Patient-specific functional genomics and disease modeling suggest a role for LRP2 in Hypoplastic Left Heart Syndrome

Xing Li, PhD*; Almudena Martinez-Fernandez, PhD*, Maria A. Missinato, PhD*, Jeanne L. Theis, PhD*, Georg Vogler, PhD*, Tanja Nielsen, BS, Stanley M. Walls, PhD, Katja Birker, BS, Jared M. Evans, MS, Megan M. O'Byrne, MA, Zachary C. Fogarty, Karen Ocorr, PhD, André Terzic, MD, PhD, Rolf Bodmer[#], PhD, Alexandre R. Colas[#], PhD, Timothy J. Nelson, MD, PhD[#], Timothy M. Olson, MD[#]

Development, Aging and Regeneration Program, Sanford-Burnham-Prebys Medical Discovery Institute, La Jolla, CA, USA (GV, MAM, PA, SS, KB, KO, ARC, RB), Department of Cardiovascular Medicine (AMF, AT, TMO), Cardiovascular Genetics Research Laboratory (JLT, TMO), Center for Regenerative Medicine (AT, TJN), Division of Biomedical Statistics and Informatics (XL, JME, MMO, ZCF), Division of Pediatric Cardiology, Department of Pediatric and Adolescent Medicine (TMO, TJN), Division of General Internal Medicine (TJN), Department of Molecular and Pharmacology and Experimental Therapeutics (AT, TMO, TJN), Mayo Clinic, Rochester, MN, USA.

* **Equal contributions**, listed in alphabetical order: Xing Li, Almudena Martinez-Fernandez, Maria Missinato, Jeanne Theis, Georg Vogler

[#] **Corresponding authors**, listed in alphabetical order**:** Rolf Bodmer, Alexandre R. Colas, Timothy J. Nelson, Timothy M. Olson

Total Word Count: 3338

Running title: Functional genomics for congenital heart disease

ABSTRACT

Background: Congenital heart diseases, such as hypoplastic left heart syndrome (HLHS), are considered to have complex genetic underpinnings that are poorly understood. Here, an integrated multi-disciplinary approach was applied to identify novel genes and underlying mechanisms associated with HLHS.

Methods: A family-based strategy was employed that coupled whole genome with RNA sequencing of patient-derived induced pluripotent stem cells (iPSCs) from an HLHS proband-parent trio to identify, prioritize and functionally evaluate candidate genes in model systems.

Results: Consistent with the hypoplastic phenotype, the proband's iPSCs had reduced proliferation capacity. Mendelian inheritance modeling identified 10 genes with recessive rare variants and altered expression compared to the parents' iPSCs. siRNA/RNAi-mediated knockdown in generic human iPSC-derived cardiac progenitors and in the *in vivo Drosophila* heart model revealed that LDL receptor related protein *LRP2* and apolipoprotein *APOB* are required for robust hiPSC-derived cardiomyocyte proliferation and *Drosophila* heart function, respectively, possibly involving an oligogenic mechanism via growth-promoting WNT and SHH signaling. Burden analysis of rare damaging variants in the 2 genes and 80 interacting partners in a cohort of 130 HLHS probands and 861 controls identified significant enrichment in *LRP2* ($p<0.001$), a gene associated with poor clinical outcomes in 30% of cases.

Conclusions: Collectively, this cross-functional genetic approach to complex congenital heart disease revealed disruption of LRP2 function as a novel genetic driver of HLHS, and hereby established a scalable approach to decipher the oligogenic underpinnings of maladaptive left heart development.

Key Words: lipoproteins, *Drosophila*, human iPSC, cardiogenesis, heart disease, cardiomyocytes.

Abbreviations

HLHS = hypoplastic left heart syndrome

CHD = congenital heart disease

WGS = whole genome sequencing

SHH = sonic hedgehog

WNT = wingless/integrated

INTRODUCTION

Hypoplastic left heart syndrome (HLHS) is a congenital heart disease (CHD) characterized by underdevelopment of the left ventricle, mitral and aortic valves, and aortic arch. Variable phenotypic manifestations and familial inheritance patterns, together with the numerous studies linking it to a diverse array of genes,[1-4] suggest that HLHS is genetically heterogeneous and may have significant environmental contributors. In this scenario, synergistic combinations of filtering and validation approaches are necessary to prioritize candidate genes and gene variants that may affect cardiogenic pathways throughout the dynamic process of human heart development.

Although the cellular mechanisms causing HLHS remain poorly characterized, a recent study reported generation of the first animal model of HLHS. Based on a digenic mechanism, mice deficient for HDAC-associated protein-encoding *Sap130* and protocadherin-coding *Pcdha9* exhibited left ventricular (LV) hypoplasia that was likely due – at least in part – to defective cardiomyocyte proliferation and differentiation, and increased cell death.[5] Similarly in humans, Gaber et al[6] provide evidence that HLHS-LV samples have more DNA damage and senescence with cell cycle arrest, and fewer cardiac progenitors and myocytes than controls. These observations suggest that impaired cardiomyocyte proliferation could be a mechanism contributing to HLHS pathogenesis, although pathogenic genes controlling this process in humans remain to be identifed and validated. Therefore, new synergistic experimental approaches that functionally validate cellular mechanisms of defective cardiogenesis are needed to probe the oligogenic hypothesis of left-sided heart defects, such as HLHS.[7, 8]

