Published in final edited form as:

Circ Genom Precis Med. 2019 May ; 12(5): e002497. doi:10.1161/CIRCGEN.119.002497.

GWAS-driven Gene-set Analyses, Genetic and Functional Follow-up Suggest *GLIS1* as a Susceptibility Gene for Mitral Valve Prolapse

Mengyao Yu, Msc[1,2], **Adrien Georges, PhD**[1,2], **Nathan R. Tucker, PhD**[3,4], **Sergiy Kyryachenko, PhD**[1,2], **Katelyn Toomer, PhD**[5], **Jean-Jacques Schott, PhD**[6,7,8], **Francesca N. Delling, MD**[9], **Leticia Fernandez-Friera, MD**[10,11], **Jorge Solis, MD**[10,11], **Patrick T. Ellinor, MD**[3,4], **Robert A. Levine, MD**[12], **Susan A. Slaugenhaupt, PhD**[13], **Albert A. Hagège, MD, PhD**[1,2,14], **Christian Dina, PhD**[6,7,8], **Xavier Jeunemaitre, MD, PhD**[1,2,15], **David J. Milan, MD**[3,*], **Russell A. Norris, PhD**[5,*], and **Nabila Bouatia-Naji, PhD**[1,2,*]

[1]INSERM, UMR970, Paris Cardiovascular Research Center;

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

[3]Cardiovascular Research Center, Cardiology Division, Massachusetts General Hospital, Harvard Medical School, Boston;

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

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

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

[7]CNRS, UMR 6291, Nantes, France;

[8]Université de Nantes, Nantes, France;

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

[10]HM Hospitales-Centro Integral de Enfermedades Cardiovasculares HM-CIEC;

[11]Centro Nacional de Investigaciones Cardiovasculares Carlos III (CNIC), Madrid, Spain;

[12]Cardiac Ultrasound Laboratory, Cardiology Division, Massachusetts General Hospital, Harvard Medical School, Boston;

[13]Center for Human Genetic Research, Massachusetts General Hospital and Department of Neurology, Harvard Medical School, Boston, MA;

[14]Assistance Publique – Hôpitaux de Paris, Department of Cardiology, Paris, France

Corresponding author: Nabila BOUATIA-NAJI, PhD, Paris Cardiovascular Research Center, INSERM UMR970, 56 Rue Leblanc, F-75015, Paris, France, Tel. +33 1 53 98 79 95, Fax. +33 1 53 98 79 52, nabila.bouatia-naji@inserm.fr.

*joint last authors

Disclosures: None

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

Abstract

Background: Mitral valve prolapse (MVP) is a common heart valve disease, the most frequent indication for valve repair and/or replacement. MVP is characterized by excess extracellular matrix secretion and cellular disorganization which leads to bulky valves that are unable to co-apt correctly during ventricular systole resulting in mitral regurgitation, and it is associated with sudden cardiac death. Here we aim to characterize globally the biological mechanisms underlying genetic susceptibility to MVP to better characterize its triggering mechanisms.

Methods: Using pathway enrichment tools i-GSEA4GWAS and DEPICT applied to MVP GWAS. We followed-up the association with MVP in an independent dataset of cases and controls. This research was conducted using the UK Biobank Resource. Immunohistochemistry staining for Glis-1 was conducted in developing heart of mice. Knockdown (KD) of *Glis1* using morpholinos were performed in zebrafish animals 72 hours post fertilization.

Results: We show that genes at risk loci are involved in biological functions relevant to actin filament organization, cytoskeleton biology and cardiac development. The enrichment for positive regulation of transcription, cell proliferation and migration motivated the follow-up of *GLIS1*, a transcription factor from the Krüppel-like zinc finger family. In combination with previously available data, we now report a genome-wide significant association with MVP (OR=1.20, P=4.36 $\times 10^{-10}$). indicated that *Glis1* is expressed during embryonic development predominantly in nuclei of endothelial and interstitial cells of mitral valves in mouse. We also show that *Glis1* knockdown causes atrioventricular regurgitation in developing hearts in zebrafish.

Conclusions: Our findings define globally molecular and cellular mechanisms underlying common genetic susceptibility to MVP and implicate established and unprecedented mechanisms. Through the *GLIS1* association and function, we point at regulatory functions during cardiac development as common mechanisms to mitral valve degeneration.

