMG20A (-2.1)
34	MP:0004624	Abnormal thoracic cage morphology			0.03	MN1 (2.3),PBX1 (2.2)
35	MP:0004876	Decreased mean systemic arterial blood pressure			0.01	FHL5 (3.8),FOXL1 (2.8),PTPRM (2.4),TACR3 (2.3),LINC00299 (2.3)
35	MP:0008874	Decreased physiological sensitivity to xenobiotic			0.02	SLC7A11 (3.2),FOXL1 (2.6),ABLIM2 (2.0)
35	GO:0035150	Regulation of tube size	Regulation of tube size	Decreased mean systemic arterial blood pressure	0.03	FHL5 (3.5),FOXL1 (3.0)
35	GO:0050880	Regulation of blood vessel size			0.03	FHL5 (3.6),FOXL1 (3.0)
35	GO:0042310	Vasoconstriction			0.03	FHL5 (4.0),FOXL1 (3.2)
35	MP:0001613	Abnormal vasodilation			0.04	FHL5 (2.7),PIEZO1 (2.2),PTPRM (2.1),CBR1 (2.0),HMG20A (-2.0)
36	ENSG00000166930	MS4A5 PPI subnetwork			0.01	NA
36	ENSG00000173702	MUC13 PPI subnetwork	MS4A5 PPI subnetwork	MS4A5 PPI subnetwork	0.01	NA
36	ENSG00000169306	IL1RAPL1 PPI subnetwork			0.01	ANKRD20A11P (2.3)
36	ENSG00000198478	SH3BGRL2 PPI subnetwork			0.03	NA

Supplementary Table 4. Enrichment of tissue and cell types from DEPICT. Suggestive enrichment (FDR < 0.20) was observed for 21 categories of tissues and/or cell types. Two categories where GILS1 was highly expressed were added. Genes listed were annotated to a given tissue/cell types and are within an associated locus. Each gene has a z-score representing the gene's likelihood to be highly expressed in the tissue/cell types. Genes displayed are those with a |z-score| > 1.96 that corresponds to a P-value<0.05 as the minimum threshold, or the gene with higher absolute value of z-score (*). MeSH: Medical Subject Headings, FDR: false discovery rate.