Over the last decade, induced pluripotent stem cells (iPSCs) have provided a revolutionary experimental tool to reveal aspects of the cellular manifestations associated with disease pathogenesis.[9-11] Progress in next-generation sequencing technology allows rapid whole genome DNA and RNA sequencing, thereby providing access to high-resolution and personalized genetic information. However, the interpretation of patient-specific sequence variants is often challenged by uncertainty in establishing a pathogenic link between biologically relevant variant(s) and a complex disease.[12]

Testing numerous potentially disease-causing genetic variants from patient sequence analysis in a mammalian *in vivo* model has been challenging because of high costs and low throughput. *Drosophila* with its genetic tools has emerged as the low cost, high throughput model of choice for human candidate disease gene testing, including neurological and cardiac diseases.[13-15] *Drosophila* has been established as an efficient model system to identify key genes and mechanisms critical for heart development and function that served as prototypes for mammalian studies, due to high degree of conservation of genetic pathways and reduced genetic complexity,[16] e.g the first cardiogenic transcription factor *Nkx2-5/tinman*[17, 18] that marked the beginning of a molecular-genetic understanding of cardiac specification.[19-21]

For this study, we combined whole-genome sequencing (WGS), iPSC technology and model system validation with a family-based approach to address potential HLHS disease mechanisms. We found that loss-of-function of 2 genes, *LRP2*, coding for a multi-ligand receptor, and *APOB*, a circulating lipoprotein ligand, resulted in reduced proliferation of hiPSCs and functional defects of the *Drosophila* heart. Furthermore, system-based analysis suggested underlying developmental (SHH and WNT) and growth-associated (TP53) pathways in HLHS. Finally, *LRP2* was enriched in variant burden analysis, suggesting a novel link between LRP2 in

cardiac development and the etiology of HLHS. Our integrated multidisciplinary high-throughput approach establishes a scalable and synergistic gene discovery platform in genetically complex human heart diseases.

METHODS

Study subjects

Written informed consent was obtained for the index family and an HLHS cohort, under a research protocol approved by the Mayo Clinic Institutional Review Board. Cardiac anatomy was assessed by echocardiography. Candidate genes were identified and prioritized by WGS of genomic DNA and RNA sequencing of patient-specific iPSC and cardiomyocytes. For variant burden analysis, controls were obtained from the Mayo Clinic Center for Individualized Medicine's Biobank. Methods for genomic analyses, RNA Sequencing, iPSC technology, bioinformatics and statistics are described in the Online Appendix. Data are available in NCBI SRA database (see below for SRA Accession IDs).

Drosophila heart studies

Drosophila orthologs were determined using the DIOPT database,[22] and RNAi lines were obtained from the Vienna Resource Drosophila Center (VDRC) stock center and crossed to Hand$^{4.2}$-Gal4 (heart-specific driver) alone or in combination with one copy of the *tinman* loss-of-function allele *tin*^{EC40}.[18] Fly hearts were filmed and analyzed according to standard protocols.[13]

hiPSC-CM, siRNA transfection, EdU assay, Immunostaining, qRT-PCR

hiPSC-derived cardiomyocytes were produced as previously described.[23, 24] Cardiomyocytes were plated in 384 wells and transfected with siRNAs (see Online Appendix). Two days after transfection, EdU was added to the media for 24 hours and fixed. EdU was detected using Click-it Plus EdU Imaging Kit (Life Technologies). For differentiation experiments, human cardiac mesoderm progenitors were fixed after 7days and stained with anti-α-Actinin, and anti-Transgelin antibodies. For qRT-PCR experiments, total RNA was extracted using TRIzol and chloroform. 1ug of RNA was converted in cDNA using QuantiTect Reverse Transcription kit (Qiagen). Human primer sequences for qRT-PCR were obtained from Harward Primer Bank. At least 3 independent biological replicates were performed for each experiment.

Statistical analysis

All statistical analysis for iPSC-derived cardiomyocytes were performed using GraphPad Prism version 7.0 (GraphPad Software, San Diego CA, USA). Statistical significance was analyzed by unpaired Student's *t*-test, and one-way ANOVA and shown as mean ± SEM. P-values were considered significant when < 0.05.

RESULTS

Transcriptome and cell cycle activity are altered in HLHS patient-derived iPSCs and CMs

This study analyzed a family quintet comprised of unrelated parents and their three offspring (Figure 1A). The male proband (II.3) was diagnosed with non-syndromic HLHS by physical examination and echocardiography, which demonstrated aortic and mitral valve atresia, virtual absence of the left ventricular cavity, and severe aortic arch hypoplasia. He was born

prematurely at 29 weeks gestation and underwent staged surgical palliation at 2 and 11 months of age. Conversion to a fenestrated Fontan circulation at 3 years of age failed owing to systolic and diastolic heart failure, necessitating early take-down. The patient subsequently died of multi-organ system failure. Echocardiography revealed structurally and functionally normal hearts in the proband's mother (I.2), father (I.1) and siblings (II.1 and II.2). Maternal history is notable for 4 miscarriages.

Patient-derived iPSCs are a valuable tool to investigate heart defects, such as those observed in HLHS.[3, 25] In this study, iPSCs from the mother (I.2), father (I.1) and HLHS proband (II.3) were generated [26] to investigate differences in transcriptional profiles potentially associated with HLHS, cells from the proband-parent trio were differentiated to day 25, using a cardiogenic differentiation protocol and processed for subsequent RNA sequencing (Figure 1B). In this *in vitro* cellular context, bioinformatic analysis revealed that 5,104 differentially expressed transcripts (DETs) in D25 differentiated samples between proband vs. mother/father (Online Table 1, Benjamini-corrected $p<0.001$). We found that 1,401 DETs were concordantly differentially expressed between proband and both parents (Figure 1C**,** Online Figure 1A-C, Online Table 1 and 2). Consistent with previous observations in HLHS fetuses,[6] KEGG analysis revealed TP53 pathway enrichment (Figure 1D), including cell cycle inhibition (Figure 1E), suggesting that cell proliferation may be affected in proband cells.