Journal Subject Terms:

Valvular Heart Disease; Genetic, Association Studies; Developmental biology; Computational Biology; Gene Expression and Regulation

Keywords

mitral valve; genetics, bioinformatics; genetics, association studies; Genome Wide Association Study; developmental biology

Introduction

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.[4,5] MVP in

associated with infective endocarditis in native valves[6] and is an established a risk factor for heart failure and sudden death in community-based studies.[2,4,7] Prospective studies have showed that asymptomatic patients with low-risk presentation (e.g., moderate mitral regurgitation and ejection fraction ≥50%) can develop adverse MVP-related events, indicating wide heterogeneity in outcomes among individuals.[4,8]

The process that leads to MVP and myxomatous degeneration is poorly understood. Previous reports have suggested that a potential mechanism is activation of quiescent valvular interstitial cells 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.[8] However, the triggering mechanisms of myxomatous degeneration are still to be identified, as are the etiological mechanisms of advert valvular complications in patients.

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*)[9] followed by research in mice[10] 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.[11] 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.[12] *DCHS1* deficiency in valvular interstitial cells altered migration and cellular patterning from patients and provoked loss of cell polarity during valve development in mice and zebrafish models.[12] 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.[13] 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*).[13] However, a broad assessment of the biological functions and potential mechanisms provided by the GWAS data is still missing.

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 unknown mechanisms relevant to MVP onset. We applied computational-based analytical methods to the currently available GWAS data to 1) highlight globally enriched biological mechanisms at MVP loci, 2) characterize their expression pattern in tissues and 3) identify biologically pertinent genes for follow-up.

Methods

Details on the populations and methods are available in the Supplementary methods. Because of the sensitive nature of the data collected for this study, requests to access the human MVP datasets from qualified researchers trained in human subject confidentiality protocols may be sent to the study leaders of each cohort and/or to the corresponding author. The experimental data that support the findings of this study are available from the corresponding author upon reasonable request. Clinical and genetic data were obtained after

full written consents in each cohort studied and after approval of studies protocols from local ethics committees (See and[14] for details). All animal experiments were performed under protocols approved by the Institutional Animal Care and Use Committee, Medical University of South Carolina. Prior to cardiac resection, mice were euthanized in accordance with the Guide for the Care and Use of Laboratory Animals (NIH Publication No. 85–23, revised 1996.). Zebrafish experiments were performed in accordance with approved Institutional Animal Care and Use Committee (IACUC) protocols.

Results

Gene-set enrichment analyses at MVP loci

We first employed i-GSEA4GWAS (Improved Gene Set Enrichment Analysis 4 GWAS), a tool for meta-analysis of GWAS summary statistics that presents the advantage of calculating enrichment based on all association results generated by the meta-analysis of GWAS, without restricting the analysis to a significance threshold.[15] This method highlighted 244 pathways as significantly enriched (false discovery rate (FDR) < 0.05) for MVP (Supplementary Table 1). The most enriched pathway was Biocarta EDG-1 pathway with 26 out of 27 genes harbouring associated variants. EDG-1, also known as Sphingosine 1-phosphate receptor (S1PR1), is a G-protein receptor reported to interact with Filamin A,[16] which is mutated in familial forms of MVP.[17] Globally, enriched gene-sets clustered into four main groups related to valve biology according to a semantic similarity between GO Biological Processes terms using REVIGO webserver.[18] Two groups supported known mechanisms in MVP and contained terms “actin filament organization” and “membrane organization”, “cytoskeleton-dependent intracellular transport” and “cell migration” (Figure 1). Interestingly, one cluster contained several enriched terms related to regulation of transcription (e.g “positive regulation of transcription”, DNA-template). Finally, this analysis highlights several enriched GO terms for developmental mechanisms including “heart development” and “vasculature development”, in addition to related terms “regulation of angiogenesis” (Figure 1, Supplementary Table 1).

We then applied the integrative method Data-driven Expression Prioritized Integration for Complex Traits (*DEPICT*) that allows to consider larger resources of gene-sets and takes into account tissue expression and mouse knockout phenotypes.[19] We used the recommended association cut-off (P_{GWAS}-value $<10^{-5}$) that identified 39 independent MVP loci harboring 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 2). Overall, these enrichment analyses and clustering were consistent with the i-GSEA4GWAS analyses, and supports cell-cell and cell-extracellular matrix interactions, vascular development and regulatory genes to be highly relevant for MVP genetic susceptibility (Supplementary Table 2).

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 (Supplementary Figure 1, Supplementary Table 3).

