Human rare variant and zebrafish CRISPR/Cas9-mediated mutant analyses reveal novel functions for *API5*, *HSPB7*, and *LMO2* in heart failure

Short title: *API5*, *HSPB7*, and *LMO2* in heart failure

Authors: Matthew J. Winter[1#*a], Yosuke Ono[2#], Jonathan S. Ball[1#], Anna Walentinsson[3], Erik Michaelsson[4], Anna Tochwin[1], Steffen Scholpp[2], Charles R. Tyler[1], Steve Rees[5], Malcolm J Hetheridge[1], Mohammad Bohlooly-Y[6*b]

Affiliations

[1]Biosciences, College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom.

[2]Living Systems Institute, College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom.

[3]Translational Science and Experimental Medicine, Research and Early Development, Cardiovascular, Renal and Metabolism, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden.

[4]Early Clinical Development, Research and Early Development, Cardiovascular, Renal and Metabolism, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

[5]Discovery Biology, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Cambridge, United Kingdom.

[6]Translational Genomics, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden.

[#]Contributed to this work equally

***Corresponding authors:**

[a]Matthew J. Winter (m.winter@exeter.ac.uk)

Biosciences, College of Life and Environmental Sciences, University of Exeter, Geoffrey Pope Building, Stocker Rd, Exeter, Devon, EX4 4QD, United Kingdom. Tel: +44 (0)1392 726258.

https://orcid.org/0000-0002-3668-6564

[b]Mohammad Bohlooly (Mohammad.Bohlooly@astrazeneca.com)

Translational Genomics, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, SE-431 83, Sweden. Tel: +46 730715171.

Keywords: CRISPR/Cas9; zebrafish; heart failure; drug target validation; cardiovascular disease

Summary Statement: Human rare variant analysis and CRISPR/Cas9-generated mutant phenotyping in zebrafish identifies genes as potential drug targets in cardiovascular disease.

Abstract

The clinical heterogeneity of heart failure has challenged our understanding of the underlying genetic mechanisms of this disease. To gain further insights into this complex pathophysiology we combined human rare variant analysis and *in vivo* CRISPR/Cas9-mediated mutant phenotyping in zebrafish to identify and investigate the role of 3 genes.

Whole-exome sequencing of patients identified *API5*, *HSPB7*, and *LMO2* as causally associated with heart failure and these genes were further investigated, alongside the positive control *gata5*, using CRISPR/Cas9-mediated multi-locus *in vivo* mutation in zebrafish. Following effective somatic mutation, we observed multiple impacts on cardiovascular development and function in F0 embryos including reductions in ventricle size, pericardial oedema, and chamber malformation. In the case of *lmo2*, there was also a significant impact on heart function.

Our analysis suggests novel functions for *API5*, *HSPB7*, and *LMO2* in human cardiovascular disease and identifies them as potential drug targets. Our data also supports *in vivo* CRISPR/Cas9-mediated multi-locus gene mutation analysis in F0 zebrafish as a rapid and effective primary screen for assessing gene function, as part of an integrated multi-level drug target validation strategy.

Introduction

At least half of all heart failure cases fall outside of the definition of heart failure associated with reduced ejection fraction (or HFrEF), and although the aetiology of these patients is largely unknown, it has been proposed that the root cause is extracardiac (Senni et al., 2014). To gain further insights into potential contributing pathophysiologies, we used a novel approach combining *in silico* assessment of human gene variant data and *in vivo* CRISPR/Cas9-mediated gene mutation analysis in zebrafish. This approach was adopted as functional analysis of identified causal genetic variants using traditional target validation strategies is impractical when multiple gene candidates require assessment. Specifically, *in vitro* cell based approaches preclude organ system-dependent functional analysis, and the use of *in vivo* assessment in mouse is prohibitively slow and expensive. The embryo-larval zebrafish combines genetic tractability, higher throughout amenability, and optical transparency allowing the relatively simple assessment of organ system morphology and functionality across multiple gene targets (Gut et al., 2017). Indeed, the zebrafish is widely considered an appropriate surrogate for studying mammalian cardiovascular biology (MacRae and Peterson, 2015) and recent studies have demonstrated its utility in CRISPR/Cas9 mediated F0 screens as a rapid, highly reproducible and scalable knock-out model (Kroll et al., 2021).

Whole exome sequencing (WES) and subsequent bioinformatic analysis was applied to identify genes from clinical cohorts that were considered to be causally related to chronic heart failure (Povysil et al., 2020). A subset of these genes (*API5*, *HSPB7*, and *LMO2*) were predicted to play a broader role in heart failure aetiology and these were selected for *in vivo* phenotypic assessment in zebrafish. The positive control gene selected was *GATA5*, which has a critical role in heart development and has been implicated in multiple human cardiovascular disease aetiologies (Zhang et al., 2015, Wei et al., 2013, Gu et al., 2012). Functional knockout of *gata5* resulted in F0 zebrafish larvae exhibiting the expected heart phenotype, and mutation of the 3 target genes resulted in a negative impact on the physiology and/or development of the zebrafish cardiovascular system. The evidence presented supports our approach as a powerful primary screen for assessing gene function during target

validation, and the data generated provides *in vivo* evidence to support the investigation of these genes as novel drug targets for treating human cardiovascular disease.

Results

In silico analysis

Key results of the *in silico* analysis are summarised in **Table 1** with snapshots of the findings shown in **Fig. 1**. Additional data are contained within **Supplementary figures S1-S2 and supplementary data S1-S4**, and literature-derived information summarised in **Supplementary Table S1**.

As expected, heart *GATA5* mRNA expression (**Fig. 1B)** was significantly altered in association with various cardiovascular disease states, the most significant being downregulated expression in dilated cardiomyopathy. The most pronounced changes in cardiac and blood *API5* expression included downregulated expression in myocardial infarction, and upregulated heart *API5* expression in ischemic and non-ischemic cardiomyopathies. Heart *HSPB7* expression was largely downregulated in various cardiomyopathies, heart failure and atrial fibrillation, although upregulation was reported in a number of published dilated cardiomyopathy studies. *LMO2* blood mRNA expression predominantly showed upregulation, including in association with ischemic cardiomyopathy, coronary artery disease and myocardial infarction, whereas cardiac *LMO2* mRNA expression showed both increases and decreases across the same range of conditions.

Analysis of common variant gene-level associations (**Fig. 1D**) revealed that *GATA5* was significantly associated with various lung function-and haematological traits. No significant common variant associations to heart disease, however, were found, although rare loss-of-function mutations in *GATA5* have been reported to cause congenital heart defects (Jiang et al., 2013). Assessment of *API5* revealed a significant association with body mass index (BMI) and a borderline significant association with neutrophil count (p-value=$9x10^{-7}$). Among the cardiovascular traits assessed, low frequency 3'UTR or intron *API5* genetic variants showed a significant association with "Cause of death: atrial

fibrillation and flutter" ($p=6.4^{-22}$), and "Cause of death: acute and subacute infective endocarditis" ($p=6.7^{-13}$). *HSPB7* common variant data showed a significant association with cardiovascular-, renal-, haematological- and musculoskeletal traits with the most significant association to cardiac function including left ventricular ejection fraction, and left ventricular end systolic volume. Other GWAS data revealed significant associations between common variants in *HSPB7* intron 3' and 5'UTR and idiopathic dilated and sporadic cardiomyopathy ($p=5.3^{-13}$ and $p=1.4^{-9}$ respectively), as well as with systolic blood pressure ($p=7^{-12}$). Finally, analysis of the *LMO2* data revealed significant associations with various haematological traits including mean corpuscular volume, haemoglobin and red blood cell count, although the most significant cardiovascular trait association was only suggestive (P-wave duration at $p=2.1^{-5)}$. Further analysis did, however, reveal that *LMO2* intron variants were significantly associated with "Cause of death: cardiomegaly" ($p=3.0^{-9}$) and "Cause of death: dilated cardiomyopathy" ($p=1.3^{-8}$).

In vivo target gene mutation efficiency

There was clear cleavage of the PCR product and high efficiency mutagenesis observed across all guide designs (100% in the presence of Cas9 according to the T7E1 assay), and sequencing confirmed that effective site-specific gene mutation was achieved for all combinations of gRNAs, but not in the case of the Cas9 controls (**Fig. 2**).

Morphology and function of mutants

At 2 days post fertilization (dpf), the most prominent *gata5* mutant phenotypes occurred after injection of the combined gRNAs (g#1,2,3), followed by injection of g#1 alone (**Fig. S3**) and by 4dpf, all *gata5* mutant animals showed pericardial oedema, high incidences of misshapen and small heart chambers and a frequent lack of chamber definition amongst other non-cardiac morphological abnormalities (**Fig. 3**). Histology (**Fig. 4 and Fig. S4**) revealed that *gata5* mutants generally exhibited cardiac hypoplasia, chamber malformation and pericardial distension, and there was also a reduction in end systolic and end diastolic ventricle length in the g#1,2,3-injected animals (**Fig. 4**). Furthermore, *cmlc2::DsRed2-nuc gata5* mutants (**Fig. 4 and Fig. S5**) exhibited weaker fluorescence and reduced

numbers of cardiomyocytes. For all endpoints measured, with the exception of vessel diameter (only measurable where flow occurred), *gata5* mutants showed reduced cardiovascular function compared with the Cas9 controls (**Fig. 5**, with accompanying videos in **Supplementary Videos**). Collectively these data supported a negative impact of *gata5* mutation on 4dpf zebrafish cardiovascular physiology, supporting the validity of our F0 zebrafish screening approach.

At 2dpf the strongest *api5* mutant phenotypes were observed after injection of the combined gRNAs (g#1,2,3), followed by g#1 alone (**Fig. S3**), although by 4dpf there was little indication of any gross morphological effects (**Fig. 3**). Histology (**Fig. 4 and Fig. S4**) suggested that in some cases there was enlargement of the heart chambers with myocardial wall thinning, although this was not observed consistently. Chamber enlargement was, however supported by a small increase in ventricle diameter in the *api5* mutants (**Fig. 4**), and confocal assessment of *cmlc2::DsRed2-nuc* animals (**Fig. 4 and Fig. S5**) revealed that *api5* mutants occasionally exhibited mild pericardial oedema, and in some cases slightly enlarged ventricles. Supporting the mild impact of *api5* mutation, there were no significant effects on cardiovascular function detected (**Fig. 5 and Supplementary Videos**).

The most prominent *hspb7* mutant phenotypes at 2dpf occurred after injection of the combined gRNAs (g#1,2,3), followed by g#3 alone (**Fig. S3**), and at 4dpf there were widespread developmental abnormalities across multiple tissues compared with the Cas9 controls (**Fig. 3**). In particular there was a high incidence of pericardial oedema, and in 20% of cases in the g#1,2,3 group, misshapen, poorly-defined and small heart chambers. Histological (**Fig. 4 and Fig. S4**) and confocal microscopic analysis (**Fig. 4 and Fig. S5**), however, did not reveal any ultrastructural abnormalities, although a significant reduction in ventricle diameter was detected (**Fig. 4**). Despite the high incidence of pericardial oedema observed in the *hspb7* mutants, the impact on cardiovascular function in these animals was mild (**Fig. 5 and Supplementary Videos**), with a small reduction in surrogate stroke volume in the g#1,2,3-injected animals detected suggesting (along with blood pooling at 2dpf) a small reduction in pumping efficiency.

Assessment of *lmo2* mutants at 2dpf revealed the most prominent phenotypes after injection of the combined gRNAs (g#1,2,3), followed by g#2 alone (**Fig. S3**). The high mortality exhibited in the g#1,2,3-injected animals (76% in run 1), however, supported the use of g#2 for the second confirmatory run. At 4dpf, there was a 100% incidence of heart, craniofacial, neural tube, jaw, swim bladder and yolk ball abnormalities and a high incidence of other non-cardiovascular abnormalities across all *lmo2* treatments (**Fig. 3**). Histological analysis supported an extreme cardiac phenotype with frequent observation of cardiac hypoplasia, chamber malformation and pericardial distension in the mutant animals (**Fig. 4 and Fig. S4**), along with altered cardiomyocyte shape in the g#1,2,3- and pyknotic nuclear structure in g#2-injected larvae. This impact of *lmo2* mutation was also reflected in significantly smaller ventricle sizes in mutant animals (**Fig.4**) and confocal assessment of *cmlc2::DsRed2-nuc* animals (**Fig. 4 and Fig. S5**) revealed disorganisation of cardiomyocytes in the myocardium, and smaller heart chambers. The cardiovascular structural impact in *lmo2* mutants was reflected in the functional assessment (**Fig. 5 and Supplementary Videos)** with both mutant groups showing heavily reduced cardiovascular functionality compared with the Cas9 injection controls. It was also notable that the *lmo2* mutants lacked visible circulating erythrocytes, which is consistent with the key role of this gene in haematopoiesis.