To test this hypothesis, we measured cell cycle activity in proband and parent hiPSC-derived cardiomyocytes (hiPSC-CMs) using an EdU-incorporation assay (Figure 1F; see Online Figure 4 for fibroblasts). Indeed, proband hiPSC-CMs exhibited reduced number of EdU-positive cells (Figure 1G-I). To evaluate TP53-dependency of cell cycle defects, we performed siRNA-mediated TP53 or CDKN1A knockdown (KD)in hiPSC-CMs. We observed significantly

elevated EdU incorporation in proband and parents (Online Figure 1D-M). Collectively, our findings suggest HLHS may involve TP53 pathway-dependent blockade of CM proliferation, consistent with observations made in both HLHS fetuses[6] and a HLHS mouse model.[5]

Family-based WGS, variant filtering, and transcriptional profiling identified 10 candidates

Array comparative genome hybridization ruled out a chromosomal deletion or duplication in the proband. WGS was performed to identify potentially pathogenic coding or regulatory single nucleotide variants (SNVs) or insertion/deletions (INDELs). Variants were filtered for rarity, predicted impact on protein strucuture or expression, and *de novo* or recessive inheritance, yielding 114 variants in 61 genes (Figure 2, Online Table 3). We next prioritized genes most likely to drive downstream pathways of dysregulated cardiogenesis in the HLHS proband by cross-referencing these candidate genes with 3,816 DETs identified in undifferentiated iPSC at day 0 (d0) (Online Table 4) and 5,104 DETs identified at day 25 (d25) differentiated cell lineages (Online Table 1). Ten genes harboring recessive variants were differentially expressed within the HLHS proband's iPSCs at d0 and d25: *HSPG2*, *APOB*, *LRP2*, *PRTG*, *SLC9A1*, *SDHD*, *JPT1*, *ELF4*, *HS6ST2* and *SIK1* (Figure 2).

Knockdown of candidate gene orthologs in *Drosophila* heart

In order to determine whether these variants occured within genes that could be important for cardiac differentiation *in vivo*, we took advantage of our established *Drosophila* heart development model and functional analysis tools (Figure 3A).[14] We hypothesized that genes critical for the *Drosophila* heart have conserved roles also in humans, as previously observed[19, 20, 27]. Predicted by DIOPT database[22] to have orthologs in *Drosophila* (Figure 3B), we analyzed 9

genes using heart-specific RNAi KD. By *in vivo* heart structure and function analysis,[13] we found that KD of *LRP2* (*mgl*), and *APOB* (*apolpp*) resulted in arrhythmias (Figure 3C,D, E; Online movies S1, S2, S3). KD of *JPT1 (CG1943)* resulted in dilation, suggesting a developmental or remodeling defect (Figure 3F).

Since HLHS is likely oligogenic,[28, 29] some genes in HLHS might only become apparent in combination with mutations in other cardiac-relevant genes. To test this, we examined the candidates also in the heterozygous background for *tinman/NKX2-5*, which in humans is well-known to contribute to a variety of CHD/HLHS manifestations.[1, 25, 30, 31] In this *in vivo* context, heart-specific knockdown of 2 out of 9 genes, namely HSPG2/Perlecan (*trol*), involved in extracellular matrix assembly,[32] and the SDHD (*SdhD*), exhibited a constricted phenotype (Figure 3G,H). These findings demonstrate that our bioinformatic candidate gene prioritization identified several candidates as cardiac relevant.

APOB and LRP2 regulate proliferation in human iPSC-derived cardiomyocytes

Decreased proliferation of left ventricular cardiomyocytes is a phenotypic hallmark of HLHS;[5, 6] see Figure 1G), suggesting that cell cycle impairment likely contributes to the disease. Thus, we asked whether siRNA-mediated candidate gene KD affects hiPSC-CM proliferation (Figure 4A).[23] Indeed, siRNAs directed against all 10 candidate genes together caused a marked reduction of EdU+ cardiomyocytes (ACTN1+) and cardiac nuclei numbers (Figure 4B-E), concomitant with downregulation of cell cycle genes (Figure 4F). In addition, proliferation inhibitors (*TP53*) and apoptosis markers were increased (Figure 4G). Subsequently, KD of the individual candidates revealed that siAPOB and siLRP2 were each able to recapitulate the reduced hiPSC-CM proliferation and increased apoptosis phenotype (Figure 4H-M).

In order to determine the impact of *APOB* and *LRP2* during the process of cardiogenesis, both genes were simultaneously knocked down in 5 day old hiPSC-derived cardiac progenitors and cardiac differentiation was assessed at day 12 (Figure 4N). Lineage quantification was performed using ACTN1 as a marker of cardiomyocytes, and TAGLN as a marker of fibroblasts. Notably, APOB and LRP2 KD reduced cardiomyocyte numbers and differentiation and increased fibroblast density (Figure 4O-Y). Collectively, these data identify *APOB* and *LRP2* as major regulators of cell cycle, apoptosis, and cardiogenesis in hiPSC-CMs, potentially contributing to the developmental cardiac impairment in HLHS patients.