Prioritization of *GLIS1* as a susceptibility gene for MVP

Prioritization from contribution to enrichment analyses—Among the loci that contributed to the enrichment analyses, we provide a focus on an association signal located on chromosome 1 that deserved prioritization in the light of genetic association results, gene function candidacy and clues from the results obtained in the enrichment analyses. The association context at the *GLIS family zinc finger 1* (*GLIS1*) locus is presented in Figure 2a. The most associated SNP with MVP is rs1879734 and locates in the first intron, as reported previously (OR=1.23, 95%CI=1.14–1.33, P-value =1.29×10^{-7}).[13] *GLIS1* encodes a GLI-related *Kruppel*-like zinc finger protein that functions as an activator and repressor of transcription. 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, a biological mechanism that we found to be enriched for MVP loci (Figure 1, Supplementary Tables 1 and 4). We also found that *GLIS1* significantly contributes to the enrichment of several tissues and cell types namely the connective tissue (*GLIS1* Z-score=2.3) and cells (*GLIS1* Z-score=2.8), and mesenchymal stem cells (Z-score=2.3) (Supplementary Table 3). We indicate in Supplementary Table 4 all 6 enriched gene sets at an FDR<0.05 where GLIS1 was found to be the best ranked gene by i-GSEA4GWAS analysis.

Follow-up of the association and functional annotation at the *GLIS1* locus

We followed-up the association between the *GLIS1* locus in a UK Biobank MVP case control dataset and obtained an overall genome-wide significant association with MVP (Lead SNP: rs18797434 $OR_{Overall}$=1.20, 95%CI=1.12–1.27, P =4.36×10^{-10}, Figure 3). To further characterize this association signal, we performed functional annotation for all 111 SNPs in moderate LD with rs18797434 (r^2>0.5) that showed nominal association in the meta-analysis of GWAS with MVP (Pvalue<0.05). All SNPs mapped to introns 1 and 2 in *GLIS1* (Supplementary Table 5, Figure 2b). We found that 47 SNPs locate in DNase Hypersensitive region (DNASEV) in diverse normal tissues, 16 locate in transcription factor binding sites (TFBS) and 55 are nominal (P<0.05) expression quantitative trait loci (eQTLs) for *GLIS1* in heart atrial appendage tissue from GTEx (N=264). We note that rs1879734 is located in a TFBS for endothelial transcription factor GATA2 and shows nominal eQTL association with GLIS1 expression in GTEx (p=0.048). The examination of histone marks (Figure 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 some SNPs in the locus. 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) (Figure 3b). *In silico* 4C experiment using four cell types including THP-1,[20] HUVEC,[21] H1-hESC[22] and NHEK[21] (Supplementary Figure 2) showed that the associated SNPs at this locus physically interact

only with *GLIS1* regulatory sequences suggesting it is likely to be the potential target and causal gene at this locus.

Glis1 Expression during heart development in mouse—GLIS1 belongs to GLI-similar 1–3 (GLIS1–3) subfamily of Krüppel-like zinc finger proteins that act either as activators or repressors of gene transcription and are involved in multiple physiological processes and diseases.[23] However, little is known about GLIS1 and its function in the cardiovascular system. In the absence of data about its role in valve biology, we performed immunohistochemistry experiments of the mouse ortholog, *Glis1* during embryonic, fetal and adult time points to study its pattern of expression during valve development. We found that Glis1 is expressed predominantly in nuclei of endothelial cells of the valves as well as the valvular interstitial cells during embryonic development (Figure 4). As the valves mature during fetal 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 analyze 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 (Supplementary Figure 3). Knockdown of *glis1b* at 72 hours post fertilization was robust but slightly less efficient (37.3% and 36.5% reduction Figure 5a, 5b) and had minimal effects on the overt morphology of the developing embryo (Figure 5c). Compared to controls, this knockdown resulted in a significant increase in the incidence of severe atrioventricular regurgitation with a combined fold increase of 1.6 (P=0.01). (Figure 5d, Supplementary Video).

Discussion

Our study defines globally molecular and cellular mechanisms underlying common genetic susceptibility to MVP and implicate established and unprecedented mechanisms. Through the genetic association and the study of the expression during valve development, we point at a potential mechanism of mitral valve degeneration involving *GLIS1*, a regulatory gene from the Krüppel-like zinc finger family of proteins.

Our gene-set enrichment and integrative analyses applied to MVP currently available meta-analysis of GWAS confirmed that genes in associated loci are predominantly involved in the regulation of cell adherence and migration, cytoskeleton biology, focal adhesion and the interaction with the extracellular matrix. An example of this enrichment is the EDG-1/SIPR1 pathway that we describe as highly enriched for MVP associated SNPs. SIPR1 is well-

known for its critical role in heart development and vascular maturation[24] as a key signaling pathway for actin assembly and lamellipodia formation. Interestingly, the reported interaction between SIPR1 and Filamin A[16] may encourage further functional investigation of the potential role of EDG1 in mitral valve degenerative process.