Discussion

Using human WES in conjunction with bioinformatics, we identified three genes with a plausible link to human cardiovascular disease. We then employed CRISPR/Cas9-mediated *in vivo* mutation and morphological and functional phenotyping in F0 zebrafish to demonstrate a role for these genes in the development, and pathophysiology, of the vertebrate cardiovascular system. Collectively, this approach affords an extremely powerful tool for the rapid provision of *in vivo* gene function data during early-stage target validation.

In vivo mutation of the positive control gene, *gata5*, resulted in an expected impact on cardiovascular development and function in 4dpf zebrafish. Our data are consistent with the known link between

GATA5 variants and multiple human cardiovascular pathologies including familial dilated cardiomyopathy (Zhang et al., 2015) and congenital ventricular-septal defects (Wei et al., 2013). These data also add to previous work in zebrafish including demonstration of prominent defects in myocardial differentiation and the formation of ectopic beating myocardial tissue after loss and gain of function, respectively (Reiter et al., 1999). Here, the impact of *gata5* mutation was also evident beyond the cardiovascular system and this extracardiac impact is supported by the spatiotemporal expression of g*ata5* in developing mouse (Chen et al., 2009), and by previous work in zebrafish demonstrating the central role of *gata5* in endodermal morphogenesis more broadly (Reiter et al., 2001).

Our *in silico* analysis suggested an association between *API5*, which encodes human apoptosis inhibitor-5 protein (Bong et al., 2020), and some human cardiovascular pathologies. Published evidence for a role for *API5* in human cardiovascular disease is limited, however, *API5*-signalling has been implicated in the altered vascular endothelial cell apoptosis involved in certain cardiovascular disease aetiologies (Lu et al., 2016, Mao et al., 2020). We found *in vivo* mutation of *api5* resulted in mild pericardial oedema and some evidence of a small increase in ventricle size. From the histopathological assessment it was not possible to determine whether cardiomyocyte hypertrophy or hyperplasia was occurring, however, the latter is conceivable given the role of *api5* as a putative anti-apoptotic gene regulator in zebrafish. Although published data on *api5* function in zebrafish are limited, *api5* has been shown to be modestly upregulated in the hearts of adult zebrafish following hypoxic insult (Marques et al., 2008) supporting a cardio-protective role against tissue injury. Along with our data this suggests that *api5* does not play a critical role in early cardiovascular development, and may be more associated with organ-system protection under conditions of physiological stress, or as a consequence of tissue injury.

*HSPB*7 encodes small heat shock protein 7, and although highly expressed in the developing and adult mammalian heart, its cardiac function remains obscure (Mercer et al., 2018). *HSPB*7 gene variants have been implicated in a range of human cardiovascular diseases including heart failure (Cappola et

al., 2010, Aung et al., 2019) and dilated cardiomyopathy (Villard et al., 2011, Esslinger et al., 2017). Our *in silico* analysis supported an association with various human cardiovascular pathologies and a central involvement in cardiovascular disease is also supported by knockout studies in mouse (Wu et al., 2017, Liao et al., 2017). *In vivo hspb7* mutation resulted in widespread impacts on cardiac morphology although perhaps surprisingly only a mild impact on cardiac pumping efficiency. Our CRISPR-mediated mutant data agree with a previous zebrafish Morpholino-based knockdown study (Rosenfeld et al., 2013) in which a key role for *hsbp7* in cardiac development was demonstrated. In addition, the mild impact on cardiac function may suggest that, as is the case in mouse, impaired functionality may be more severe in older animals, or given its proposed role as a cardio-protective chaperone (Ke et al., 2011), under conditions of organ stress. The latter is supported by previous work using TALENs-mediated *hspb7* adult zebrafish knockouts which, despite showing normal heart development, were more susceptible to exercise-induced stress (Mercer et al., 2018).

*LMO*2 is highly conserved amongst vertebrate lineages, and encodes the Lim-domain only 2 nuclear transcriptional co-regulator crucial in early embryonic erythropoiesis and angiogenic remodelling (Chambers and Rabbitts, 2015). Beyond highlighting its well-established role in haematopoiesis, our *in silico* data supported an association between *LMO2* variants and cardiomegaly and cardiomyopathy, although comparative expression data were less conclusive. Published data on cardiac-specific functions for *LMO2* are limited, however, overexpression of *lmo2* (with *scl/Tal1* with which a transcriptional complex is formed) in zebrafish has been shown to result in enhanced ectopic blood and endothelial development at the expense of the somitic, pronephric and cardiogenic mesoderm (Gering et al., 2003). This apparent wider influence on mesodermal development was supported here by widespread and severe embryonic malformations in mutants across multiple tissues including the heart. The specific impact on cardiovascular development was also reflected in the significant impairment of cardiovascular function in the *lmo2* mutant embryos supporting a crucial role for this gene in early organ-system development and functionality.

In conclusion, our approach of combining rare gene variant analysis following whole exome sequencing from patients with heart failure, followed by rapid *in vivo* CRISPR/Cas9-mediated mutation in zebrafish, has provided new target validation data on *API5, HSPB7* and *LMO2*. Mutation of all three genes resulted, to a greater or lesser degree, in measurable impacts on the development and/or function of the zebrafish cardiovascular system, warranting further investigation of these genes as potential targets for the development of drugs to treat human cardiovascular disease.

Materials and Methods

Case-control collapsing analysis

Initially target genes were identified by WES of heart failure patients from two clinical trials; Candesartan in Heart Failure-Assessment of Reduction in Mortality and Morbidity (CHARM) (Pfeffer et al., 2003) and Controlled Rosuvastatin Multinational Trial in Heart Failure (CORONA) (Kjekshus et al., 2007). 5942 heart failure cases from these trials were compared to controls without reported heart disease using gene-based rare-variant collapsing analysis (Povysil et al., 2020). One gene, TTN (encoding Titin), reached study-wide significance, with the strongest association in the dominant protein-truncating variant (PTV) model ($p = 3.35x10^{-13}$), a finding that was replicated in the UK Biobank WES data (Povysil et al., 2020) and validated our *in silico* approach. An additional 255 genes that were considered suggestively causal, with a significance of $p<1x10^{-4}$, were prioritised for bioinformatics assessment as described below.

Bioinformatics analysis

Firstly, each gene was assessed for genetic association to human disease phenotypes based on large-scale genome-wide association (GWAS) and WES studies encompassing common to low frequency variants (Common Metabolic Diseases Knowledge Portal or CMDKP, GWAS Catalog, Phenoscanner). Rare variant associations reported in Online Mendelian Inheritance in Man (OMIM) and ClinVar were also captured. Baseline tissue and cellular expression of candidate targets were investigated based on bulk and single cell RNA sequencing data from human tissues. Studies of

expression dysregulation in cardiovascular disease were also conducted using patient transcriptomics data deposited in NCBI Gene Expression Omnibus (GEO) using QIAGEN's OmicSoft DiseaseLand (release HumanDisease_B37_20191215 _v14a), which applies generalized linear models on log2 transformed intensities (microarray data), and DESeq2 for raw counts data (RNAseq data). Genes were considered significantly differentially expressed at an adjusted $p<0.05$. For mechanistic inference assessment, network-based functional enrichment analysis was performed using three separate tools (STRING, Harmonizome and GeneMANIA) all relying on multiple data types including protein-protein interactions, co-expression, database and text mining. Further details on data types and sources used are summarised in **Table 1**. Collectively, from these *in silico* analyses *API5*, *HSPB7* and *LMO2* were selected for *in vivo* assessment.

Guide RNA design and preparation

For the zebrafish orthologues of each gene (*gata5*~ENSDARG00000017821; *api5*~ENSDARG00000033597; *hspb7*~ENSDARG00000104441; *lmo2*~ENSDARG00000095019), three individual guide RNAs (gRNAs) were designed (https://chopchop.cbu.uib.no) to target discrete sections of coding exon 1, except for *lmo*2 which used exon 2 (**Fig. S6**). Each gRNA was applied alone (termed g#1, g#2 or g#3) to assess the consistency of phenotypes across different target sites, and as a combined injection containing all three gRNAs (termed g#1,2,3) to ensure functionally-effective mutation, alongside Cas9-only injection controls. Gene-specific crRNA and tracrRNA (Integrated DNA Technologies Inc. Coralville, USA) were diluted to a final concentration of 12µM in nuclease-free duplex buffer and the resultant gRNA mixture incubated at 95°C for 5 min. Immediately prior to use, 5µl of gRNA mixture was mixed with Cas9-NLS protein (final concentration of 5µM. New England Biolabs, Ipswich, USA), 2M KCl (final concentration of 300mM), and 0.5% v/v Phenol red solution (Sigma Aldrich UK Ltd. Poole, UK), and the resultant mixture incubated at 37 °C for 10 minutes to assemble the gRNA/Cas9 ribonucleoprotein complex. The resultant mixture was held at room temperature until use.

Zebrafish culture and microinjection

Adult WIK (Wild-type India Kolkata) strain zebrafish (*Danio rerio*), originally obtained from the Zebrafish International Resource Center (ZIRC, University of Oregon, Eugene, USA), were held under optimal spawning conditions (14 hour light: 10 hour dark cycle, with 20 minute dusk-dawn transition periods, 28 ± 1°C), in groups of males and females in flow through aquaria in the Aquatic Resources Centre, at the University of Exeter. This line was chosen as previous work suggested good reproduction in the laboratory with few spontaneous developmental abnormalities. Embryos were collected from individual male-female pairs, and used for injection at the one-cell stage (**see below**). In addition, *cmlc2::DsRed2-nuc* transgenic zebrafish (Mably et al., 2003) were also used for confocal microscopic analysis of the impact of gene mutation on cardiomyocyte development and morphology (**see below**). These were subject to the same methods as used for the WIK embryos.

Microinjection needles were prepared from thin wall borosilicate glass capillaries with filament (Outer diameter 1.0mm, inner diameter 0.75mm. World Precision Instruments, Sarasota, USA) on a micropipette puller (P-1000,Sutter instruments, Novato, USA) using the following settings: Heat 501, Pull 60, Velocity 60, Time 20, Pressure 300, Ramp 499.

From pairs of spawning zebrafish, eggs were assessed for the desired development stage (1-cell) and for general condition before being transferred in batches of 50-60 into the furrows of an injection mould-imprinted agar plate. Next, the injection needle was loaded with the injection mixture calibrated using a microscale graticule to deliver 0.5-1.7nl per injection. Each egg was injected once, close to the cell/yolk boundary layer (FemtoJet 4x, Eppendorf, Hamburg, Germany), with successful injection indicated by the presence of phenol red. Injected eggs were then transferred to a Petridish containing culture water and methylene blue (2 drops per litre of water) and cultured on a black background, under a 14-hour light: 10 hour dark cycle with 20 minute dusk-dawn transition periods, at 28 ± 1°C in Petri dishes for the first 24 hours. At the end of day 0, all unfertilised and dead embryos were removed. Next, 50-60 viable embryos per treatment, selected at random, were individually transferred to wells of 48-well microplates (each in 1ml) for later assessment.

Morphological assessment at 2 and 4dpf

At 2dpf, all embryos across all the 3 individual and 1 combined guide RNA (gRNA) injected groups (named g#1, g#2, g#3 and g#1,2,3 respectively) were assessed morphologically according to the following procedure. The main aim of morphological assessment at 2dpf was to identify the most effective guides (i.e. those resulting in the most prominent phenotype versus the Cas9-only injected controls) for full morphological and functional assessment of the resultant phenotype at 4dpf. In addition, 8 embryos were removed from each treatment and individually placed into PCR tubes for analysis of gene mutation efficiency (**see below**). 2dpf embryos were scored (without anaesthesia) using a dissecting microscope against a list of criteria shown in **supplementary Table S2**.

At 3dpf, if necessary, embryos were manually dechorionated using fine forceps allowing the spine to straighten to facilitate full morphological analysis at 4dpf. At 4dpf, 10 animals were selected at random from the 2 most effective treatments groups, alongside 10 embryos from the Cas9-only injected animal group for full morphological scoring according to the basic method of Gustafson *et al*., (Gustafson et al., 2012) and Ball *et al*. (Ball et al., 2014). To facilitate scoring, animals were lightly anaesthetised by immersion in 0.165g/L tricaine methanesulfonate (pH 7.5) and scored according to the criteria shown in **supplementary Table S3**. Images were taken from representative animals within each treatment group. After scoring, each animal was directly transferred to a solution of benzocaine (1g/L in 1% ethanol) for euthanasia without recovery, and then fixed for later histological analysis (**see below**). The guide-injected group providing the strongest phenotype was also selected for a second run to confirm the observed effect in a separate batch of embryos.