Potential regulatory network of validated gene canditates

In order to delineate how potentially pathogenic mutations in the proband's six genes validated in model systems might affect signaling homeostasis in proband cells, we assembled a gene regulatory network containing *APOB*, *HS6ST2*, *HSPG2*, *JPT1*, *LRP2*, *SLC9A1* and their first neighbors (genetic and protein-protein interactions, using BioGRID data for *Drosophila melanogaster*) (Figure 5A; Online Table 5). Strikingly, the resulting network (82 genes; Online Figure 5) connected all 6 genes to WNT and SHH signaling cascades, both key regulators of cardiac differentiation and proliferation.[33, 34] Consistent with these findings, RNA-seq analysis of the family trio cells reveiled that negative regulator of SHH pathway, *PTCH1,* was upregulated, while agonists of WNT signaling pathway, *WNT1/3a/8a/10b* and *FZD10*,[35] were downregulated, suggesting these pathways are attenuated in the proband. In addition, *TP53* and *CDKN1A* expression was upregulated in response to siAPOB or siLRP2 (Figure 4), and notably also in proband cells (Figure 1). Importantly, siTP53/CDKN1A restored cell cycle activity in proband-

derived cardiomyocytes (Online Figure 1). Thus, a WNT/SHH/TP53-mediated impairment of cardiac differentiation and proliferation may contribute to HLHS in the proband **(**Figure 5A).

Rare variant analysis in HLHS cohort reveals enrichment in *LRP2*

To evaluate whether the proband's regulatory network of genes is generalizable to HLHS, we determined the frequency of rare, predicted-damaging variants in these genes in a cohort of 130 HLHS cases and 861 controls. Of the 82 genes, 65 had at least one rare variant present in cases or controls, yet *LRP2* was the only gene with significant enrichment in cases (Online Table 6). HLHS patients had a ~3-fold increase in the frequency of *LRP2* missense variants compared to healthy controls (10% versus 3.4%; p=0.0008). Moreover, among the 13 patients (including the index case) who carried a *LRP2* variant, 4 developed early heart failure (Figure 5B, Online Table 7). Comparison of *LRP2* variants identified in cases and controls did not reveal significant differences in their population-based frequencies, predicted-damaging scores, or presence in binding domains (data not shown). Of note, 13 of the 130 HLHS cases (including the index family proband) possessed <80% of ancestral Caucasian alleles, while all controls posessed ≥80%. Four of the 13 cases had rare, predicted-damaging missense variants in *LRP2* however all variants assessed were required to be rare in all racial populations. To eliminate the potentially confounding variable of race a Caucasian-only sub-analysis was performed, resulting in a less significant p-value for rare, predicted-damaging missense variants (7.7% versus 3.4%; p=0.05). However, removal of the predicted-damaging restriction on rare *LRP2* variants among Caucasians revealed significant enrichment in cases (p=0.0035), most notably in missense and intronic variants (p=0.0178 and 0.0082, respectively) Online Table 8). Location of variants within active histone marks or transcription factor binding sites was not

different between cases and controls. The biological significance of rare *LRP2* variants in HLHS and potential race-specific differences require further study.

DISCUSSION

Integrated Multidisciplinary Disease Gene Discovery and Modeling

Unraveling the molecular-genetic etiology of HLHS pathogenesis is crucial for (1) our ability to provide reliable prenatal diagnostics to families and (2) to develop novel approaches to to treat the disease. As an important step towards these goals, our integrated multidisciplinary approach identifies rare variants in *LRP2* as a molecular signature enriched in HLHS patients, consistent with our mechanistic analysis in model systems.

For this study, we used a powerful combination of high-throughput DNA/RNA patient sequencing coupled with high-throughput functional screening in model systems enabling to probe gene function (alone or in combination) on a wide array of cellular processes that are deployed during heart formation. For validation in model systems, we have established an integrated multi-site and multidisciplinary pipeline that systematically evaluates the functional role of genes presenting rare and deleterious variants in HLHS patients in hiPSC and *Drosophila* heart models. As a main objective – identify genes and potential mechanisms associated with HLHS - our pioneer study highlights rare, predicted-damaging *LRP2* missense variants as 3-fold enriched in 130 HLHS patients compared to 861 controls. Validation of this gene in hiPSC and *Drosophila* heart models demonstrates a requirement in cardiac proliferation, differentiation, and rhythm. Mutations in *LRP2* have been previously associated with cardiac defects in mouse[36] and in Donnai-Barrow Syndrome in humans.[36, 37] However, *LRP2* has not previously been linked to HLHS within curated bioinformatic networks.

One pre-requisite to reduce the knowledge gap between patient genomes and clinical phenotypes is to establish reliable/quantifiable phenotypic links between HLHS-associated genes and their role during normal cardiac development. Also, given that each large scale genomic study to identify CHD-associated genes can generate hundreds of candidates, we reasoned that our cardiac phenotypical platform should enable high throughput functional screening strategies to accomodate rapid testing of a large number of genes. Although overall heart structure in flies differs from that in vertebrates, the fundamental mechanisms of heart development and function are remarkably conserved, including a common transcriptional regulatory network,[19, 20] and a shared protein composition [38] as well as electrical and metabolic properties.[14, 39, 40] Using this 'convergent biology' approach identified APOB and LRP2 as hits in both the *in vitro* and *in vivo* cardiac model systems. Moreover, for further mechanistic understanding of complex CHD characterized by oligogenic etiologies this dual approach enables assessment of gene function combinatorially and in sensistized genetic backgrounds (e.g. *tinman*/NKX2-5).