We also show that MVP genes are highly expressed in the cardiovascular system, connective and mesenchymal tissues. Several top contributing genes at MVP loci were 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 that we confirmed as a new genome-wide significant 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 from mouse somatic cells in the presence of OCT1/2, SOX2 and KLF4.[11] 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.[25] Interestingly, a common variant in *WNT8A*, member of the WNT genes family, is associated with atrial fibrillation[26] and was previously shown to be regulated by GLIS1/[11].

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 analyzed at risk loci and excluded more distant genes of interest. There is established evidence in favor of the functional role of distant long-range regulatory variants in predisposition to complex diseases.[27] 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.[13]

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 follow-up of *GLIS1*, a transcription factor with potential role in mechanisms linked for endothelial to mesenchymal transition cell migration during heart development and valve degenerative process.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments:

We acknowledge the contribution of the Leducq Foundation, Paris for supporting a transatlantic consortium investigating the physiopathology of mitral valve disease, for which the genome-wide association study was a major project (coordinators: R.A.L. and A.A.H.). We acknowledge investigators who contributed access to validation in cohorts: Yohan Bossé and Philippe Pibarot from *Institut universitaire de cardiologie et de pneumologie de Québec-Université Laval*, Quebec, Canada, Ramachandran S. Vasan and Emilia J. Benjamin from the Framingham Heart Study, USA and Ronen Durst, Hassadah Hebrew University Medical Center, Jerusalem, IL.

Sources of Funding: This study was supported by a PhD scholarship from the Chinese Scientific Council to MY, French Agency of Research (ANR-16-CE17-0015-02) to NB-N. AG and NB-N are supported by a European Research Council grant (ERC-Stg-ROSALIND-716628). 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). NIH K23 HL116652 (FND), contracts HHSN268201500001I, contract No1-HL25195, R01-HL-080124 and R01HL126136–01A1 (VSR), NIH R01 HL128099 (RAL), NIH HL141917 (RAL), Fondation Leducq Grant number 07CVD04 (RAL, AAH, SAS, XJ, JJS) and the Ellison Foundation, Boston, MA (RAL). The recruitment of the MVP France cohort was supported by the French Society of Cardiology (SFC).

References:

1. Delling FN, et al. Familial clustering of mitral valve prolapse in the community. Circulation. 2015;131:263–8. [PubMed: 25361552]
2. Freed LA, et al. Prevalence and clinical outcome of mitral-valve prolapse. N Engl J Med. 1999;341:1–7. [PubMed: 10387935]
3. Guy TS, Hill AC. Mitral valve prolapse. Annu Rev Med. 2012;63:277–292. [PubMed: 22248324]
4. Avierinos J, et al. Natural history of asymptomatic mitral valve prolapse in the community. Circulation. 2002;106:1355–1361. [PubMed: 12221052]
5. Gammie JS, et al. Isolated Mitral Valve Surgery: The Society of Thoracic Surgeons Adult Cardiac Surgery Database Analysis. Ann Thorac Surg. 2018;106:716–727. [PubMed: 30032907]
6. Katan O, et al. Incidence and Predictors of Infective Endocarditis in Mitral Valve Prolapse: A Population-Based Study. Mayo Clin Proc 2016;91:336–42. [PubMed: 26856780]
7. Freed LA, et al. Mitral valve prolapse in the general population: the benign nature of echocardiographic features in the Framingham Heart Study. J Am Coll Cardiol. 2002;40:1298–1304. [PubMed: 12383578]
8. Levine RA, et al. Mitral valve disease--morphology and mechanisms. Nat Rev Cardiol. 2015;12:689–710. [PubMed: 26483167]
9. Kyndt F, et al. Mutations in the gene encoding filamin A as a cause for familial cardiac valvular dystrophy. Circulation. 2007;115:40–49. [PubMed: 17190868]
10. Sauls K, et al. Developmental basis for filamin-A-associated myxomatous mitral valve disease. Cardiovasc Res. 2012;96:109–119. [PubMed: 22843703]
11. Maekawa M, et al. Direct reprogramming of somatic cells is promoted by maternal transcription factor Glis1. Nature. 2011;474:225–9. [PubMed: 21654807]
12. Durst R, et al. Mutations in DCHS1 cause mitral valve prolapse. Nature. 2015;525:109–113. [PubMed: 26258302]
13. Dina C, et al. Genetic association analyses highlight biological pathways underlying mitral valve prolapse. Nat Genet. 2015;47:1206–11. [PubMed: 26301497]
14. Sudlow C, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12:e1001779. [PubMed: 25826379]