Analysis of mutation efficiency

Genomic DNA was extracted from individual 2dpf larvae using the HotSHOT method (Meeker et al., 2007). Briefly, all water was removed from each PCR tube containing an embryo, 50µl of 50 mM NaOH added, and the sample heated for 10mins at 95°C. The samples were then vigorously vortexed and subsequently cooled on ice. Next 5µl of 1M Tris-HCl (pH 8.0) was added, the samples well

mixed, centrifuged and the supernatant containing the genomic DNA removed and stored at -20 °C until further processing. The PCR primers were designed and obtained from Eurofins Genomics (Ebersberg, Germany). The primers used were: *gata5* (forward: GGAAACCATCGCATTTGGAG and reverse: AGGGCACTTCCATATTGATC); *api5* (forward: ATACAGCGGAAGTATCCGAC and reverse: TCAATTCTCGCTCAGGCTTG; *hspb7* (forward: GAATAAGAACTTGATCACCGG and reverse: GCATATAGCTTTCCACTCAC); and *lmo2* (Forward: TGGATGAGGTGCTCCAGATG and reverse: ATCTCTCCTGCACAGCTTTC). The PCR mixture was prepared as follows: 10µl of 2x PCRBIO Taq Mix Red (PCR Biosystems Ltd., London, UK); 0.8µl each of 10µM forward and reverse primers; 1µl of genomic DNA; and 7.4µl of water. The PCR machine settings were as follows: 1min at 95°C; 30 cycles of 15sec at 95°C, 15sec at 58°C, 15sec at 72°C; and finally 1min at 72°C. DNA amplification was checked on a 3% agar gel. T7endonuclease I assays were undertaken to detect heteroduplexes in the genomic DNA. For this, PCR products were denatured at 95°C for 5min and then the T7E1 reaction mixture made as follows: 10µl of each PCR product; 1.5µl NEBuffer 2 (New England Biolabs, Ipswich, USA); 0.3µl T7E1 enzyme at 10 Units/µl (New England Biolabs, Ipswich, USA); and 3.2µl water. Next digestion was undertaken for 1min at 37°C and the resultant products assessed on a 3% agar gel. A sub sample of amplified DNA from each PCR product was also sent for Sanger sequencing by Eurofins Genomics (Ebersberg, Germany) using the abovementioned forward or reverse primers.

Histology

After morphological analysis, 4dpf animals were transferred into fixation tubes in which all water was replaced with 10% neutral buffered formalin for 4 hours, followed by 70% ethanol/methanol in which they were stored at 4°C until further processing. Animals were then transferred into agar moulds for orientation (Copper et al., 2018, Sabaliauskas et al., 2006), and subsequently into tissue cassettes and embedded in paraffin using an automatic tissue processor (Thermo Scientific Excelsior AS, Thermo Fisher Scientific Ltd., Waltham, USA). The sequence of fixation and embedding steps applied are summarised in **supplementary Table S4**. Following fixation, 5µm sections were cut from each paraffin block using a microtome (Shandon AS325, ThermoFisher Scientific, Waltham, USA), and

the resultant sections haemoxylin and eosin (H&E) stained on an automatic stainer (Shandon Varistain 24-4, ThermoFisher Scientific, Waltham, USA) using the sequence summarised in **supplementary Table S5**. After staining, images of each section were captured on a binocular microscope (Axioskop 40, Zeiss, Oberkochen, Germany) equipped with a colour digital camera (DP70, Olympus, Tokyo, Japan) to allow histopathological analysis.

Cardiovascular functional assessment at 4dpf

Cardiovascular function was assessed in embryos at 4dpf using the method of Parker *et al*. (Parker et al., 2014). For this, 10 animals were selected at random from the same treatment groups selected for full morphological analysis at 4dpf. Briefly, each animal was lightly anaesthetised by immersion in tricaine methanesulfonate (0.1g/L pH 7.5) and then transferred into low melting point agarose (1g/100ml of the same tricaine methanesulfonate solution to maintain anaesthesia during imaging) and then deposited on its side on a clear microscope slide. Imaging was undertaken on an inverted light microscope (Leica DM IRB, Leica Microsystems UK Ltd., Milton Keynes, UK) equipped with two video cameras, at 10x magnification. One camera recorded the heart at 25 frames per second (fps. Grasshopper® GRAS-50S5C-C, Point Grey, Richmond, Canada) and the second was positioned to capture the dorsal aorta at 120fps (Grasshopper® GRAS-03K2M-C, Point Grey, Richmond, Canada). Recording was then undertaken for 10 minutes following which animals were directly transferred to a solution of benzocaine (1g/L in 1% ethanol) for euthanasia without recovery.

Heart videos were subsequently analysed using MicroZebraLab™ (v3.5, ViewPoint, Lyon, France) as detailed previously (Parker et al., 2014). From this, beat frequencies were provided for each of the atrium (atrial beat rate or ABR) and the ventricle (ventricular beat rate or VBR) independently to allow a global heart rate output and also the detection of certain arrhythmias, for example the decoupling of atrial and ventricular beat frequencies which has been previously described in association with exposure to some QT-prolonging drugs. Blood flow videos were analysed using ZebraBlood™ (v1.3.2, ViewPoint, Lyon, France), also as previously detailed (Parker et al., 2014). This provided measures of blood flow (nl/sec), blood linear velocity and vessel diameter. In addition

to the directly determined parameters, estimates of stroke volume and cardiac output were calculated using measurements of heart rate and blood flow, and these were termed surrogate stroke volume (SSV) and surrogate cardiac output (SCO). Normally stroke volume is precisely calculated using the difference between end systolic and end diastolic volumes. However, using our system a surrogate measure of SSV was calculated by dividing the dorsal aorta flow rate (in nL/second), by the VBR per second (bpm/60). Similarly, cardiac output is normally calculated by dividing the stroke volume by the heart rate to provide a volume of blood pumped per minute by a surrogate measure SCO, was calculated here by multiplying the SSV by the VBR in bpm.

Ventricle diameter measurements

Using the videos of the heart captured for functional analyses, five animals were randomly selected per treatment from which a manual measurement of ventricle diameters was undertaken. Using VirtualDub (http://www.virtualdub.org/) to pause videos, 10 time points were selected at random across the full video length, from each animal. At each of these 10 points, the minimum (systolic) and maximum (diastolic) ventricle chamber length was measured along the longest axis to respectively represent the end systolic and end diastolic ventricle size. As videos were used for analysis, precise determination of chamber edges was difficult and as such the actual dimensions should be considered to be approximations. In addition, it was not possible to normalise these measurements to the length of the animal used for functional assessment, however ventricle measurements have been considered within the context of the average total body length of embryos within the same batch of animals recorded during the 4dpf morphological assessment.

Confocal analysis of cardiomyocyte development and morphology

At 4dpf, *cmlc2::*DsRed2-nuc transgenic zebrafish embryos were analysed using confocal microscopy. For each gene and for Cas9-only injected animals, 10 randomly selected embryos were assessed on a Nikon A1R laser scanning confocal microscope (Nikon, Tokyo, Japan) using 568nm laser excitation (power 90, PMT 85) and transmitted light (PMT 25). At 20x magnification, Z-slices were taken every 5μm through the heart from which maximum intensity z-projections were then generated. For

imaging, the embryo's hearts were stopped by immersion in 1g/L tricaine methanesulfonate (pH 7.5) after which they were transferred to 1g/100ml low melting point agarose made using the same tricaine methanesulfonate solution for immobilization during imaging. As before, after imaging larvae were directly transferred to a solution of benzocaine (1g/L in 1% ethanol) for euthanasia without recovery.

Statistical analysis

All measured functional parameters were averaged across the assessment period to provide a series of individual values per animal. Each group was first tested for normality (Anderson-Darling Test) and homogeneity of variance (Levene's, Bartlett's, or F-test). Each treatment was then compared with the Cas9-injected control group using either the Student's T-tests or 1-way ANOVA and Tukey's HSD tests (parametric), or the Mann Whitney U-tests or Kruskal Wallis and Dunn's Tests (non-parametric). All analyses were undertaken using Minitab™. Throughout data are shown as the mean, ± standard error of the mean (n), with a minimum α level of 0.05 applied (with a Bonferroni correction for multiple comparisons).

Animal Ethics statement

All animal work was carried out in accordance with the European Communities Council Directive (2010/63/EU) and UK Animals Scientific Procedures Act (ASPA) 1986. Experimental procedures were carried out under personal and project licenses granted by the UK Home Office under ASPA, and ethically approved by the Animal Welfare and Ethical Review Body at the University of Exeter.

Acknowledgements

The authors would like to thank the staff in the Aquatic Resources Centre at the University of Exeter, for the supply and maintenance of the zebrafish. The *cmlc2::DsRed2-nuc* fish were generously provided by Dr. Geoff Burns, Boston Children's Hospital and Harvard Medical School. The authors would also like to thank Drs Keith Cheng and Alex Yu-Shun Lin, at Pennsylvania State University College of Medicine, for their advice on the use of the moulds for the histology.

Competing interests

SR, MB, AW and EM are employees of AstraZeneca PLC. There are no other potential conflicts of interest.

Funding

This work was supported by AstraZeneca, a Royal Society short Industry Fellowship (SIF\R2\192004) awarded to MJW, and the University of Exeter, College of Life and Environmental Sciences. Research in the Scholpp lab is supported by the BBSRC (Research Grant, BB/S016295/1) and by the Living Systems Institute, University of Exeter.

Data availability

All data associated with this study are available in the article or the supplementary materials.

Author contributions statement

SR, MB and MJH conceived the project and developed the concept. SR, MB, MJH, SS, CRT and MJW obtained the funding. AW, EM and MB undertook the *in silico* work. MJW, YO, JSB, MJH, SS, CRT and MB designed the zebrafish experiments. MJW, YO, JSB and AT undertook the zebrafish experiments and analysed the resultant data. Data interpretation and manuscript production was undertaken by all authors.

References

Aung, N., Vargas, J. D., Yang, C., Cabrera, C. P., Warren, H. R., Fung, K., Tzanis, E., Barnes, M. R., Rotter, J. I., Taylor, K. D., Manichaikul, A. W., Lima, J. a. C., Bluemke, D. A., Piechnik, S. K., Neubauer, S., Munroe, P. B. & Petersen, S. E. 2019. Genome-Wide Analysis of Left Ventricular Image-Derived Phenotypes Identifies Fourteen Loci Associated With Cardiac Morphogenesis and Heart Failure Development. *Circulation,* 140**,** 1318-1330.

Ball, J. S., Stedman, D. B., Hillegass, J. M., Zhang, C. X., Panzica-Kelly, J., Coburn, A., Enright, B. P., Tornesi, B., Amouzadeh, H. R., Hetheridge, M., Gustafson, A. L. & Augustine-Rauch, K. A. 2014. Fishing for teratogens: a consortium effort for a harmonized zebrafish developmental toxicology assay. *Toxicol Sci,* 139**,** 210-9.

Bong, S. M., Bae, S. H., Song, B., Gwak, H., Yang, S. W., Kim, S., Nam, S., Rajalingam, K., Oh, S. J., Kim, T. W., Park, S., Jang, H. & Lee, B. I. 2020. Regulation of mRNA export through API5 and nuclear FGF2 interaction. *Nucleic Acids Res,* 48**,** 6340-6352.

Cappola, T. P., Li, M., He, J., Ky, B., Gilmore, J., Qu, L., Keating, B., Reilly, M., Kim, C. E., Glessner, J., Frackelton, E., Hakonarson, H., Syed, F., Hindes, A., Matkovich, S. J., Cresci, S. & Dorn, G. W., 2nd 2010. Common variants in HSPB7 and FRMD4B associated with advanced heart failure. *Circ Cardiovasc Genet,* 3**,** 147-54.

Chambers, J. & Rabbitts, T. H. 2015. LMO2 at 25 years: a paradigm of chromosomal translocation proteins. *Open Biol,* 5**,** 150062.

Chen, B., Yates, E., Huang, Y., Kogut, P., Ma, L., Turner, J. R., Tao, Y., Camoretti-Mercado, B., Lang, D., Svensson, E. C., Garcia, J. G., Gruber, P. J., Morrisey, E. E. & Solway, J. 2009. Alternative promoter and GATA5 transcripts in mouse. *Am J Physiol Gastrointest Liver Physiol,* 297**,** G1214-22.