5H pathogenic pathway: a potential role for SHH, WNT, p53 and cell proliferation in HLHS

Our current understanding of the molecular genetic causes of HLHS is very limited, despite clear genetic orgins of disease.[41] Past research on HLHS has yielded very few high-confidence gene candidates that may contribute to HLHS: *NOTCH1*, *NKX2-5* and *MYH6* have been implicated,[1, 3, 4] but they are also associated with other CHDs.

Heart development is a complex process that involves the interaction of many pathways and tissue interactions, and a large number of genes have been implicated in various types of CHDs.[42] The postulated oligogenic nature of HLHS likely is the result of an unfavorable

combination of disease genes, and such a combination of alleles in turn could affect several, successive steps of heart development. This makes it extremely difficult to model the disease by single gene mutations. Current hypotheses of the etiology of HLHS include changes in cell cycle progression of myocytes, as well as altered blood flow ("no flow – no grow") as a consequence of defects of valves or the outflow tract.

Interestingly, the only mouse HLHS model to date, a digenic mutant for sap130 and pcdha9,[5] has a penetrance of less than thirty percent, indicating a profound role for subtle differences between genetic backgrounds. This mouse model suggests a separate mechanism with pcdha9 affecting aortic growth, whereas sap130 can exert a more severe HLHS-like phenotype, which might reflect a modular etiology of HLHS that separates valve and ventricular defects.

The gene network analysis that we have conducted points to the possibility that several of the prioritized candidate genes identified in the index patient can have a modulatory impact on WNT and/or SHH via a diverse set of mechanisms.[33] There is evidence that the three of the candidates with fly heart phenotypes - Trol/HSPG2, Mgl/LRP2 and Apolpp/APOB - can alter WNT and SHH signaling,[43, 44] consistent with both pathways being attenuated in the proband iPSCs. We hypothesize that a collective of likely hypomorphic mutations affects heart development leading to HLHS. Impaired ventricular growth could in addition be caused by changes in the p53 pathway, and our analysis of iPSC-derived cardiomyocytes suggests that p53 indeed depends on APOB and LRP2 levels. Such a multi-hit model of HLHS caused by sub-threshold hypomorphic alleles represents an attractive explanation of the disease.

In summary, this integrated multidisciplinary strategy of functional genomics using patient-specific iPSC combined with *in vivo* and human cellular model systems of functional validation postulates novel genetic pathways and potential polygenic interactions underlying HLHS. These complementory experimental approaches enabled the deconvolution of the complex polygenic risk factors underlying HLHS genotype/phenotype and establish the novel groundwork for definitive mechanistic studies of interacting risk factors that contribute to arrested cardiac development and adverse outcomes. This scalable approach promises more efficient discovery of novel genes associated more broadly with HLHS as multiple HLHS families can now be multiplexed in future studies.

Study Limitations

HLHS candidate gene selection was based on in silico predictive algorithms to filter for functional coding and regulatory variants. Mendelian inheritance modeling, a strategy used to identify major-effect variants, did not include consideration of incompletely penetrant heterozygous alleles inherited from an unaffected parent. Differential gene expression, which was functionally validated as a powerful filter for candidate variant prioritization, excluded functional variants that do not alter gene expression. Not all human genes are conserved in *Drosophila*, but >70% of disease-causing human genes have a fly ortholog. While structural differences exist between hiPSC-CM, *Drosophila* heart and human newborn cardiomyocytes, our combinatorial approach allows to uncover gene networks and interactions that is not feasible in mammalian model systems.

Data availability

Sequencing data are deposited in the NCBI Sequence Read Archive (SRA) database with accession numbers: SRS1417684 (proband iPSCs), SRS1417685 (Dad iPSCs), SRS1417686 (Mom iPSCs), SRS1417695 (proband d25 differentiated cells), SRS1417696 (Dad d25 differentiated cells), SRS1417714 (Mom d25 differentiated cells).

Acknowledgements: The authors acknowledge support from the Todd and Karen Wanek Family Program for Hypoplastic Left Heart Syndrome and the Medical Genome Facility Sequencing Core and Biobank within the Mayo Clinic Center for Individualized Medicine, Rochester, MN. We gratefully acknowledge the patient and family who participated in this study. We thank Sean Spearing, Prashila Amatya, Marco Tamayo and Bosco Trinh for their excellent technical assistance.

Sources of Funding: This work was supported by the Todd and Karen Wanek Family Program for Hypoplastic Left Heart Syndrome, Mayo Clinic Foundation, Rochester, MN (SAN-233970 to R.B. and A.R.C). This work was also supported by National Institutes of Health (R01-HL054732 to RB).

Disclosures: None.

Affiliations: T.J.N and T.M.O: Pediatric Cardiology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905, USA. Phone: 507-538-4891, Fax 507-284-8432, Email: Nelson.Timothy@Mayo.edu

R.B. and A.R.C.: Development, Aging and Regeneration Program, Sanford-Burnham-Prebys Medical Discovery Institute, 10901 North Torrey Pines Road, La Jolla, CA92037, USA. Phone: 858-795-5295. Email: rolf@SBPdiscovery.org.

REFERENCES

1. Elliott DA, Kirk EP, Yeoh T, Chandar S, McKenzie F, Taylor P, Grossfeld P, Fatkin D, Jones O, Hayes P, Feneley M and Harvey RP. Cardiac homeobox gene NKX2-5 mutations and congenital heart disease: associations with atrial septal defect and hypoplastic left heart syndrome. *J Am Coll Cardiol*. 2003;41:2072-2076.