15. Zhang K, et al. 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. 2010;38:W90–W95. [PubMed: 20435672]
16. Maceyka M, et al. Filamin A links sphingosine kinase 1 and sphingosine-1-phosphate receptor 1 at lamellipodia to orchestrate cell migration. Mol Cell Biol 2008;28:5687–97. [PubMed: 18644866]
17. Le Tourneau T, et al. New insights into mitral valve dystrophy: a Filamin-A genotype-phenotype and outcome study. Eur Heart J. 2018;39:1269–1277. [PubMed: 29020406]
18. Supek F, et al. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One. 2011;6:e21800. [PubMed: 21789182]
19. Pers TH, et al. Biological interpretation of genome-wide association studies using predicted gene functions. Nat Commun. 2015;6:5890. [PubMed: 25597830]
20. Phanstiel DH, et al. Static and Dynamic DNA Loops form AP-1-Bound Activation Hubs during Macrophage Development. Mol Cell. 2017;67:1037–1048 e6. [PubMed: 28890333]
21. Rao SSP, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell. 2014;159:1665–1680. [PubMed: 25497547]
22. Dixon JR, et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature. 2012;485:376–80. [PubMed: 22495300]
23. Jetten AM. GLIS1–3 transcription factors: critical roles in the regulation of multiple physiological processes and diseases. Cell Mol Life Sci. 2018;75:3473–3494. [PubMed: 29779043]
24. Cannavo A, et al. Sphingosine Kinases and Sphingosine 1-Phosphate Receptors: Signaling and Actions in the Cardiovascular System. Front Pharmacol. 2017;8:556. [PubMed: 28878674]
25. Scoville DW, et al. GLIS1–3: emerging roles in reprogramming, stem and progenitor cell differentiation and maintenance. Stem Cell Investig. 2017;4:80.
26. Ellinor PT, et al. Meta-analysis identifies six new susceptibility loci for atrial fibrillation. Nat Genet. 2012;44:670–5. [PubMed: 22544366]
27. Kleinjan DA, van Heyningen V. Long-range control of gene expression: emerging mechanisms and disruption in disease. Am J Hum Genet. 2005;76:8–32. [PubMed: 15549674]