Copper, J. E., Budgeon, L. R., Foutz, C. A., Van Rossum, D. B., Vanselow, D. J., Hubley, M. J., Clark, D. P., Mandrell, D. T. & Cheng, K. C. 2018. Comparative analysis of fixation and embedding techniques for optimized histological preparation of zebrafish. *Comp Biochem Physiol C Toxicol Pharmacol,* 208**,** 38-46.

Esslinger, U., Garnier, S., Korniat, A., Proust, C., Kararigas, G., Müller-Nurasyid, M., Empana, J. P., Morley, M. P., Perret, C., Stark, K., Bick, A. G., Prasad, S. K., Kriebel, J., Li, J., Tiret, L., Strauch, K., O'regan, D. P., Marguiles, K. B., Seidman, J. G., Boutouyrie, P., Lacolley, P., Jouven, X., Hengstenberg, C., Komajda, M., Hakonarson, H., Isnard, R., Arbustini, E., Grallert, H., Cook, S. A., Seidman, C. E., Regitz-Zagrosek, V., Cappola, T. P., Charron, P., Cambien, F. & Villard, E. 2017. Exome-wide association study reveals novel susceptibility genes to sporadic dilated cardiomyopathy. *PLoS One,* 12**,** e0172995.

Gering, M., Yamada, Y., Rabbitts, T. H. & Patient, R. K. 2003. Lmo2 and Scl/Tal1 convert non-axial mesoderm into haemangioblasts which differentiate into endothelial cells in the absence of Gata1. *Development,* 130**,** 6187-99.

Gu, J. Y., Xu, J. H., Yu, H. & Yang, Y. Q. 2012. Novel GATA5 loss-of-function mutations underlie familial atrial fibrillation. *Clinics (Sao Paulo),* 67**,** 1393-9.

Gustafson, A. L., Stedman, D. B., Ball, J., Hillegass, J. M., Flood, A., Zhang, C. X., Panzica-Kelly, J., Cao, J., Coburn, A., Enright, B. P., Tornesi, M. B., Hetheridge, M. & Augustine-Rauch, K. A. 2012. Inter-laboratory assessment of a harmonized zebrafish developmental toxicology assay - progress report on phase I. *Reprod Toxicol,* 33**,** 155-64.

Gut, P., Reischauer, S., Stainier, D. Y. R. & Arnaout, R. 2017. Little Fish, Big Data: Zebrafish as a Model for Cardiovascular and Metabolic Disease. *Physiological Reviews,* 97**,** 889-938.

Jiang, J. Q., Li, R. G., Wang, J., Liu, X. Y., Xu, Y. J., Fang, W. Y., Chen, X. Z., Zhang, W., Wang, X. Z. & Yang, Y. Q. 2013. Prevalence and spectrum of GATA5 mutations associated with congenital heart disease. *Int J Cardiol,* 165**,** 570-3.

Ke, L., Meijering, R. a. M., Hoogstra-Berends, F., Mackovicova, K., Vos, M. J., Van Gelder, I. C., Henning, R. H., Kampinga, H. H. & Brundel, B. J. J. M. 2011. HSPB1, HSPB6, HSPB7 and HSPB8 Protect against RhoA GTPase-Induced Remodeling in Tachypaced Atrial Myocytes. *PLOS ONE,* 6**,** e20395.

Kjekshus, J., Apetrei, E., Barrios, V., Böhm, M., Cleland, J. G., Cornel, J. H., Dunselman, P., Fonseca, C., Goudev, A., Grande, P., Gullestad, L., Hjalmarson, A., Hradec, J., Jánosi, A., Kamenský, G., Komajda, M., Korewicki, J., Kuusi, T., Mach, F., Mareev, V., Mcmurray, J.

J., Ranjith, N., Schaufelberger, M., Vanhaecke, J., Van Veldhuisen, D. J., Waagstein, F., Wedel, H. & Wikstrand, J. 2007. Rosuvastatin in older patients with systolic heart failure. *N Engl J Med,* 357**,** 2248-61.

Kroll, F., Powell, G. T., Ghosh, M., Gestri, G., Antinucci, P., Hearn, T. J., Tunbak, H., Lim, S., Dennis, H. W., Fernandez, J. M., Whitmore, D., Dreosti, E., Wilson, S. W., Hoffman, E. J. & Rihel, J. 2021. A simple and effective F0 knockout method for rapid screening of behaviour and other complex phenotypes. *Elife,* 10.

Liao, W.-C., Juo, L.-Y., Shih, Y.-L., Chen, Y.-H. & Yan, Y.-T. 2017. HSPB7 prevents cardiac conduction system defect through maintaining intercalated disc integrity. *PLOS Genetics,* 13**,** e1006984.

Lu, W., Huang, S. Y., Su, L., Zhao, B. X. & Miao, J. Y. 2016. Long Noncoding RNA LOC100129973 Suppresses Apoptosis by Targeting miR-4707-5p and miR-4767 in Vascular Endothelial Cells. *Scientific reports,* 6**,** 21620-21620.

Mably, J. D., Mohideen, M. A., Burns, C. G., Chen, J. N. & Fishman, M. C. 2003. heart of glass regulates the concentric growth of the heart in zebrafish. *Curr Biol,* 13**,** 2138-47.

Macrae, C. A. & Peterson, R. T. 2015. Zebrafish as tools for drug discovery. *Nat Rev Drug Discov,* 14**,** 721-31.

Mao, H. Y., Liu, L. N. & Hu, Y. M. 2020. Mesenchymal stem cells-derived exosomal miRNA-28-3p promotes apoptosis of pulmonary endothelial cells in pulmonary embolism. *Eur Rev Med Pharmacol Sci,* 24**,** 10619-10631.

Marques, I. J., Leito, J. T., Spaink, H. P., Testerink, J., Jaspers, R. T., Witte, F., Van Den Berg, S. & Bagowski, C. P. 2008. Transcriptome analysis of the response to chronic constant hypoxia in zebrafish hearts. *J Comp Physiol B,* 178**,** 77-92.

Meeker, N. D., Hutchinson, S. A., Ho, L. & Trede, N. S. 2007. Method for isolation of PCR-ready genomic DNA from zebrafish tissues. *Biotechniques,* 43**,** 610, 612, 614.

Mercer, E. J., Lin, Y.-F., Cohen-Gould, L. & Evans, T. 2018. Hspb7 is a cardioprotective chaperone facilitating sarcomeric proteostasis. *Developmental Biology,* 435**,** 41-55.

Parker, T., Libourel, P. A., Hetheridge, M. J., Cumming, R. I., Sutcliffe, T. P., Goonesinghe, A. C., Ball, J. S., Owen, S. F., Chomis, Y. & Winter, M. J. 2014. A multi-endpoint in vivo larval zebrafish (Danio rerio) model for the assessment of integrated cardiovascular function. *J Pharmacol Toxicol Methods,* 69**,** 30-8.

Pfeffer, M. A., Swedberg, K., Granger, C. B., Held, P., Mcmurray, J. J., Michelson, E. L., Olofsson, B., Ostergren, J., Yusuf, S. & Pocock, S. 2003. Effects of candesartan on mortality and morbidity in patients with chronic heart failure: the CHARM-Overall programme. *Lancet,* 362**,** 759-66.

Povysil, G., Chazara, O., Carss, K. J., Deevi, S. V. V., Wang, Q., Armisen, J., Paul, D. S., Granger, C. B., Kjekshus, J., Aggarwal, V., Haefliger, C. & Goldstein, D. B. 2020. Assessing the Role of Rare Genetic Variation in Patients With Heart Failure. *JAMA Cardiol*.

Reiter, J. F., Alexander, J., Rodaway, A., Yelon, D., Patient, R., Holder, N. & Stainier, D. Y. 1999. Gata5 is required for the development of the heart and endoderm in zebrafish. *Genes Dev,* 13**,** 2983-95.

Reiter, J. F., Kikuchi, Y. & Stainier, D. Y. 2001. Multiple roles for Gata5 in zebrafish endoderm formation. *Development,* 128**,** 125-35.

Rosenfeld, G. E., Mercer, E. J., Mason, C. E. & Evans, T. 2013. Small heat shock proteins Hspb7 and Hspb12 regulate early steps of cardiac morphogenesis. *Developmental Biology,* 381**,** 389-400.

Sabaliauskas, N. A., Foutz, C. A., Mest, J. R., Budgeon, L. R., Sidor, A. T., Gershenson, J. A., Joshi, S. B. & Cheng, K. C. 2006. High-throughput zebrafish histology. *Methods,* 39**,** 246-54.

Senni, M., Paulus, W. J., Gavazzi, A., Fraser, A. G., Díez, J., Solomon, S. D., Smiseth, O. A., Guazzi, M., Lam, C. S., Maggioni, A. P., Tschöpe, C., Metra, M., Hummel, S. L., Edelmann, F., Ambrosio, G., Stewart Coats, A. J., Filippatos, G. S., Gheorghiade, M., Anker, S. D., Levy, D., Pfeffer, M. A., Stough, W. G. & Pieske, B. M. 2014. New strategies for heart failure with preserved ejection fraction: the importance of targeted therapies for heart failure phenotypes. *Eur Heart J,* 35**,** 2797-815.

Villard, E., Perret, C., Gary, F., Proust, C., Dilanian, G., Hengstenberg, C., Ruppert, V., Arbustini, E., Wichter, T., Germain, M., Dubourg, O., Tavazzi, L., Aumont, M. C., Degroote, P., Fauchier,

L., Trochu, J. N., Gibelin, P., Aupetit, J. F., Stark, K., Erdmann, J., Hetzer, R., Roberts, A. M., Barton, P. J., Regitz-Zagrosek, V., Aslam, U., Duboscq-Bidot, L., Meyborg, M., Maisch, B., Madeira, H., Waldenström, A., Galve, E., Cleland, J. G., Dorent, R., Roizes, G., Zeller, T., Blankenberg, S., Goodall, A. H., Cook, S., Tregouet, D. A., Tiret, L., Isnard, R., Komajda, M., Charron, P. & Cambien, F. 2011. A genome-wide association study identifies two loci associated with heart failure due to dilated cardiomyopathy. *Eur Heart J,* 32**,** 1065-76.

Wei, D., Bao, H., Zhou, N., Zheng, G. F., Liu, X. Y. & Yang, Y. Q. 2013. GATA5 loss-of-function mutation responsible for the congenital ventriculoseptal defect. *Pediatr Cardiol,* 34**,** 504-11.

Wu, T., Mu, Y., Bogomolovas, J., Fang, X., Veevers, J., Nowak, R. B., Pappas, C. T., Gregorio, C. C., Evans, S. M., Fowler, V. M. & Chen, J. 2017. HSPB7 is indispensable for heart development by modulating actin filament assembly. *Proceedings of the National Academy of Sciences,* 114**,** 11956-11961.

Zhang, X. L., Dai, N., Tang, K., Chen, Y. Q., Chen, W., Wang, J., Zhao, C. M., Yuan, F., Qiu, X. B., Qu, X. K., Yang, Y. Q. & Xu, Y. W. 2015. GATA5 loss-of-function mutation in familial dilated cardiomyopathy. *Int J Mol Med,* 35**,** 763-70.

Figure legends

Fig. 1. Snapshot of the results from the *in silico* assessment of the heart failure candidate genes *API5* (top right), *HSPB7* (bottom left), *LMO2* (bottom right), and positive control gene *GATA5* (top left panel). A. Top 10 tissues showing expression based on RNA sequencing data from the Human Protein Atlas (HPA) and EMBL-EBI Expression Atlas as summarized by Open Targets Platform (https://www.targetvalidation.org). **B.** Gene expression changes in cardiovascular disease conditions based on publicly available transcriptome studies from NCBI Gene Expression Omnibus (GEO). Disease conditions are shown on the Y-axis and log2 fold changes (vs normal controls) on the X-axis. Icons are coloured by direction of change; red and green represent up- and downregulation in disease, respectively. Icon shapes represent tissue type subjected to transcriptomics; circles and triangles represent heart and blood, respectively. Finally, icon size reflects statistical significance; the larger the icon the lower the P-value. All findings shown are significant (Adjusted P-value<0.05). **C.** Network-based functional predications using STRING. Colour of network edges indicate evidence types; dark pink: experimental, light blue: database, yellow: text mining, black: co-expression, violet: protein homology. **D.** Common variant gene locus association data from 190 datasets and 251 traits in Common Metabolic Diseases Knowledge Portal (CMDKP). Traits considered genome-wide significant (p-value $\leq 5x10^{-8}$) are highlighted (border-line significant traits are marked with *).