2. Lascone M, Ciccone R, Galletti L, Marchetti D, Seddio F, Lincesso AR, Pezzoli L, Vetro A, Barachetti D, Boni L, Federici D, Soto AM, Comas JV, Ferrazzi P and Zuffardi O. Identification of de novo mutations and rare variants in hypoplastic left heart syndrome. *Clin Genet*. 2012;81:542-554.

3. Theis JL, Hrstka SC, Evans JM, O'Byrne MM, de Andrade M, O'Leary PW, Nelson TJ and Olson TM. Compound heterozygous NOTCH1 mutations underlie impaired cardiogenesis in a patient with hypoplastic left heart syndrome. *Hum Genet*. 2015;134:1003-1011.

4. Theis JL, Zimmermann MT, Evans JM, Eckloff BW, Wieben ED, Qureshi MY, O'Leary PW and Olson TM. Recessive MYH6 mutations in hypoplastic left heart with reduced ejection fraction. *Circ Cardiovasc Genet*. 2015;8:564-571.

5. Liu X, Yagi H, Saeed S, Bais AS, Gabriel GC, Chen Z, Peterson KA, Li Y, Schwartz MC, Reynolds WT, Saydmohammed M, Gibbs B, Wu Y, Devine W, Chatterjee B, Klena NT, Kostka D, de Mesy Bentley KL, Ganapathiraju MK, Dexheimer P, Leatherbury L, Khalifa O, Bhagat A, Zahid M, Pu W, Watkins S, Grossfeld P, Murray SA, Porter Jr GA, Tsang M, Martin LJ, Benson DW, Aronow BJ and Lo CW. The complex genetics of hypoplastic left heart syndrome. *Nat Genet*. 2017;49:1152.

6. Gaber N, Gagliardi M, Patel P, Kinnear C, Zhang C, Chitayat D, Shannon P, Jaeggi E, Tabori U, Keller G and Mital S. Fetal reprogramming and senescence in hypoplastic left heart syndrome and in human pluripotent stem cells during cardiac differentiation. *Am J Pathol*. 2013;183:720-734.

7. Hinton RB, Martin LJ, Tabangin ME, Mazwi ML, Cripe LH and Benson DW. Hypoplastic Left Heart Syndrome is heritable. *J Am Coll Cardiol*. 2007;50:1590-1595.

8. McBride KL, Marengo L, Canfield M, Langlois P, Fixler D and Belmont JW. Epidemiology of noncomplex left ventricular outflow tract obstruction malformations (aortic valve stenosis, coarctation of the aorta, hypoplastic left heart syndrome) in Texas, 1999–2001. *Birth Defects Research Part A: Clinical and Molecular Teratology*. 2005;73:555-561.

9. Matsa E, Ahrens JH and Wu JC. Human induced pluripotent stem cells as a platform for personalized and precision cardiovascular medicine. *Physiol Rev*. 2016;96:1093-1126.

10. Mercola M, Colas A and Willems E. Induced pluripotent stem cells in cardiovascular drug discovery. *Circ Res*. 2013;112:534-548.

11. Moretti A, Laugwitz K-L, Dorn T, Sinnecker D and Mummery C. Pluripotent stem cell models of human heart disease. *Cold Spring Harb Perspect Med*. 2013;3.

12. Cooper GM and Shendure J. Needles in stacks of needles: finding disease-causal variants in a wealth of genomic data. *Nat Rev Genet*. 2011;12:628-640.

13. Fink M, Callol-Massot C, Chu A, Ruiz-Lozano P, Izpisua Belmonte JC, Giles W, Bodmer R and Ocorr K. A new method for detection and quantification of heartbeat parameters in Drosophila, zebrafish, and embryonic mouse hearts. *Biotechniques*. 2009;46:101-113.

14. Ocorr K, Vogler G and Bodmer R. Methods to assess Drosophila heart development, function and aging. *Methods*. 2014;68:265-272.

15. Şentürk M and Bellen HJ. Genetic strategies to tackle neurological diseases in fruit flies. *Current Opinion in Neurobiology*. 2018;50:24-32.

16. Bodmer R and Frasch M. *Development and aging of the Drosophila heart*. Amsterdam: Elsevier; 2010.

17. Azpiazu N and Frasch M. Tinman and bagpipe: two homeo box genes that determine cell fates in the dorsal mesoderm of Drosophila. *Genes Dev*. 1993;7:1325-1340.

18. Bodmer R. The gene tinman is required for specification of the heart and visceral muscles in Drosophila. *Development*. 1993;118:719-729.

19. Bodmer R. Heart development in Drosophila and its relationship to vertebrates. *Trends Cardiovasc Med*. 1995;5:21-28.

20. Cripps RM and Olson EN. Control of cardiac development by an evolutionarily conserved transcriptional network. *Dev Biol*. 2002;246:14-28.

21. Benson DW, Silberbach GM, Kavanaugh-McHugh A, Cottrill C, Zhang Y, Riggs S, Smalls O, Johnson MC, Watson MS, Seidman JG, Seidman CE, Plowden J and Kugler JD. Mutations in the cardiac transcription factor NKX2.5 affect diverse cardiac developmental pathways. *J Clin Invest*. 1999;104:1567-1573.

22. Hu Y, Flockhart I, Vinayagam A, Bergwitz C, Berger B, Perrimon N and Mohr SE. An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics*. 2011;12:357.