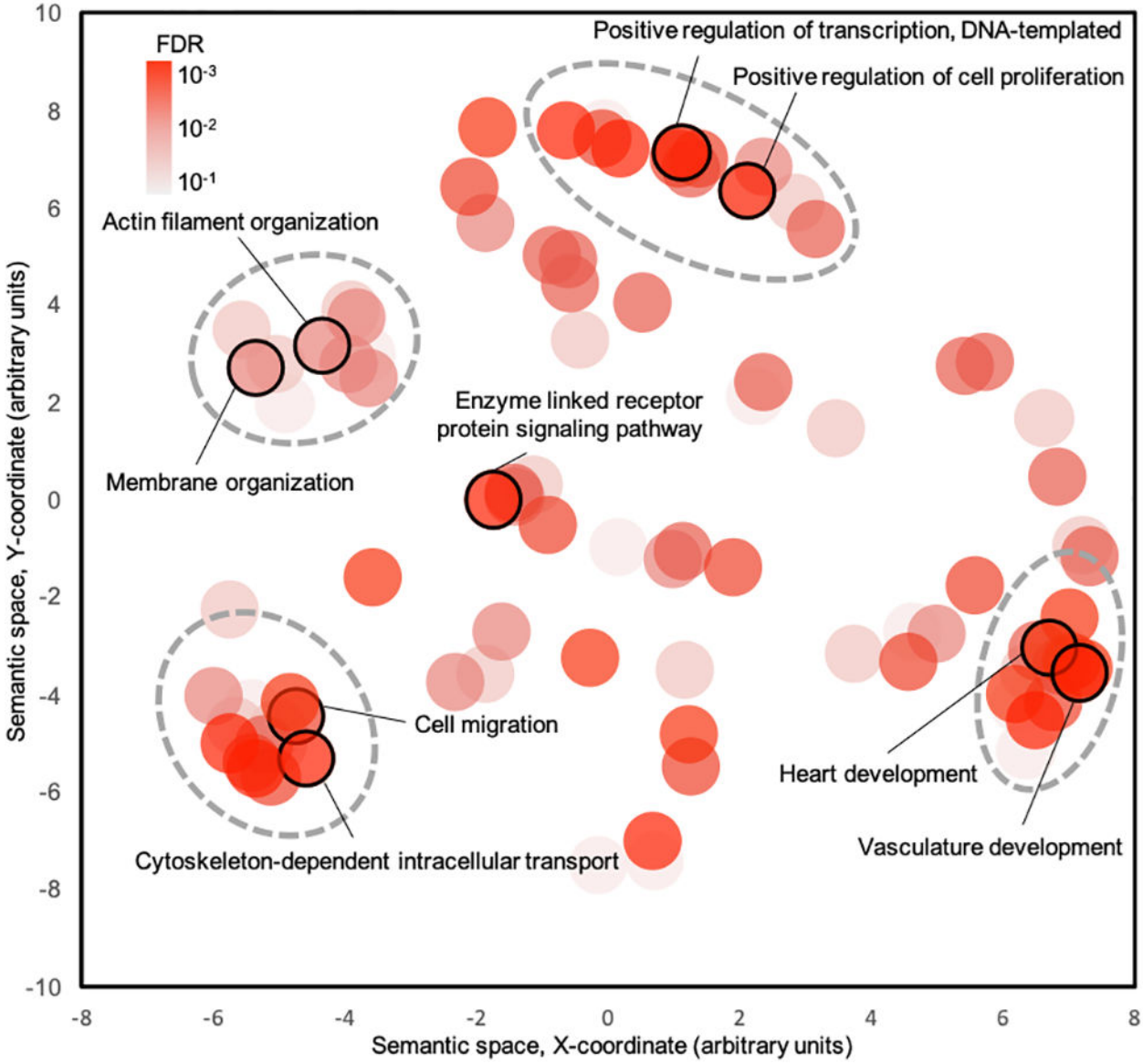

Figure 1: Semantic proximity between GO terms enriched for MVP associated risk loci.

Each GO biological process term identified as enriched in MVP-associated variant with FDR>0.05 was compared to other associated terms to measure their semantic proximity. REVIGO was used to group GO Biological Processes terms according to the semantic similarity. The matrix of semantic proximity was collapsed into two main component axis (X and Y) represented here. Each bubble represents one GO Biological Processes (GOBP) term, and the diameter of the bubble is proportional to the inverse logarithm of FDR. Main groups were delimited, and representative GOBP terms are indicated for each group. The full list of enriched terms is provided as Supplementary table