MeSH term	Name	MeSH first level term	MeSH second level term	P value	FDR	Contributing genes
A15.145.846	Serum	Hemic and Immune Systems	Blood	2.61E-03	<0.20	LINC01013 (2.9),UBTD1 (2.6),ERCC4 (2.0),UFL1 (-2.1)
A10.690.467	Muscle Smooth	Tissues	Muscles	3.94E-03	<0.20	FOXL1 (3.2)
A07.231	Blood Vessels	Cardiovascular System	Blood Vessels	4.25E-03	<0.20	LINC01013 (4.1),UBTD1 (2.3)
A11.620.520	Myocytes Smooth Muscle	Cells	Muscle Cells	5.36E-03	<0.20	FOXL1 (4.8),TACR3 (2.6),LINC00973 (2.4),LINC01648 (2.1),SRR (2.0)
A11.620	Muscle Cells	Cells	Muscle Cells	5.36E-03	<0.20	FOXL1 (4.8),TACR3 (2.6),LINC00973 (2.4),LINC01648 (2.1),SRR (2.0)
A11.872.190.260	Embryoid Bodies	Cells	Stem Cells	5.92E-03	<0.20	TACR3 (2.6)
A11.329.228	Fibroblasts	Cells	Connective Tissue Cells	6.72E-03	<0.20	FOXL1 (2.2),SRR (2.1)
A11.329	Connective Tissue Cells	Cells	Connective Tissue Cells	8.20E-03	<0.20	*SLC7A11 (1.3)
A07.231.908	Veins	Cardiovascular System	Blood Vessels	9.70E-03	<0.20	LINC01013 (5.8),UBTD1 (2.6),SLC7A11 (2.0)
A10.165.450.300.425	Keloid	Tissues	Connective Tissue	1.00E-02	<0.20	MN1 (3.4),PEAK1 (3.1),XPC (2.7),HDGFL1 (2.5),LINC02064 (2.3),**GLIS1 (2.3)**,LINC01648 (2.3),SLC7A11 (2.0),TRIB1 (-2.5)
A05.360.444.492	Penis	Urogenital System	Genitalia	1.00E-02	<0.20	*LINC01921 (1.9)
A05.360.444.492.362	Foreskin	Urogenital System	Genitalia	1.00E-02	<0.20	LINC01921 (2.0)
A07.231.908.670.874	Umbilical Veins	Cardiovascular System	Blood Vessels	1.00E-02	<0.20	LINC01013 (6.3),UBTD1 (2.8),SLC7A11 (2.2)
A07.231.908.670	Portal System	Cardiovascular System	Blood Vessels	1.00E-02	<0.20	LINC01013 (6.3),UBTD1 (2.8),SLC7A11 (2.2)
A10.165.450.300	Cicatrix	Tissues	Connective Tissue	1.00E-02	<0.20	MN1 (3.1),PEAK1 (2.9),HDGFL1 (2.4),XPC (2.3),LINC02064 (2.2),LINC01648 (2.0),TRIB1 (-2.3)
A10.165.450	Granulation Tissue	Tissues	Connective Tissue	1.00E-02	<0.20	MN1 (3.1),PEAK1 (2.9),HDGFL1 (2.4),XPC (2.3),LINC02064 (2.2),LINC01648 (2.0),TRIB1 (-2.3)
A11.436.275	Endothelial Cells	Cells	Epithelial Cells	1.00E-02	<0.20	LINC01013 (3.8),UBTD1 (2.4)
A10.690	Muscles	Tissues	Muscles	2.00E-02	<0.20	LINC00312 (3.3),MN1 (2.7),CAND2 (2.4),LMCD1 (2.4),ABLIM2 (2.4),LINC02064 (2.3)
A07.231.114	Arteries	Cardiovascular System	Blood Vessels	2.00E-02	<0.20	FOXL1 (2.7),FHL5 (2.4),LINC01013 (2.0)
A11.329.629	Osteoblasts	Cells	Connective Tissue Cells	2.00E-02	<0.20	LINC01013 (4.1),ERCC4 (2.9),**GLIS1 (2.8)**,PEAK1 (2.7),SLC7A11 (2.5),LMCD1 (2.3),SIPA1L1 (2.1),PTPRM (2.1),SRR (2.0)
A11.872.580	Mesenchymal Stem Cells	Cells	Stem Cells	3.00E-02	<0.20	SRR (2.3),**GLIS1 (2.3)**,PEAK1 (2.2)
A07.541.510.110	Aortic Valve	Cardiovascular System	Heart	1.30E-01	>=0.20	**GLIS1 (6.2)**,LINC00312 (2.2),PTPRM (1.9)
A07.541.510	Heart Valves	Cardiovascular System	Heart	1.30E-01	>=0.20	**GLIS1 (6.2)**,LINC00312 (2.2),PTPRM (1.9)

Supplementary Table 5. The 6 enriched gene sets (FDR < 0.05) where *GLIS1* was found to be the best-ranked gene by i-GSEA4GWAS. Enrichment significance (p-value and FDR) of pathways / gene sets are displayed. Genes where SNPs are associated with MVP (P-value < 0.01 from the GWAS meta-analysis) are listed from the most to the less significant.