23. Cunningham TJ, Yu MS, McKeithan WL, Spiering S, Carrette F, Huang C-T, Bushway PJ, Tierney M, Albini S, Giacca M, Mano M, Puri PL, Sacco A, Ruiz-Lozano P, Riou J-F, Umbhauer M, Duester G, Mercola M and Colas AR. Id genes are essential for early heart formation. *Genes Dev*. 2017;31:1325-1338.

24. Yu MS, Spiering S and Colas AR. Generation of first heart field-like cardiac progenitors and ventricular-like cardiomyocytes from human pluripotent stem cells. *J Vis Exp*. 2018:e57688.

25. Hrstka SCL, Li X, Nelson TJ and Wanek Program Genetics Pipeline G. NOTCH1-dependent nitric oxide signaling deficiency in Hypoplastic Left Heart Syndrome revealed through patient-specific phenotypes detected in bioengineered cardiogenesis. *Stem Cells*. 2017;35:1106-1119.

26. Takahashi K and Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell*. 2006;126:663-676.

27. Qian L, Wythe JD, Liu J, Cartry J, Vogler G, Mohapatra B, Otway RT, Huang Y, King IN, Maillet M, Zheng Y, Crawley T, Taghli-Lamallem O, Semsarian C, Dunwoodie S, Winlaw D, Harvey RP, Fatkin D, Towbin JA, Molkentin JD, Srivastava D, Ocorr K, Bruneau BG and Bodmer R. Tinman/Nkx2-5 acts via miR-1 and upstream of Cdc42 to regulate heart function across species. *J Cell Biol*. 2011;193:1181-1196.

28. Blue GM, Kirk EP, Giannoulatou E, Sholler GF, Dunwoodie SL, Harvey RP and Winlaw DS. Advances in the genetics of congenital heart disease: A clinician's guide. *J Am Coll Cardiol*. 2017;69:859-870.

29. Gelb BD and Chung WK. Complex genetics and the etiology of human congenital heart disease. *Cold Spring Harb Perspect Med*. 2014;4.

30. Benson DW. Genetic origins of pediatric heart disease. *Pediatr Cardiol*. 2010;31:422-429.

31. Kobayashi J, Yoshida M, Tarui S, Hirata M, Nagai Y, Kasahara S, Naruse K, Ito H, Sano S and Oh H. Directed differentiation of patient-specific induced pluripotent stem cells identifies the transcriptional repression and epigenetic modification of NKX2-5, HAND1, and NOTCH1 in Hypoplastic Left Heart Syndrome. *PLoS One*. 2014;9:e102796.

32. Sasse P, Malan D, Fleischmann M, Roell W, Gustafsson E, Bostani T, Fan Y, Kolbe T, Breitbach M, Addicks K, Welz A, Brem G, Hescheler J, Aszodi A, Costell M, Bloch W and Fleischmann BK. Perlecan is critical for heart stability. *Cardiovasc Res*. 2008;80:435-444.

33. Briggs LE, Burns TA, Lockhart MM, Phelps AL, Van den Hoff MJB and Wessels A. Wnt/β-catenin and sonic hedgehog pathways interact in the regulation of the development of the dorsal mesenchymal protrusion. *Dev Dyn*. 2016;245:103-113.

34. Gessert S and Kühl M. The multiple phases and faces of Wnt signaling during cardiac differentiation and development. *Circ Res*. 2010;107:186-199.

35. Dawson K, Aflaki M and Nattel S. Role of the Wnt-Frizzled system in cardiac pathophysiology: a rapidly developing, poorly understood area with enormous potential. *J Physiol*. 2013;591:1409-1432.

36. Baardman ME, Zwier MV, Wisse LJ, Gittenberger-de Groot AC, Kerstjens-Frederikse WS, Hofstra RM, Jurdzinski A, Hierck BP, Jongbloed MR, Berger RM, Plosch T and DeRuiter MC. Common arterial trunk and ventricular non-compaction in Lrp2 knockout mice indicate a crucial role of LRP2 in cardiac development. *Dis Model Mech*. 2016;9:413-425.

37. Kantarci S, Ragge NK, Thomas NS, Robinson DO, Noonan KM, Russell MK, Donnai D, Raymond FL, Walsh CA, Donahoe PK and Pober BR. Donnai-Barrow syndrome (DBS/FOAR) in a child with a homozygous LRP2 mutation due to complete chromosome 2 paternal isodisomy. *Am J Med Genet A*. 2008;146A:1842-1847.

38. Cammarato A, Ahrens CH, Alayari NN, Qeli E, Rucker J, Reedy MC, Zmasek CM, Gucek M, Cole RN, Van Eyk JE, Bodmer R, O'Rourke B, Bernstein SI and Foster DB. A mighty small heart: the cardiac proteome of adult Drosophila melanogaster. *PLoS One*. 2011;6:e18497.

39. Ocorr K, Reeves NL, Wessells RJ, Fink M, Chen HS, Akasaka T, Yasuda S, Metzger JM, Giles W, Posakony JW and Bodmer R. KCNQ potassium channel mutations cause cardiac arrhythmias in Drosophila that mimic the effects of aging. *Proc Natl Acad Sci U S A*. 2007;104:3943-8.

40. Diop SB and Bodmer R. Gaining insights into diabetic cardiomyopathy from Drosophila. *Trends Endocrinol Metab*. 2015;26:618-627.

41. Yagi H, Liu X, Gabriel GC, Wu Y, Peterson K, Murray SA, Aronow BJ, Martin LJ, Benson DW and Lo CW. The genetic landscape of Hypoplastic Left Heart Syndrome. *Pediatr Cardiol*. 2018;39:1069-1081.