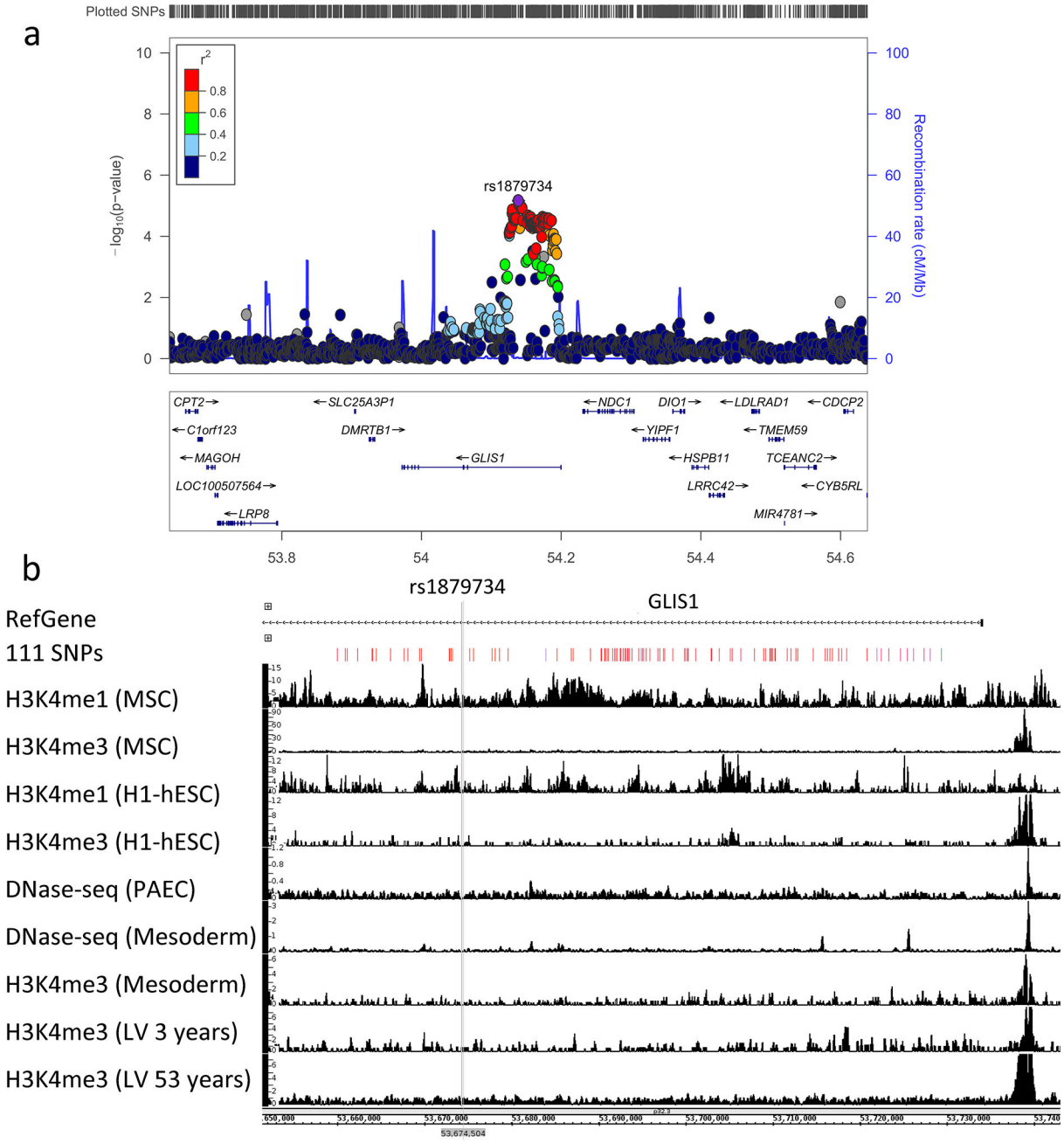

Figure 2. Genomic context and functional annotation of the association signal observed in the GWAS meta-analysis.

(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.

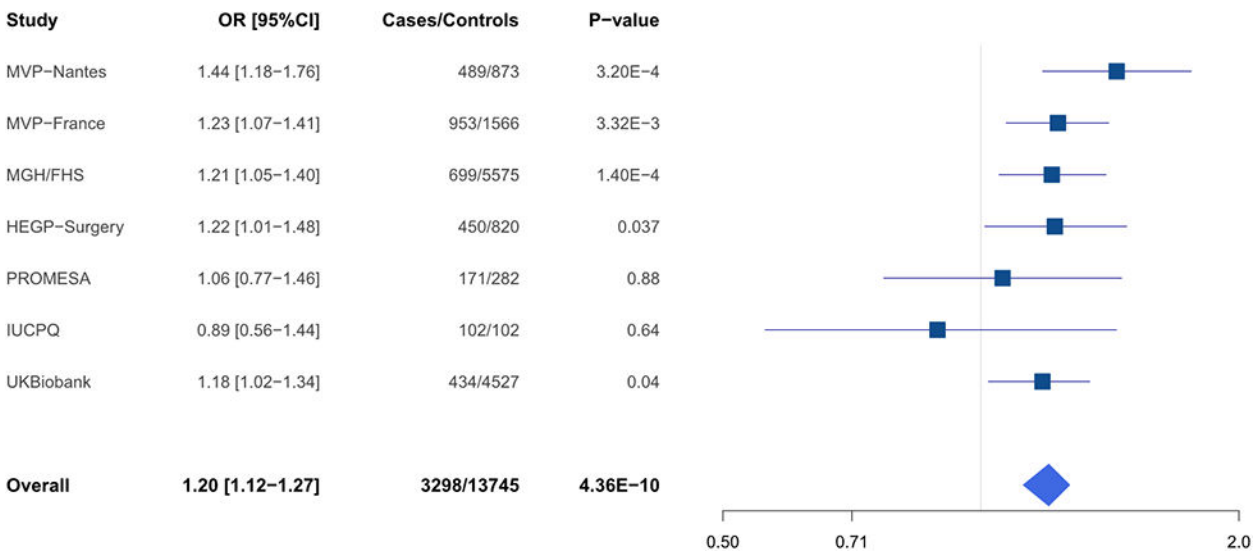

Figure 3. Forest plot shows associations in individual studies and the overall association between rs1879734 and mitral valve prolapse.

OR: odds ratio, CI: confidence interval. All cases control cohorts were previously described, except UK Biobank, which is novel and reported in details in methods.