Fig. 2. Mutation efficiency of the gRNAs for each target gene. For each gene, the left hand gel images show (upper image) the bands obtained following targeted PCR of genomic DNA extracted from 4 individual animals injected with the two most effective CRISPR gRNAs + Cas9 (based on 2dpf morphological analysis), compared with the Cas9 injected control animals. The lower gel images show the same samples following T7E1 assay undertaken to reveal the amount of heteroduplex DNA present. The right hand image in each case shows the result of Sanger sequencing undertaken on representative genomic DNA samples from the most effective gRNA + Cas9 per target gene compared with that from a representative Cas9-injected control animal (top image). Data are shown

for A) *gata5*, the positive control gene, B) *api5*, C) *hspb7* and D) *lmo2*. Note in all cases the most effective gRNA was the combined guide group (g#1,2,3) except for *lmo2* which due to high mortality of the combined guide group, the g#2-injected animals were selected for full analysis. On the right hand edge of each panel, the number of animals assessed that showed efficient mutation of the target gene is shown, which in all cases was 4/4 (100%) in the presence of Cas9, but zero in the Cas9 control group animals.

Fig. 3. Results of the morphological analysis of 4dpf mutant *gata5* (positive control), *api5, hspb7* and *lmo2* zebrafish larvae versus Cas9-injected control animals. Panel A: General whole body morphological endpoints measured following injection of Cas9 alone, or after mutation of each of the target genes. Data are shown as the % incidence of abnormalities under each category. Note different n-numbers present as two runs were undertaken for the Cas9 control and the gRNA + Cas9-injected group showing the strongest phenotype from run 1 (for *lmo2* g#2 was run twice due to concerns about excessive mortality in the g#1,2,3 group). The guide combination used for two runs in each case is shown in the lower panel of the example images for each gene. **Ai**: Expansion of heart-specific endpoints showing the full range scored. **B**: example larvae following *gata5* mutation versus the Cas9-injected control. The yellow arrows show the position of the pericardial membrane and the extent of pericardial oedema, which was minimal in the controls, but extensive in most mutant animals (two examples are shown for g#1 + Cas9-injected animals as there was some variability in the severities seen). **C-E**: examples of larvae from other target gene treatment categories (note the apparent lack of effect of *api5* mutation on general morphology). Scale bar shown in the first image of panels B-E represents 500μm.

Fig. 4. Results of the heart pathology analysis of 4dpf mutant *gata5* (positive control), *api5, hspb7* and *lmo2* zebrafish larvae versus Cas9-injected control animals. Panel A) Example haematoxylin and eosin stained coronal sections through the heart (top, A = atrium, V = ventricle, *bulbous arteriosus) from each of the treatment groups versus the Cas9-injected controls (left hand panels). Note in particular the extreme cardiac hypoplasia after *gata5* and *lmo2* mutation (indicated by

red arrow in images) in which the atrium is not visible possibly due to the severe pericardial oedema. In each panel, animals are orientated with the head to the left, and viewed in the dorsal plane at a magnification of 40x (Scale bar shown in right hand panel represents 200μm) **Panel B)** Example images of hearts from *cmlc2::DsRed2-nuc* mutant larvae in which the cardiomyocytes are labelled red, especially prominently in the ventricle. The top row of panels shows the image with transmitted light and *cmlc2::DsRed2-nuc* fluorescence signals, and the lower row shows the *cmlc2::DsRed2-nuc* signal alone. Note the severe oedema, weaker *cmlc2::DsRed2-nuc* fluorescence signal, reduced number of cells and smaller chamber size typical of the *gata5* mutant (indicated by yellow arrow in images); the oedema, and perhaps slightly enlarged ventricle observed in some *api5* mutants; and the severe oedema, disorganisation of myocytes and smaller chamber size typical of the *lmo2* mutants (yellow arrow). *hspb7* mutant larval hearts outwardly appeared no different to the Cas9 controls. Scale bar shown in left hand image represents 50μm. **Panel C)** Mean ventricle diameter estimated in mutants from each treatment at the maximum point of contraction (end systole), the maximum point of relaxation (end diastole) and the difference between the two, in arbitrary units. Data are shown as the mean, ± SEM, n=5. *signifies a statistically significant difference versus the Cas9 control for that parameter at $p<0.05$, ** at $p<0.01$, and *** at $p<0.001$ (Student's T-test or Mann Whitney U-tests). **Note**: the overall body lengths of the *gata5* and the *lmo2* mutants were also significantly reduced ($p<0.001$).

Fig. 5. Results of the analysis of cardiovascular function in 4dpf mutant *gata5* (positive control), *api5*, *hspb7* and *lmo2* zebrafish larvae versus Cas9-injected control animals. Ai-Aiv: Images of example Cas9 control larvae alongside larvae treated with the two gRNAs + Cas9 mixtures giving the strongest phenotypes (as assessed at 2dpf). The top row shows the trunk vasculature with the position of the dorsal aorta outlined in yellow dashed lines, where blood flow and vessel diameter measurements were taken. The lower row of images shows the heart from the same animals with the atrium highlighted by a small white arrow, and the ventricle by a small yellow arrow. The large red arrows show the position of the pericardial membrane and the extent of pericardial oedema. Most Cas9 control animals exhibited normal morphology and function in contrast with many of the mutant

animals. **Bi-Biv:** cardiovascular functional endpoints quantified in the same groups of animals Note the complete absence of blood flow measured in all of the *gata5* g#1,2,3 + Cas9 and in 6/10 of the *gata5* g#1 + Cas9-injected animals., the absence of effective blood flow in the *lmo2* g#1,2,3 + Cas9 and *lmo2* g#2 + Cas9-injected animals due to the absence of erythrocytes meaning flow was not visible. Vessel diameter measurements were only not possible in animals lacking blood flow (indicated by n/a). Data are shown as the mean % increase versus the Cas9-control group (100% indicated by the red dashed line), ± SEM, n=19-20 for the Cas9 and right hand mutant treatment for each gene (data combined from two runs) and 10 for the left hand treatment group for each gene where only one run was undertaken. *signifies a statistically significant difference versus the Cas9 control at $p<0.05$, ** at $p<0.01$, and *** at $p<0.001$ (T-test or Mann Whitney U-tests for the combined guide injected groups, or 1-way ANOVA and Tukey's HSD tests or Kruskal-Wallis and Dunn's tests for the single guide injected groups in which runs 1 and 2 were combined). Scale bar shown in upper left hand image of each panel represents 100µm. Scale bar shown in lower left hand image of each panel represents 50µm.

List of supplementary information

Table S1. Summary of literature data on the structure and function of the target genes identified for zebrafish-based in vivo investigation in the current study

Table S2. List of morphological endpoints scored as a simple yes/no at 2dpf to facilitate selection of treatment groups for more in depth morphological scoring and functional phenotype analysis at 4dpf

Table S3. List of morphological endpoints scored in severity from 1 (severe) to 5 (normal) at 4dpf. Scoring criteria adapted from those of Gustafson *et al*., (2012) and Ball *et al*., (2014)

Table S4. Sequence of fixation and embedding steps applied to 4dpf zebrafish using an automated tissue embedder

Table S5. Sequence of H&E staining steps applied to 4dpf zebrafish using an automated stainer

Fig. S1. GeneMANIA network analysis using API5, HSPB7, LMO2 and GATA5 as input genes.

Fig. S2. Transcriptomics data analysis using *API5*, *HSPB7*, *LMO2* and *GATA5* as input genes.

Fig. S3. Comparison of morphological endpoints measured in 2dpf animals.

Figure S4. Haematoxylin and eosin stained coronal sections through the hearts of 4dpf CRISPR-mutant zebrafish.

Fig. S5. Confocal maximum intensity projection images of hearts from *cmlc2::DsRed2-nuc* mutant larvae.

Fig. S6. gRNA design and mutation efficient for each of the 4 targeted genes. Panel A:

Supplementary videos. Videos of cardiovascular function to accompany Figure 5 in the main manuscript.

Supplementary datasets. Data S1. Data S1_Network Analysis.xlsx; Data S2. Data S2_Differential Expression.xlsx; Data S3. Data S3_Genetic Associations.xlsx; Data S4. Data S4_Homologies.xlsx

TABLE 1

Data type(s) and source(s)		*GATA5 (positive control)*	*API5*	*HSPB7*	*LMO2*	*Suppl. ref.*
LITERATURE	PubMed (https://pubmed.ncbi.nlm.nih.gov/)	Documented role in cardiovascular and cardiac development	Documented role in apoptosis and tumorigenesis	Documented role in UPR and cardiac function (largely unknown mechanisms). Common variants associate with advanced HF and DCM.	Documented role in erythropoiesis, angiogenesis and tumorigenesis	**Table S1**
NETWORK ANALYSIS	STRING (https://string-db.org/) Harmonizome (https://maayanlab.cloud/Harmonizome/) GeneMANIA (https://genemania.org/)	Corroborates documented roles in transcriptional regulation, embryonic development, stem cell differentiation and heart morphogenesis. Additionally, predicts role in blood coagulation, lipid metabolism and intestinal function & others	Predicts role in splicing and processing of mRNA, DNA replication and repair	Corroborates role in cardiac biology including sarcomere organization, myofibril assembly, cardiac contraction and UPR. Additionally, predicts role in cell-extracellular matrix interactions, focal adhesion, calcium signalling and cellular metabolic processes (inc. TCA cycle)	Corroborates documented roles in haematopoietic stem cell differentiation, erythropoiesis and transcriptional regulation. Additionally, predicts role in translation and immune response.	**Fig. 1C** **Data S1** **Fig. S1**
PATHWAY ANALYSIS	*KEGG (https://www.genome.jp/kegg/) **REACTOME (https://reactome.org/) ***WikiPathways (https://www.wikipathways.org/)	Factors involved in megakaryocyte development and platelet production** Haemostasis**	No known participation in canonical pathways	Differentiation of white and brown adipocyte***	Haematopoietic Stem Cell Differentiation*** Transcriptional mis-regulation in cancer*	
TRANSCRIPTOMICS	**Tissue & cell-type distribution (RNAseq)** Human Protein Atlas (HPA,	Expressed by multiple tissues outside CNS. Strongest expression in intestine, gall	Ubiquitously expressed across tissues and cell types	Selectively expressed by heart and skeletal muscle	Ubiquitously expressed across tissues. Selective expression in endothelial cells, immune cells	**Fig. 1A** **Fig. S2**

	(https://www.proteinatlas.org/) EMBL-EBI Expression Atlas (https://www.ebi.ac.uk/gxa/home)	bladder, urinary bladder, testis, heart and smooth muscle			and Ito cells	
	Cellular expression in human adult heart (scRNAseq), Heart Cell Atlas (https://www.heartcellatlas.org/)	Preferentially expressed by cardiomyocytes	Expressed by vascular, immune and mesenchymal cells	Expressed by cardiomyocytes	Preferentially expressed by endothelial and immune cells	**Fig. S2**
	Expression in CVD vs controls - heart and blood (microarray, RNAseq) NCBI Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/)	Most pronounced changes are seen in heart, with down-regulation in DCM and RCM and upregulation in HCM, sinus arrhythmia and AF. No significant changes in blood	Most pronounced changes include upregulated cardiac expression in several cardiomyopathies as well as unstable angina, and downregulated expression in blood in MI, VTE, pre-eclampsia.	Most pronounced changes are seen in heart, with downregulation in multiple cardiac disease conditions including HCM, HF and AF. Little significant changes in blood	Most pronounced changes are seen in blood, with upregulation in ICM, MI and CAD In heart, downregulated expression is seen in DCM	**Fig. 1B** **Data S2**
GENOMICS	Common Metabolic Diseases Knowledge Portal (CMDKP) (https://hugeamp.org/) Phenoscanner (http://www.phenoscanner.medschl.cam.ac.uk/) Online Mendelian Inheritance in Man (OMIM), ClinVar and GWAS Catalog accessed via Ensembl	Common variants associate with lung function (FEV1/FVC and FVC) and haematological traits (RBC count, haematocrit, haemoglobin). Rare variants are linked to Congenital heart defects, multiple types, 5 (CHTD5) (OMIM: # 617912)	Common variants associate with BMI and neutrophil count (borderline significant) Low frequency 3' UTR variant associates with "Cause of death: atrial fibrillation and flutter" in UK Biobank (P-value=6.43e-22)	Common variants associate with cardiac function traits (LVEF, LVESV and LVEDV), HF (suggestive), blood pressure (SBP and DBP), haematological traits (RBC count, haematocrit, haemoglobin), estimated BMD and serum creatinine	Common variants associate with haematological traits (MCH,MCV, RBC count, eosinophil count)	**Fig. 1D** **Data S3**