Gene Set Name	P-value	FDR	Significant genes in the pathway/gene set
GO: Positive Regulation of Transcription DNA Dependent	0.001	0.001	GLIS1; RUNX1; SMAD3; MKL2; SQSTM1; NUFIP1; NRIP1; PAX8; GLIS3; TBX5; CREB5; ESRRG; MDFIC; RXRA; MYO6; ARNTL
GO: Positive Regulation of RNA Metabolic Process	0.001	0.002	GLIS1; RUNX1; SMAD3; MKL2; SQSTM1; NUFIP1; NRIP1; PAX8; GLIS3; TBX5; CREB5; ESRRG; MDFIC; RXRA; MYO6; ARNTL
GO: Positive Regulation of Transcription	0.001	0.003	GLIS1; RUNX1; SMAD3; HIVEP3; MKL2; IRF4; SQSTM1; NUFIP1; NRIP1; PAX8; GLIS3; TBX5; CREB5; ESRRG; MDFIC; RXRA; NFATC2; MYO6; ARNTL; GLIS2
GO: Positive Regulation of Nucleobasenucleosidenucleotide and Nucleic Acid Metabolic Process	0.001	0.004	GLIS1; RUNX1; SMAD3; HIVEP3; MKL2; IRF4; SQSTM1; NUFIP1; NRIP1; PAX8; GLIS3; TBX5; CREB5; ESRRG; MDFIC; ABCA1; RXRA; NFATC2; MYO6; ARNTL; GLIS2
GO: Positive Regulation of Transcription From RNA Polymerase Ii Promoter	0.001	0.009	GLIS1; RUNX1; SMAD3; MKL2; SQSTM1; NUFIP1; NRIP1; GLIS3; RXRA; MYO6; ARNTL
GO: Negative Regulation of Transcription	0.002	0.011	GLIS1; SMAD3; NFX1; MECP2; NRIP1; TGIF1; GLIS3; BCOR; VDR; TIMELESS; RFX3; ARHGAP35; SUDS3; GLIS2
GO: Negative Regulation of RNA Metabolic Process	0.023	0.089	GLIS1; SMAD3; NFX1; MECP2; NRIP1; TGIF1; GLIS3; BCOR
GO: Specific RNA Polymerase Ii Transcription Factor Activity	0.044	0.119	GLIS1; GLIS3; SOX9; RXRA

Supplementary Table 6. List of 38 significant genes in the Hedgehog signalling pathway (p-value = 0.015) identified by i-GSEA4GWAS. Chr: chromosome. Position from the Hg19 human genome assembly. The rs number and p-value from the GWAS meta-analysis are indicated for the top associated variant in the contributing genes.

Gene Name	Chr	Gene Start	Gene End	Variant Position	Associated variant	Association P-value in GWAS
WNT5B	12	1639057	1756409	1749905	rs2270038	2.07E-03
WNT9B	17	44910567	44964096	44929325	rs56344308	2.43E-03
BMP6	6	7727030	7881655	7759886	rs270402	2.78E-03
WNT4	1	22446461	22470462	22495261	rs2807352	2.79E-03
HHIP	4	145567173	145666423	145668076	rs7699504	3.18E-03
WNT11	11	75897369	75921780	75882890	rs60995418	3.58E-03
WNT7B	22	46316242	46373009	46279265	rs28371399	4.03E-03
BMP5	6	55618443	55740362	55601702	rs2093373	4.22E-03
WNT8A	5	137419581	137428054	137405781	rs7721939	4.49E-03
CSNK1E	22	38686697	38794527	38741793	rs5750597	4.75E-03
WNT16	7	120965421	120981158	120963965	rs3757552	5.08E-03
PTCH1	9	98205262	98279339	98336854	rs111896098	5.22E-03
CSNK1G2	19	1941188	1981337	1941121	rs112624009	6.95E-03
RAB23	6	57053607	57087078	57117624	rs2064245	7.47E-03
GSK3B	3	119540170	119813264	119598409	rs62264713	8.19E-03
SUFU	10	104263744	104393292	104397323	rs186502091	8.40E-03
GLI2	2	121493199	121750229	121567079	rs55922034	8.59E-03
WNT5A	3	55499743	55523973	55450917	rs35262007	9.14E-03
WNT7A	3	13857755	13921618	13892543	rs17038710	1.07E-02
PRKACB	1	84543745	84704181	84513318	rs76776503	1.10E-02
WNT9A	1	228106357	228135631	228084158	rs883583	1.12E-02
SHH	7	155592680	155604967	155642784	rs74319244	1.29E-02
LRP2	2	169983619	170219195	170184814	rs9710881	1.31E-02
PRKX	X	3522411	3631649	3609554	rs56390349	1.39E-02
GAS1	9	89559279	89562104	89575055	rs1927083	1.52E-02
BMP2	20	6748311	6760927	6676723	rs6054474	1.53E-02
BTRC	10	103113820	103317078	103126789	rs11593144	1.55E-02
WNT3A	1	228194752	228248961	228226602	rs76082193	1.56E-02
BMP7	20	55743804	55841685	55765481	rs138980309	1.57E-02