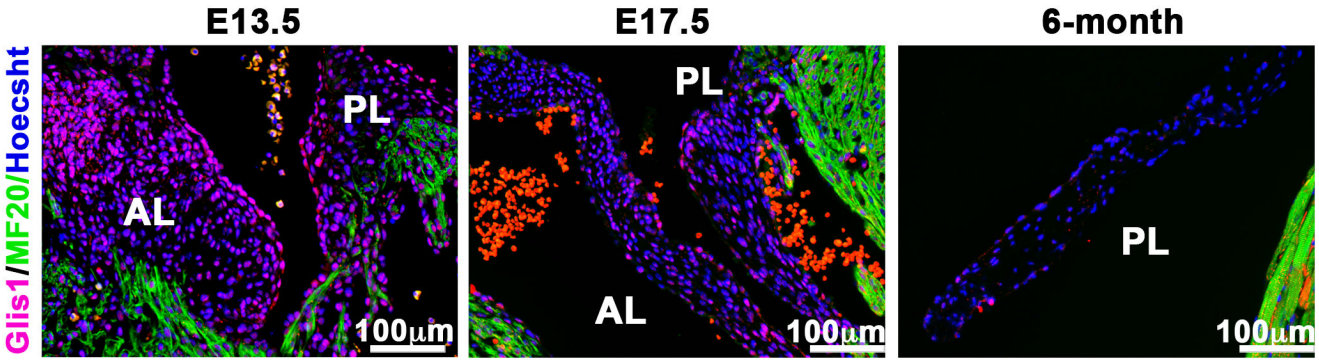

Figure 4. Glis1 expression in mouse developing and adult heart.

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 EndoMT 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.

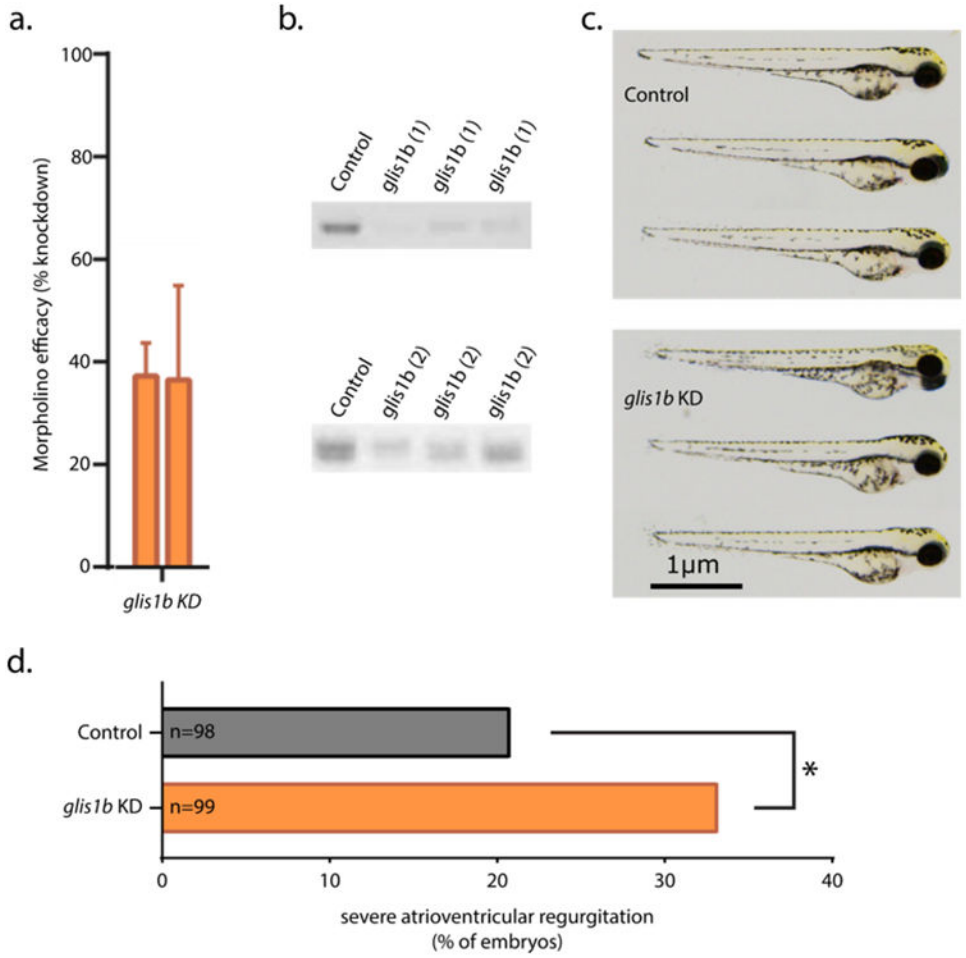

Figure 5. Assessment of cardiac regurgitation in zebrafish after morpholino knockdown of glis1b.

(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. The sample size for control and glis1b KD is 98 and 99, respectively. All results are relative to clutchmate controls.