	(http://www.ensembl.org/)					
HOMOLOGY	Ensembl (http://www.ensembl.org/)	1-to-1 orthology. Human *GATA5* protein is 52% identical to the Zebrafish orthologue *gata5*	1-to-1 orthology Human *API5* protein is 83% identical to the Zebrafish orthologue *api5*	1-to-1 orthology Human *HSPB7* protein is 41% identical to the Zebrafish orthologue *hspb7*	1-to-1 orthology Human *LMO2* protein is 62% identical to the Zebrafish orthologue *lmo2*	**Data S4**

Table 1. Summary of main findings from the *in silico* assessment of gene variant data including source of data used in each case. Additional, more extensive datasets are contained within the supplementary results section as referenced in the right-hand most column. *Abbreviations*: atrial fibrillation (AF), bone mineral density (BMD), coronary artery disease (CAD), diastolic blood pressure (DBP), dilated cardiomyopathy (DCM), forced expiratory volume in 1 second (FEV1), forced vital capacity (FVC), heart failure (HF), hypertrophic cardiomyopathy (HCM), ischemic cardiomyopathy (ICM), left ventricular ejection fraction (LVEF), left ventricular end systolic volume (LVESV), mean corpuscular haemoglobin (MCH), mean corpuscular volume (MCV), myocardial infarction (MI), red blood cell (RBC), restrictive cardiomyopathy (RCM), systolic blood pressure (SBP), unfolded protein response (UPR), venous thromboembolism (VTE)

FIGURE 1

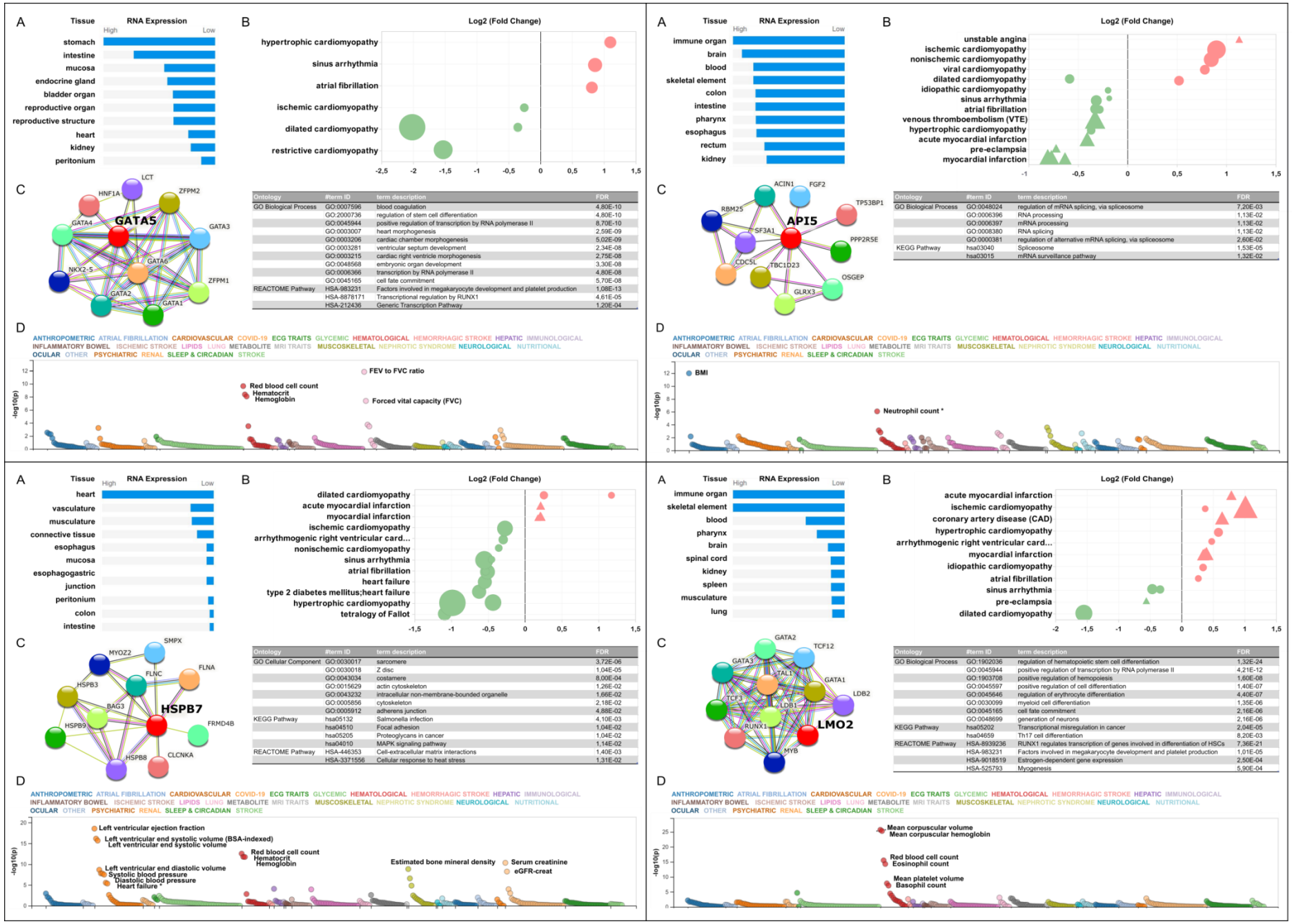

A
Tissue
RNA Expression
High
Low
stomach
intestine
mucosa
endocrine gland
bladder organ
reproductive organ
reproductive structure
heart
kidney
peritonium
B
Log2 (Fold Change)
hypertrophic cardiomyopathy
sinus arrhythmia
atrial fibrillation
ischemic cardiomyopathy
dilated cardiomyopathy
restrictive cardiomyopathy
-2,5
-2
-1,5
-1
-0,5
0
0,5
1
1,5
C
LCT
HNF1A
ZFPM2
GATA4
GATA5
GATA3
GATA6
NKX2-5
ZFPM1
GATA2
GATA1
Ontology | #term ID | term description | FDR
GO Biological Process | GO:0007596 | blood coagulation | 4,80E-10
GO:2000736 | regulation of stem cell differentiation | 4,80E-10
GO:0045944 | positive regulation of transcription by RNA polymerase II | 8,70E-10
GO:0003007 | heart morphogenesis | 2,59E-09
GO:0003206 | cardiac chamber morphogenesis | 5,02E-09
GO:0003281 | ventricular septum development | 2,34E-08
GO:0003215 | cardiac right ventricle morphogenesis | 2,75E-08
GO:0048568 | embryonic organ development | 3,30E-08
GO:0006366 | transcription by RNA polymerase II | 4,80E-08
GO:0045165 | cell fate commitment | 5,70E-08
REACTOME Pathway | HSA-983231 | Factors involved in megakaryocyte development and platelet production | 1,08E-13
HSA-8878171 | Transcriptional regulation by RUNX1 | 4,61E-05
HSA-212436 | Generic Transcription Pathway | 1,20E-04
D
ANTHROPOMETRIC ATRIAL FIBRILLATION CARDIOVASCULAR COVID-19 ECG TRAITS GLYCEMIC HEMATOLOGICAL HEMORRHAGIC STROKE HEPATIC IMMUNOLOGICAL
INFLAMMATORY BOWEL ISCHEMIC STROKE LIPIDS LUNG METABOLITE MRI TRAITS MUSCOSKELETAL NEPHROTIC SYNDROME NEUROLOGICAL NUTRITIONAL
OCULAR OTHER PSYCHIATRIC RENAL SLEEP & CIRCADIAN STROKE
-log10(p)
FEV to FVC ratio
Red blood cell count
Hematocrit
Hemoglobin
Forced vital capacity (FVC)
A
Tissue
RNA Expression
High
Low
immune organ
brain
blood
skeletal element
colon
intestine
pharynx
esophagus
rectum
kidney
B
Log2 (Fold Change)
unstable angina
ischemic cardiomyopathy
nonischemic cardiomyopathy
viral cardiomyopathy
dilated cardiomyopathy
idiopathic cardiomyopathy
sinus arrhythmia
atrial fibrillation
venous thromboembolism (VTE)
hypertrophic cardiomyopathy
acute myocardial infarction
pre-eclampsia
myocardial infarction
-1
-0,5
0
0,5
1
1,5
C
ACIN1
FGF2
RBM25
TP53BP1
API5
SF3A1
PPP2R5E
CDC5L
TBC1D23
OSGEP
GLRX3
Ontology | #term ID | term description | FDR
GO Biological Process | GO:0048024 | regulation of mRNA splicing, via spliceosome | 7,20E-03
GO:0006396 | RNA processing | 1,13E-02
GO:0006397 | mRNA processing | 1,13E-02
GO:0008380 | RNA splicing | 1,13E-02
GO:0000381 | regulation of alternative mRNA splicing, via spliceosome | 2,60E-02
KEGG Pathway | hsa03040 | Spliceosome | 1,53E-05
hsa03015 | mRNA surveillance pathway | 1,32E-02
D
ANTHROPOMETRIC ATRIAL FIBRILLATION CARDIOVASCULAR COVID-19 ECG TRAITS GLYCEMIC HEMATOLOGICAL HEMORRHAGIC STROKE HEPATIC IMMUNOLOGICAL
INFLAMMATORY BOWEL ISCHEMIC STROKE LIPIDS LUNG METABOLITE MRI TRAITS MUSCOSKELETAL NEPHROTIC SYNDROME NEUROLOGICAL NUTRITIONAL
OCULAR OTHER PSYCHIATRIC RENAL SLEEP & CIRCADIAN STROKE
-log10(p)
BMI
Neutrophil count *
A
Tissue
RNA Expression
High
Low
heart
vasculature
musculature
connective tissue
esophagus
mucosa
esophagogastric
junction
peritonium
colon
intestine
B
Log2 (Fold Change)
dilated cardiomyopathy
acute myocardial infarction
myocardial infarction
ischemic cardiomyopathy
arrhythmogenic right ventricular card...
nonischemic cardiomyopathy
sinus arrhythmia
atrial fibrillation
heart failure
type 2 diabetes mellitus;heart failure
hypertrophic cardiomyopathy
tetralogy of Fallot
-1,5
-1
-0,5
0
0,5
1
1,5
C
SMPX
MYOZ2
FLNA
FLNC
HSPB3
BAG3
HSPB7
HSPB9
FRMD4B
HSPB8
CLCNKA
Ontology | #term ID | term description | FDR
GO Cellular Component | GO:0030017 | sarcomere | 3,72E-06
GO:0030018 | Z disc | 1,04E-05
GO:0043034 | costamere | 8,00E-04
GO:0015629 | actin cytoskeleton | 1,26E-02
GO:0043232 | intracellular non-membrane-bounded organelle | 1,66E-02
GO:0005856 | cytoskeleton | 2,18E-02
GO:0005912 | adherens junction | 4,88E-02
KEGG Pathway | hsa05132 | Salmonella infection | 4,10E-03
hsa04510 | Focal adhesion | 1,04E-02
hsa05205 | Proteoglycans in cancer | 1,04E-02
hsa04010 | MAPK signaling pathway | 1,14E-02
REACTOME Pathway | HSA-446353 | Cell-extracellular matrix interactions | 1,40E-03
HSA-3371556 | Cellular response to heat stress | 1,31E-02
D
ANTHROPOMETRIC ATRIAL FIBRILLATION CARDIOVASCULAR COVID-19 ECG TRAITS GLYCEMIC HEMATOLOGICAL HEMORRHAGIC STROKE HEPATIC IMMUNOLOGICAL
INFLAMMATORY BOWEL ISCHEMIC STROKE LIPIDS LUNG METABOLITE MRI TRAITS MUSCOSKELETAL NEPHROTIC SYNDROME NEUROLOGICAL NUTRITIONAL
OCULAR OTHER PSYCHIATRIC RENAL SLEEP & CIRCADIAN STROKE
-log10(p)
Left ventricular ejection fraction
Left ventricular end systolic volume (BSA-indexed)
Left ventricular end systolic volume
Red blood cell count
Hematocrit
Hemoglobin
Left ventricular end diastolic volume
Systolic blood pressure
Diastolic blood pressure
Heart failure *
Estimated bone mineral density
Serum creatinine
eGFR-creat
A
Tissue
RNA Expression
High
Low
immune organ
skeletal element
blood
pharynx
brain
spinal cord
kidney
spleen
musculature
lung
B
Log2 (Fold Change)
acute myocardial infarction
ischemic cardiomyopathy
coronary artery disease (CAD)
hypertrophic cardiomyopathy
arrhythmogenic right ventricular card...
myocardial infarction
idiopathic cardiomyopathy
atrial fibrillation
sinus arrhythmia
pre-eclampsia
dilated cardiomyopathy
-2
-1,5
-1
-0,5
0
0,5
1
1,5
C
GATA2
TCF12
GATA3
TAL1
GATA1
LDB2
TCF3
LDB1
RUNX1
LMO2
MYB
Ontology | #term ID | term description | FDR
GO Biological Process | GO:1902036 | regulation of hematopoietic stem cell differentiation | 1,32E-24
GO:0045944 | positive regulation of transcription by RNA polymerase II | 4,21E-12
GO:1903708 | positive regulation of hemopoiesis | 1,60E-08
GO:0045597 | positive regulation of cell differentiation | 1,40E-07
GO:0045646 | regulation of erythrocyte differentiation | 4,40E-07
GO:0030099 | myeloid cell differentiation | 1,35E-06
GO:0045165 | cell fate commitment | 2,16E-06
GO:0048699 | generation of neurons | 2,16E-06
KEGG Pathway | hsa05202 | Transcriptional misregulation in cancer | 2,04E-05
hsa04659 | Th17 cell differentiation | 8,20E-03
REACTOME Pathway | HSA-8939236 | RUNX1 regulates transcription of genes involved in differentiation of HSCs | 7,36E-21
HSA-983231 | Factors involved in megakaryocyte development and platelet production | 1,01E-05
HSA-9018519 | Estrogen-dependent gene expression | 2,50E-04
HSA-525793 | Myogenesis | 5,90E-04
D
ANTHROPOMETRIC ATRIAL FIBRILLATION CARDIOVASCULAR COVID-19 ECG TRAITS GLYCEMIC HEMATOLOGICAL HEMORRHAGIC STROKE HEPATIC IMMUNOLOGICAL
INFLAMMATORY BOWEL ISCHEMIC STROKE LIPIDS LUNG METABOLITE MRI TRAITS MUSCOSKELETAL NEPHROTIC SYNDROME NEUROLOGICAL NUTRITIONAL
OCULAR OTHER PSYCHIATRIC RENAL SLEEP & CIRCADIAN STROKE
-log10(p)
Mean corpuscular volume
Mean corpuscular hemoglobin
Red blood cell count
Eosinophil count
Mean platelet volume
Basophil count