GLI3	7	42000548	42277469	41979946	rs13438132	1.71E-02
BMP8B	1	40222854	40254533	40229964	rs187082776	1.91E-02
SMO	7	128828713	128853386	128826811	rs6962740	2.24E-02
FBXW11	5	171288553	171433877	171268203	rs6891834	2.41E-02
PTCH2	1	45285516	45308735	45305492	rs11573543	2.65E-02
BMP4	14	54416454	54425479	54492146	rs72680588	2.76E-02
WNT2	7	116916685	116963343	116980769	rs2023707	3.34E-02
PRKACA	19	14202500	14228896	14209826	rs79646099	3.77E-02
CSNK1A1L	13	37677398	37679803	37660392	rs1199979	4.15E-02

Supplementary Table 7. Association of rs1879734 in the discovery meta-analysis and follow-up in four case-control studies

	MVP Cases	Controls	OR [95%CI]	P-value
MVP-Nantes	489	873	1.44 [1.18-1.76]	3.20×10^{-4}
MVP-France	953	1566	1.23 [1.07-1.41]	3.32×10^{-3}
Discovery meta-analysis				6.80×10^{-6}
MGH/FHS	699	5575	1.21 [1.05-1.40]	1.40×10^{-4}
HEGP-Surgery	450	820	1.22 [1.01-1.48]	0.037
PROMESA	171	282	1.06 [0.77-1.46]	0.88
IUPCQ	102	102	0.89 [0.56-1.44]	0.64
Global meta-analysis	2864	9218	1.23 [1.14-1.33]	1.29×10^{-7}
Heterogeneity				0.41

* N: Numbers; OR: odds ratio; 95%CI: 95% confidence interval

† Alleles are indexed to the forward strand of NCBI Build 37

‡ P-values reported are two-sided and based on an inverse-variance weighted meta-analysis model (fixed effects)

§ The P-value for heterogeneity corresponds to Cochran's Q statistic.

Supplementary Table 8 : Functional annotation of 111 SNPs at the GLIS1 locus. We provided functional annotations for all nominally associated (P-value < 0.05) SNPs located within ±500Kb and in linkage desequilibrium (r2 > 0.5) with the top associated SNP rs1879743 at the GLIS1 locus. We used 111 SNPs, all were located in intron X as input to Variant Annotation Integrator (VAI), a tool box avialable at the UCSC Genome Browser to annotate SNPs. We provide location in DNase hypersensitive region (DNASEV), cell types and descriptive and the karyotype of cells, in addition to location at transcription factor binding site (TFBS) using ENCODE data. When available, we indicated if SNPs are expression quantitative trait loci (eQTLs) in atrial appendage tissue using GTeX data