FIGURE 2

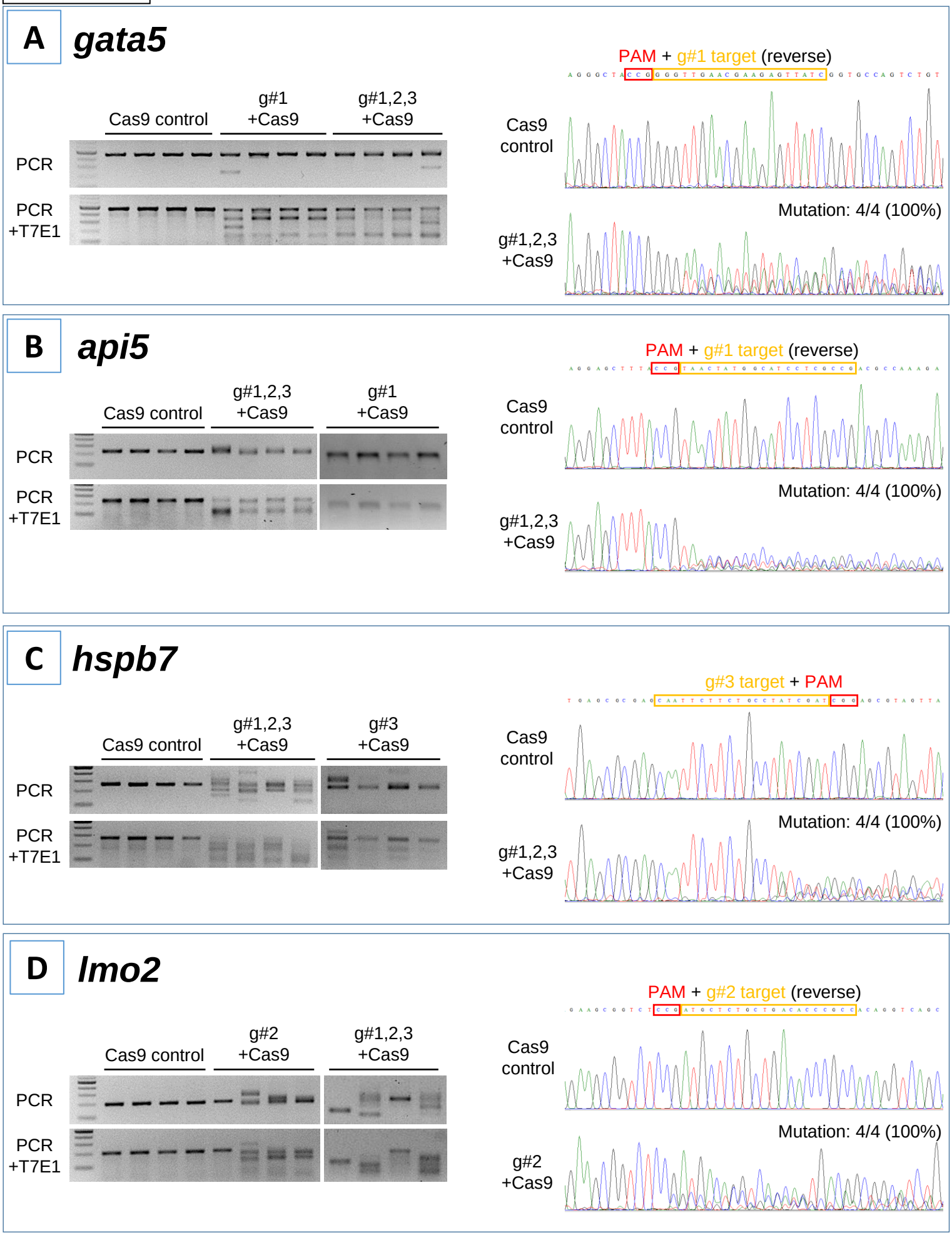

A
gata5
Cas9 control
g#1 +Cas9
g#1,2,3 +Cas9
PCR
PCR +T7E1
PAM + g#1 target (reverse)
Cas9 control
Mutation: 4/4 (100%)
g#1,2,3 +Cas9
B
api5
Cas9 control
g#1,2,3 +Cas9
g#1 +Cas9
PCR
PCR +T7E1
PAM + g#1 target (reverse)
Cas9 control
Mutation: 4/4 (100%)
g#1,2,3 +Cas9
C
hspb7
Cas9 control
g#1,2,3 +Cas9
g#3 +Cas9
PCR
PCR +T7E1
g#3 target + PAM
Cas9 control
Mutation: 4/4 (100%)
g#1,2,3 +Cas9
D
lmo2
Cas9 control
g#2 +Cas9
g#1,2,3 +Cas9
PCR
PCR +T7E1
PAM + g#2 target (reverse)
Cas9 control
Mutation: 4/4 (100%)
g#2 +Cas9

FIGURE 3

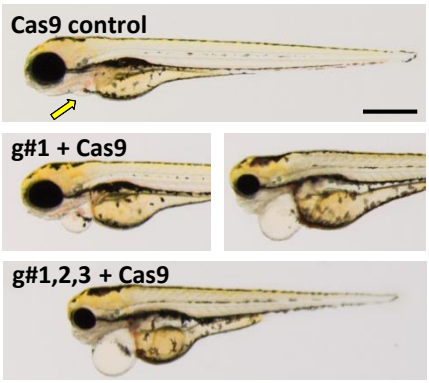
A

	Malformations at 4 days post fertilisation (4dpf)																	
	n-number (scored)	% Larvae Abnormal	Body Shape	Somites	Notochord	Tail	Fins	Heart	Face	Neural Tube	Arches/Jaws	Group Ave length (mm)	Swim bladder not inflated	Swim bladder not evident	Yolk ball _ Remnant excessive	Liver - not evident	Liver -Enlarged	Abdomen distended/yolk oedema
Cas9	20	20%	0%	0%	0%	0%	0%	20%	0%	0%	0%	3.58	50%	0%	0%	0%	0%	0%
gata5 g#1 + Cas9	10	100%	10%	50%	0%	0%	40%	100%	60%	60%	0%	3.31	100%	60%	0%	60%	0%	50%
gata5 g#1,2,3 + Cas9	20	100%	0%	80%	20%	0%	90%	100%	85%	90%	95%	3.15	100%	85%	0%	25%	0%	75%
Cas9	20	5%	0%	0%	0%	0%	0%	5%	0%	0%	0%	3.62	0%	0%	0%	0%	0%	0%
api5 g#1 + Cas9	10	40%	0%	0%	0%	0%	0%	40%	0%	0%	0%	3.69	0%	0%	0%	0%	0%	0%
api5 g#1,2,3 + Cas9	20	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3.67	0%	0%	0%	0%	0%	0%
Cas9	20	5%	0%	0%	0%	0%	0%	5%	0%	0%	0%	3.65	100%	0%	0%	0%	0%	0%
hspb7 g#3 + Cas9	10	90%	0%	0%	0%	0%	40%	80%	10%	30%	70%	3.55	100%	0%	0%	0%	0%	0%
hspb7 g#1,2,3 + Cas9	20	75%	20%	20%	20%	15%	60%	60%	30%	45%	55%	3.28	100%	15%	20%	15%	5%	5%
Cas9	20	15%	0%	0%	0%	0%	0%	10%	0%	5%	10%	3.58	100%	0%	0%	0%	0%	0%
lmo2 g#1,2,3 + Cas9	10	100%	40%	90%	50%	60%	100%	100%	100%	100%	100%	2.30	100%	100%	100%	100%	100%	0%
lmo2 g#2 + Cas9	20	100%	15%	100%	20%	20%	90%	100%	100%	100%	100%	3.08	100%	100%	100%	60%	20%	50%

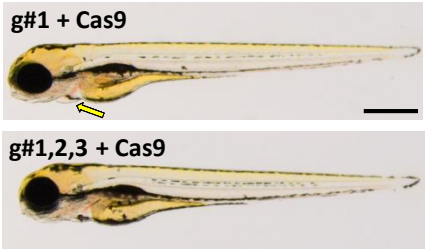
Ai

	Heart-specific malformations at 4dpf								
	n-number (scored)	% Abnormal	Slow Heartbeat	No Heartbeat	Pericardial Sac-Swollen	Chambers not well defined	Misshapen	Small	Enlarged
Cas9	20	20%	0%	0%	20%	0%	0%	0%	0%
gata5 g#1 + Cas9	10	100%	0%	0%	100%	60%	60%	60%	0%
gata5 g#1,2,3 + Cas9	20	100%	0%	5%	100%	40%	85%	85%	0%
Cas9	20	5%	0%	0%	10%	0%	0%	0%	0%
api5 g#1 + Cas9	10	40%	0%	0%	40%	0%	0%	0%	0%
api5 g#1,2,3 + Cas9	20	0%	0%	0%	0%	0%	0%	0%	0%
Cas9	20	5%	0%	0%	45%	0%	0%	0%	0%
hspb7 g#3 + Cas9	10	80%	0%	0%	90%	0%	0%	0%	0%
hspb7 g#1,2,3 + Cas9	20	60%	5%	5%	70%	15%	20%	20%	0%
Cas9	20	10%	0%	0%	5%	0%	0%	0%	0%
lmo2 g#1,2,3 + Cas9	10	100%	30%	0%	100%	100%	100%	100%	0%
lmo2 g#2 + Cas9	20	100%	0%	0%	100%	100%	100%	100%	0%

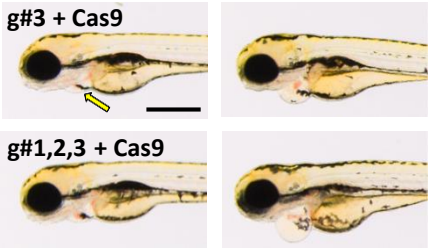
B ***gata5* (+ve)**



C ***api5***



D ***hspb7***



E ***lmo2***

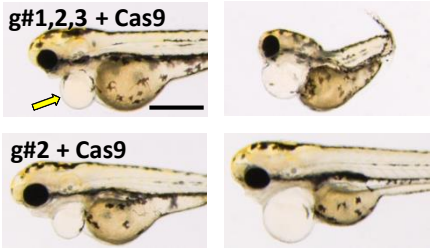


FIGURE 4

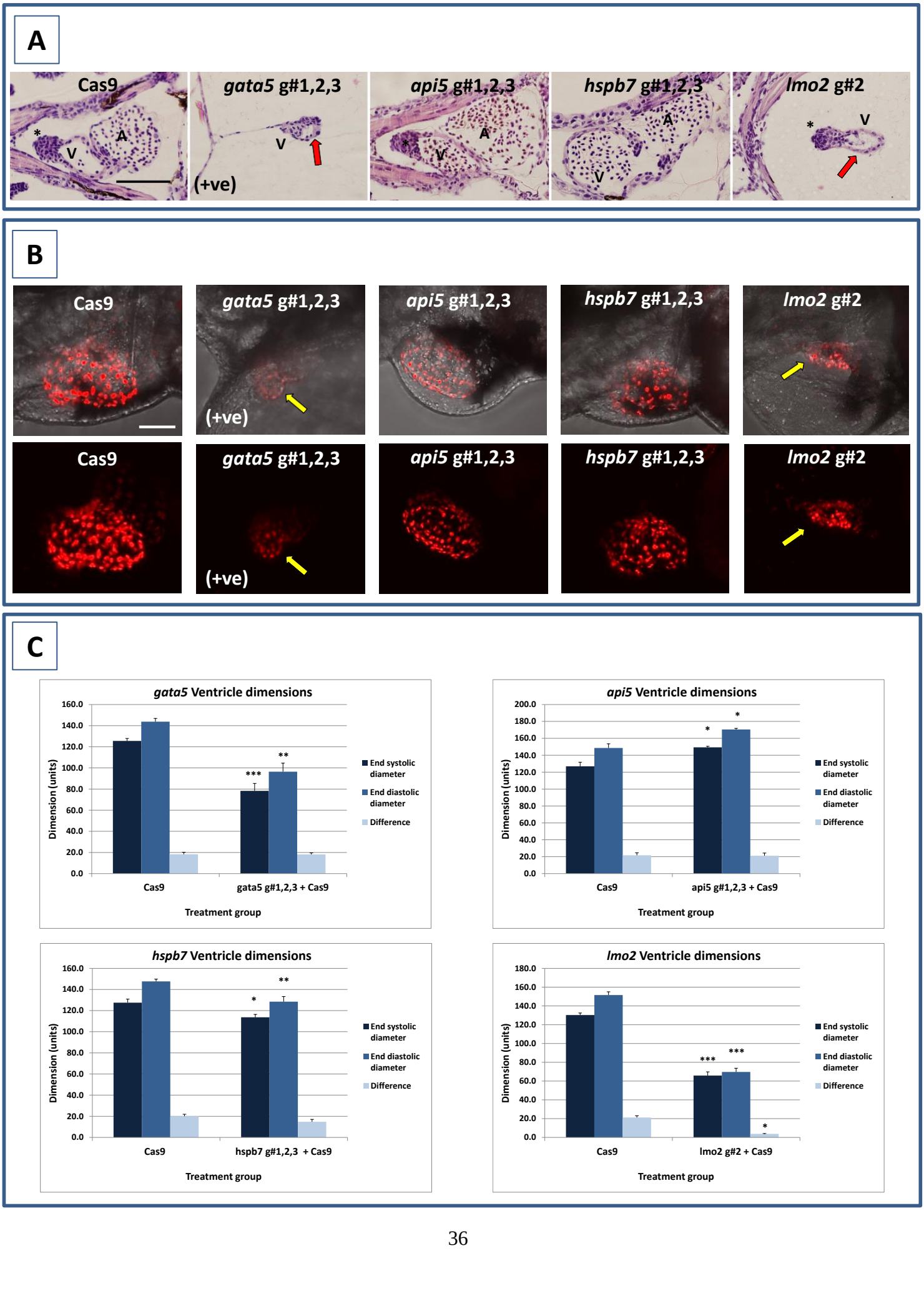
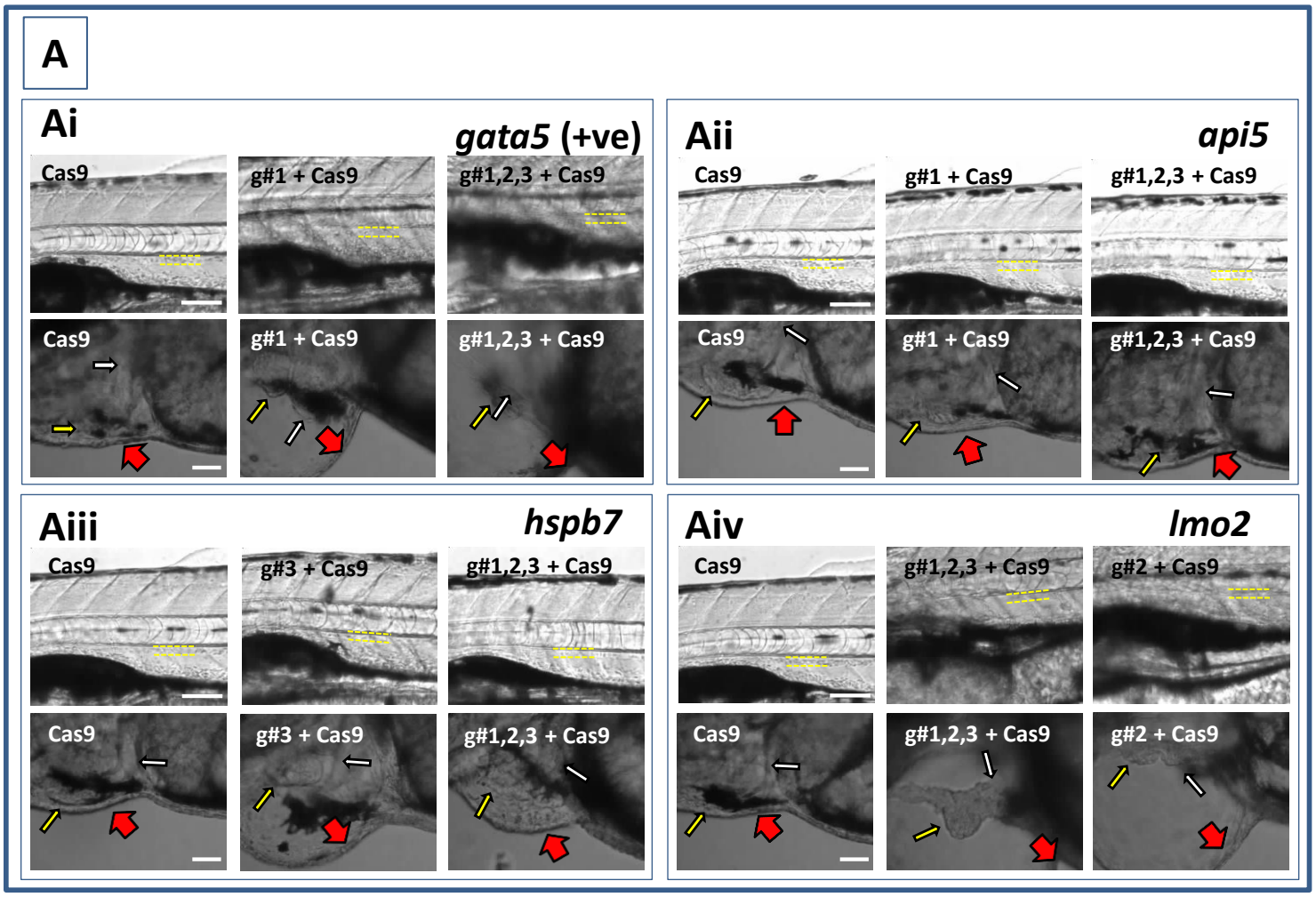
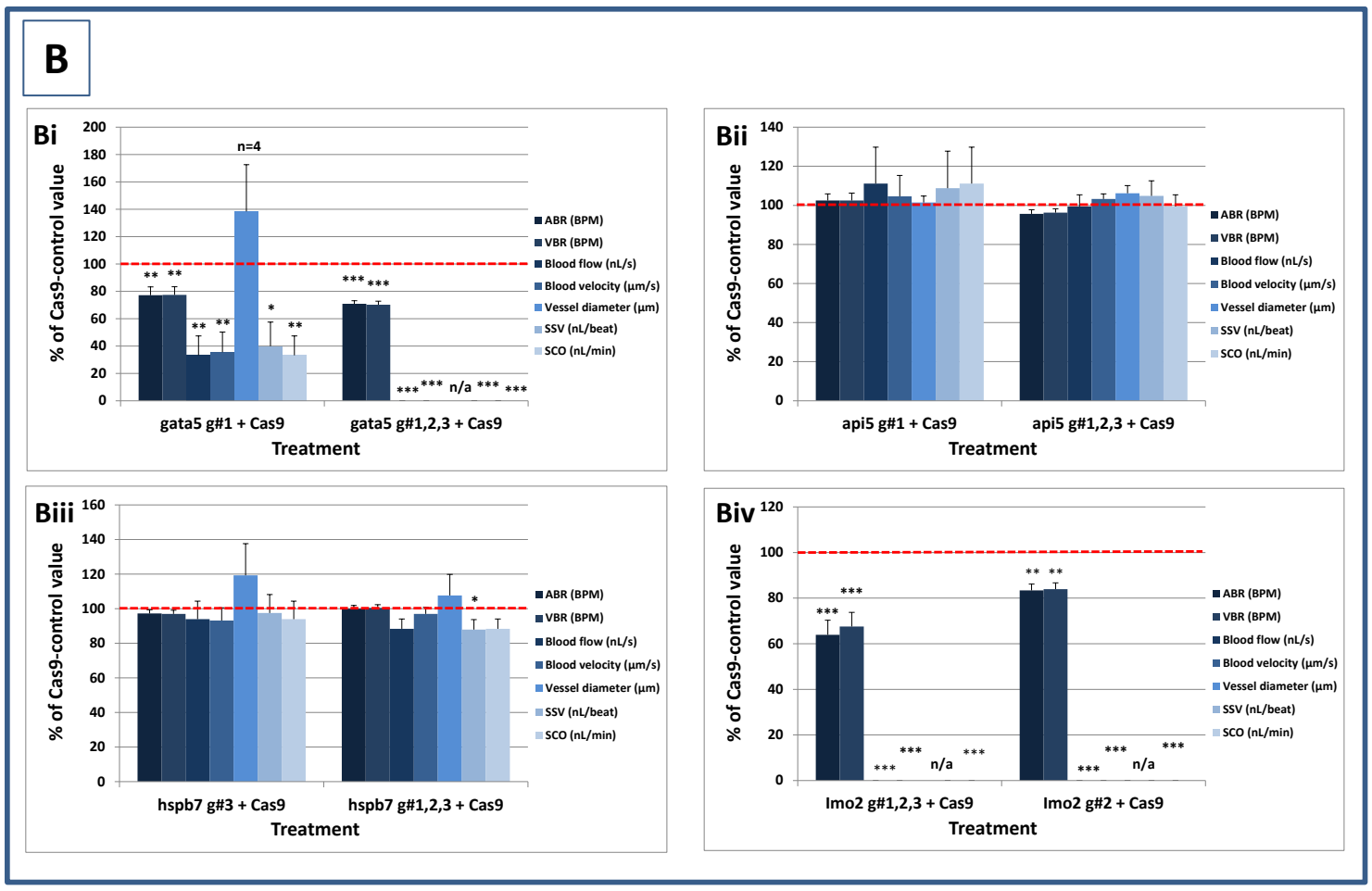

A
Cas9
gata5 g#1,2,3
(+ve)
api5 g#1,2,3
hspb7 g#1,2,3
lmo2 g#2
V
A
*
B
Cas9
gata5 g#1,2,3
(+ve)
api5 g#1,2,3
hspb7 g#1,2,3
lmo2 g#2
C
gata5 Ventricle dimensions
Dimension (units)
Treatment group
Cas9
gata5 g#1,2,3 + Cas9
End systolic diameter
End diastolic diameter
Difference
***
**
api5 Ventricle dimensions
Cas9
api5 g#1,2,3 + Cas9
hspb7 Ventricle dimensions
Cas9
hspb7 g#1,2,3 + Cas9
lmo2 Ventricle dimensions
Cas9
lmo2 g#2 + Cas9

FIGURE 5


A
Ai
gata5 (+ve)
Cas9
g#1 + Cas9
g#1,2,3 + Cas9
Aii
api5
Aiii
hspb7
g#3 + Cas9
Aiv
lmo2
g#2 + Cas9


B
Bi
% of Cas9-control value
n=4
gata5 g#1 + Cas9
gata5 g#1,2,3 + Cas9
Treatment
Bii
api5 g#1 + Cas9
api5 g#1,2,3 + Cas9
Biii
hspb7 g#3 + Cas9
hspb7 g#1,2,3 + Cas9
Biv
lmo2 g#1,2,3 + Cas9
lmo2 g#2 + Cas9
n/a
ABR (BPM)
VBR (BPM)
Blood flow (nL/s)
Blood velocity (µm/s)
Vessel diameter (µm)
SSV (nL/beat)
SCO (nL/min)