SNP	Position (bp on Chr1)	DNASEV in cell types	Cell types name for DNASEV	Detail of cell types for DNASEV	Karyotype of DNASEV	TFBS	LD with rs1879734 (r^2)	eQTL P-value	MVP GWAS P-value
rs28743002	54125615	1	Epithelium	Epithelium cell from a lung carcinoma	Cancer		0.69	-	7.85E-05
rs17383325	54126552	2	Skin	Fetal buttock fibroblast	Normal		0.73	-	6.50E-05
rs12026298	54126754						0.73	-	7.47E-05
rs11206187	54127969	3	Lung	Fetal lung fibroblast	Normal	CTCF	0.73	-	5.24E-05
rs2950241	54129591	20	Blood	Lymphoblastoid	Normal	NR3C1	0.95	0.009	1.79E-05
rs12405835	54129732						0.73	0.039	4.65E-05
rs3006905	54130119						0.95	0.033	1.38E-05
rs4927025	54131753	4	Skin	Toe fibroblast	Normal		0.77	0.024	2.72E-05
rs6658469	54133345						0.77	0.046	3.00E-05
rs67319173	54133723						0.77	0.024	2.69E-05
rs1465037	54135078						0.77	0.025	2.67E-05
rs1554750	54135334	30	Brain	Gliblastoma	Cancer	TCF7L2, BHLHE40, USF1	0.77	0.024	2.66E-05
rs2141080	54138489						1.00	-	6.83E-06
rs2141081	54138552	1	Epithelium	Epithelium cell from a lung carcinoma	Cancer		0.91	-	6.84E-06
rs2141082	54138698						1.00	-	8.79E-06
rs1879734	54138854					GATA2	1.00	0.048	6.80E-06
rs4927026	54140843	2	Skin	Fetal buttock fibroblast	Normal		0.73	-	5.35E-05
rs4927027	54141274	3	Lung	Fetal lung fibroblast	Normal		1.00	-	9.52E-06
rs7543166	54143423						1.00	-	1.24E-05
rs2950244	54143759						0.95	-	1.16E-05
rs55888542	54144297	10	Brain	Neuroblastoma	Cancer		1.00	-	1.21E-05
rs7549345	54145288	15	Blood	Acute megakaryocytic leukemia cell	Cancer		0.65	0.046	3.10E-05
rs2948044	54149577						0.50	0.022	6.73E-04
rs7518284	54150890	68	Epithelium	Renal epithelial cell	Normal	EZH2, SUZ12, ZNF263, TCF7L2, KAP1	0.77	0.045	3.64E-05
rs3006882	54152487						0.95	-	2.53E-05
rs2948042	54152703						0.95	-	2.11E-05
rs882879	54154726					CTCF	0.95	-	2.39E-05
rs3013771	54155902	14	Blood	Chronic lymphocytic leukemia cell	Cancer	JUND, NR2F2, GATA2, MYC, POLR2A, TEAD4, CCNT2, RCOR1, TAL1	1.00	-	2.60E-05
rs3013770	54155959					JUND, MYC, POLR2A, TEAD4, CCNT2	0.86	-	4.66E-05
rs3013769	54156323	10	Brain	Neuroblastoma	Cancer	REST	0.95	-	2.64E-05
rs3013768	54156434	10	Brain	Neuroblastoma	Cancer		0.95	-	2.65E-05
rs2948043	54156443	10	Brain	Neuroblastoma	Cancer		0.95	-	2.65E-05
rs12082358	54156624						0.73	-	3.40E-05
rs3013767	54156757						0.95	-	2.44E-05
rs72660653	54157283	9	Blood vessel	Aortic smooth muscle cells	Normal		0.73	-	3.94E-05
rs72660654	54157287	9	Blood vessel	Aortic smooth muscle cells	Normal		0.73	-	3.94E-05
rs3013766	54157498	2	Skin	Fetal buttock fibroblast	Normal		0.82	-	2.97E-05

rs12088451	54157640						0.73	0.047	4.04E-05
rs12076192	54157742						0.68	0.041	4.10E-05
rs11206191	54158102	2	Skin	Fetal buttock fibroblast	Normal		0.69	0.041	4.20E-05
rs7555617	54158155	2	Skin	Fetal buttock fibroblast	Normal		0.73	-	3.82E-05
rs7555711	54158220						0.73	0.041	4.23E-05
rs7555908	54158462						0.73	0.041	4.28E-05
rs7544346	54158475						0.73		3.89E-05
rs11206192	54158690						0.73	0.041	4.29E-05
rs11206193	54158762						0.65		4.61E-05
rs6658245	54159005					CTCF	0.73	0.041	4.20E-05
rs6691833	54159079	17	Fetal membrane	Chorion cells	Normal	CTCF	0.73	0.046	4.20E-05
rs6694399	54159198	17	Fetal membrane	Chorion cells	Normal	CTCF	0.73	0.041	4.16E-05
rs56368850	54159441						0.73		5.32E-05
rs10888803	54160235						0.73	0.041	4.28E-05
rs12044337	54160558	1	Epithelium	Epithelium cell from a lung carcinoma	Cancer		0.73	0.041	3.74E-04
rs10888804	54160570	1	Epithelium	Epithelium cell from a lung carcinoma	Cancer		0.73	0.041	3.79E-04
rs10888805	54160571	1	Epithelium	Epithelium cell from a lung carcinoma	Cancer		0.73	0.041	3.80E-04
rs61184814	54160709	6	Skin	Skin fibroblasts	Normal		0.73	0.041	4.31E-05
rs6673378	54160812	6	Skin	Skin fibroblasts	Normal		0.69		4.32E-05
rs6665257	54161087						0.73	-	4.09E-05
rs6695509	54161112						0.69	0.041	4.34E-05
rs4927028	54161545	8	Blood vessel	Aortic adventitial fibroblast cells	Normal		0.73	0.041	4.37E-05
rs6661728	54162471						0.73	0.041	5.09E-05
rs59020945	54162635						0.73	-	4.93E-05
rs7554633	54163128	3	Lung	Fetal lung fibroblast	Normal		0.65	0.041	4.89E-05
rs7512093	54163207	3	Lung	Fetal lung fibroblast	Normal		0.73	0.041	4.89E-05
rs7519965	54163251						0.73	0.042	4.59E-05
rs11206195	54164156	45	Epithelium	Esophageal epithelial cells	Normal		0.73	0.041	2.55E-04
rs12041221	54164165	45	Epithelium	Esophageal epithelial cells	Normal		0.73	0.028	3.78E-05
rs1615037	54165541						0.91	-	3.16E-05
rs11206196	54165624						0.73	0.041	4.48E-05
rs11206197	54165831						0.73	0.041	5.09E-05
rs11206198	54165910						0.73	0.028	4.90E-05
rs6588486	54166777						0.73	0.028	4.77E-05
rs6588487	54166790						0.73	0.028	4.35E-05
rs11206200	54168545						0.73	0.016	5.00E-05
rs12040081	54168609	1	Epithelium	Epithelium cell from a lung carcinoma	Cancer		0.69	0.028	4.97E-05
rs4927029	54169455						0.73	0.042	4.57E-05
rs12097598	54170801						0.69	0.028	4.99E-05
rs11206201	54170963	5	Gingival	Gum tissue fibroblast	Normal		0.69	0.028	3.48E-05
rs55786134	54171981	4	Skin	Toe fibroblast	Normal		0.69	0.028	1.04E-04

rs17409178	54173551						0.64	0.049	2.36E-05
rs11206202	54174607						0.64	0.034	2.79E-05
rs2694594	54174846	1	Epithelium	Epithelium cell from a lung carcinoma tissue	Cancer		0.74	-	3.58E-05
rs1780392	54175302						0.82	-	3.56E-05
rs12025857	54175360						0.64	0.034	2.78E-05
rs17109200	54175548	3	Lung	Fetal lung fibroblast	Normal		0.64	0.050	2.54E-05
rs12043690	54175629	3	Lung	Fetal lung fibroblast	Normal		0.64	0.034	2.78E-05
rs1879731	54175861						0.52	0.034	2.78E-05
rs1879732	54175909						0.64	0.034	2.78E-05
rs1879733	54175971						0.64	0.034	2.78E-05
rs7539908	54177279						0.73	-	7.54E-05
rs1780391	54177708	1	Epithelium	Epithelium cell from a lung carcinoma	Cancer		0.82	-	3.57E-05
rs12080903	54178270	9	Blood vessel	Aortic smooth muscle cells	Normal		0.64	0.050	2.59E-05
rs12080993	54178501						0.64	0.050	2.59E-05
rs2186041	54180287						0.64	0.042	2.76E-05
rs1614395	54181668	3	Lung	Fetal lung fibroblast	Normal		0.78	-	3.87E-05
rs1526911	54181942						0.64	-	2.65E-05
rs1526910	54182174						0.64	-	2.86E-05
rs56087009	54182502						0.59	-	2.69E-05
rs4927031	54183191						0.64	-	2.95E-05
rs72660679	54183531					GATA2	0.73	-	9.36E-05
rs12043520	54184097					CTCF	0.64	-	3.03E-05
rs4927032	54186489						0.73	-	3.07E-05
rs17388507	54187575						0.73	-	2.88E-04
rs11206211	54188096	18	Skin	Fibroblast from parkinson's disease	Normal		0.64	-	1.03E-04
rs6658137	54188994						0.64	-	1.11E-04
rs12059719	54190333						0.73	-	1.22E-04
rs12091931	54191103	61	Epithelium	Non-pigmentn ciliary epithelial cells	Normal	JUND, MAFF, MAFK, UBTF, MAZ, CTCF, ZBTB7A, TBP	0.64	0.036	1.32E-04
rs72660683	54191811						0.73	-	3.11E-04
rs7554636	54193023						0.60	-	1.29E-04
rs910298	54193701	6	Skin	Skin fibroblasts	Normal		0.73	-	3.76E-04
rs702491	54194992	4	Skin	Toe fibroblast	Normal	CTCF	0.51	-	4.27E-03
rs12406120	54195034	4	Skin	Toe fibroblast	Normal		0.61	0.032	4.60E-03