42. Pierpont Mary E, Brueckner M, Chung Wendy K, Garg V, Lacro Ronald V, McGuire Amy L, Mital S, Priest James R, Pu William T, Roberts A, Ware Stephanie M, Gelb Bruce D, Russell Mark W and null n. Genetic basis for congenital heart disease: Revisited: A scientific statement from the American Heart Association. *Circulation*. 2018;138:e653-e711.

43. Christ A, Christa A, Klippert J, Eule JC, Bachmann S, Wallace Valerie A, Hammes A and Willnow Thomas E. LRP2 acts as SHH clearance receptor to protect the retinal margin from mitogenic stimuli. *Dev Cell*. 2015;35:36-48.

44. Datta MW, Hernandez AM, Schlicht MJ, Kahler AJ, DeGueme AM, Dhir R, Shah RB, Farach-Carson C, Barrett A and Datta S. Perlecan, a candidate gene for the CAPB locus, regulates prostate cancer cell growth via the Sonic Hedgehog pathway. *Molecular Cancer*. 2006;5:9.

Figure Legend 1. Family-based iPSC characterization for HLHS. (A) Pedigree of family 5H: proband with HLHS (black symbol), relatives without CHD (white symbols), miscarriages (gray diamonds). (B) Schematic for family-based iPSC production and characterization. (C) Whole genome RNA sequencing identifies 1401 concordantly DETs between proband and parents. (D) KEGG pathway analysis shows enrichment of DETs in TP53 pathway. (E) Heatmap of p53 signaling pathway associated genes in probands vs parents. (F) Schematic describing EdU incorporation assay in hiPSC-CMs. (G) Graph representing quantification of EdU+ cardiomyocytes in HLHS family-derived iPSC-CMs. ***$p<0.001$ One-way ANOVA. (H,I) Representative images of iPSC-CMs derived from mother (H) and proband (I), stained for EdU, ACTN1 and DAPI. Scale bar: 50 µm.

Figure Legend 2. Identification of HLHS candidate genes from whole genome and RNA sequencing. An iterative, family-based variant filtering approach based on rarity, functional impact, and mode of inheritance yielded 61 candidate genes. RNA sequencing data from d0 iPSC and d25 differentiated iPSC were used to filter for transcriptional differences yielding 10 candidate genes.

Figure Legend 3. Phenotypic assessment of HLHS candidate genes in *Drosophila* adult hearts. (**a**) Schematic of the *Drosophila* adult heart assay. (**b**) Human candidate genes and corresponding *Drosophila* ortholog as determined by DIOPT score (*confidence score: number of databases reporting orthology). Listed are heart phenotypes upon knockdown in wildtype or *NKX2*-5/*tin*$^{+/-}$ heterozygous background. (**c-e**) RNAi-induced arrhythmicity and M-modes observed with *APOB* and *LRP2* knockdown. (**f-h**) Heart size (EDD: enddiastolic diameter) alterations upon RNAi-knockdown of *JPT1*, *HSPG2* and *SDHD* (also in NKX2-5/tin heterozygous background). Wilcoxon rank sum test: *$p<0.05$, **$p<0.01$, **$p<0.001$.

Figure Legend 4. HLHS candidate gene analysis in hiPSC-derived cardiomyocytes and progenitors

(A) Schematic of the experimental strategy. (B,C) Quantification of EdU incorporation and number of cardiac nuclei in day 28 hiPSC-CMs. (D,E) Representative picture of ACTN1, DAPI and EdU staining of hiPSC-CM transfected with siRNA directed against the ten candidate genes (si5H). (F) qRT-PCR analysis after si5H transfection. CCNE1: cyclin E1; PCNA: proliferating cell nuclear antigen; CCNB1: cyclin B1; CCNB2: cyclin B2; CDK1: cyclin-dependent kinase 1. (G) qRT-PCR analysis of cell death markers genes in si5H hiPSC-CM. CRADD: death domain receptor; TP53: tumor protein p53; CASP6: Caspase 6. (H) Deconvolution of "5H" genes in EdU incorporation with hiPSC-CMs. One-way ANOVA: **$p<0.01$. (I-K) Representative ACTN1 and EdU staining upon siAPOB and siLRP2 transfection. (L,M) qRT-PCR analysis of cell cycle, tumor suppressor and cell cycle inhibitor genes upon siAPOB/siLRP2 transfection. (N) Schematic of differentiation potential of cardiac mesoderm progenitors. (O-Q) Quantification of ACTN1 (cardiomyocytes), TAGLN (fibroblasts) and DAPI (nuclei). (R-Y) Representative images of day12 human cardiac mesoderm progenitors after siAPOB, siLRP2 and siControl transfection, immunostained for DAPI, ACTN1 and TAGLN. All experiments were in biological triplicates. Scale bar: 50µm. Student's T-test (except H): *$p<0.05$, **$p<0.01$, ***$p<0.001$,****$p<0.0001$.

Figure Legend 5. Potential role for SHH, WNT and LRP2 in HLHS. (**a**) A gene network integrating family-centric HLHS candidate genes with heart development. ORANGE – genes with cardiac phenotypes in iPSC/*Drosophila* assays. YELLOW – other candidate genes tested. RED – Genes up-regulated in proband iPSCs vs. parents. BLUE – Genes downregulated down-regulated in proband iPSCs vs. parents. (**b**) Cohort-wide analysis of LRP2 variants shows significant enrichment for SNVs in HLHS patients compared to control populations. Variants (blue / magenta) are found throughout LRP2 protein.

Figure 1

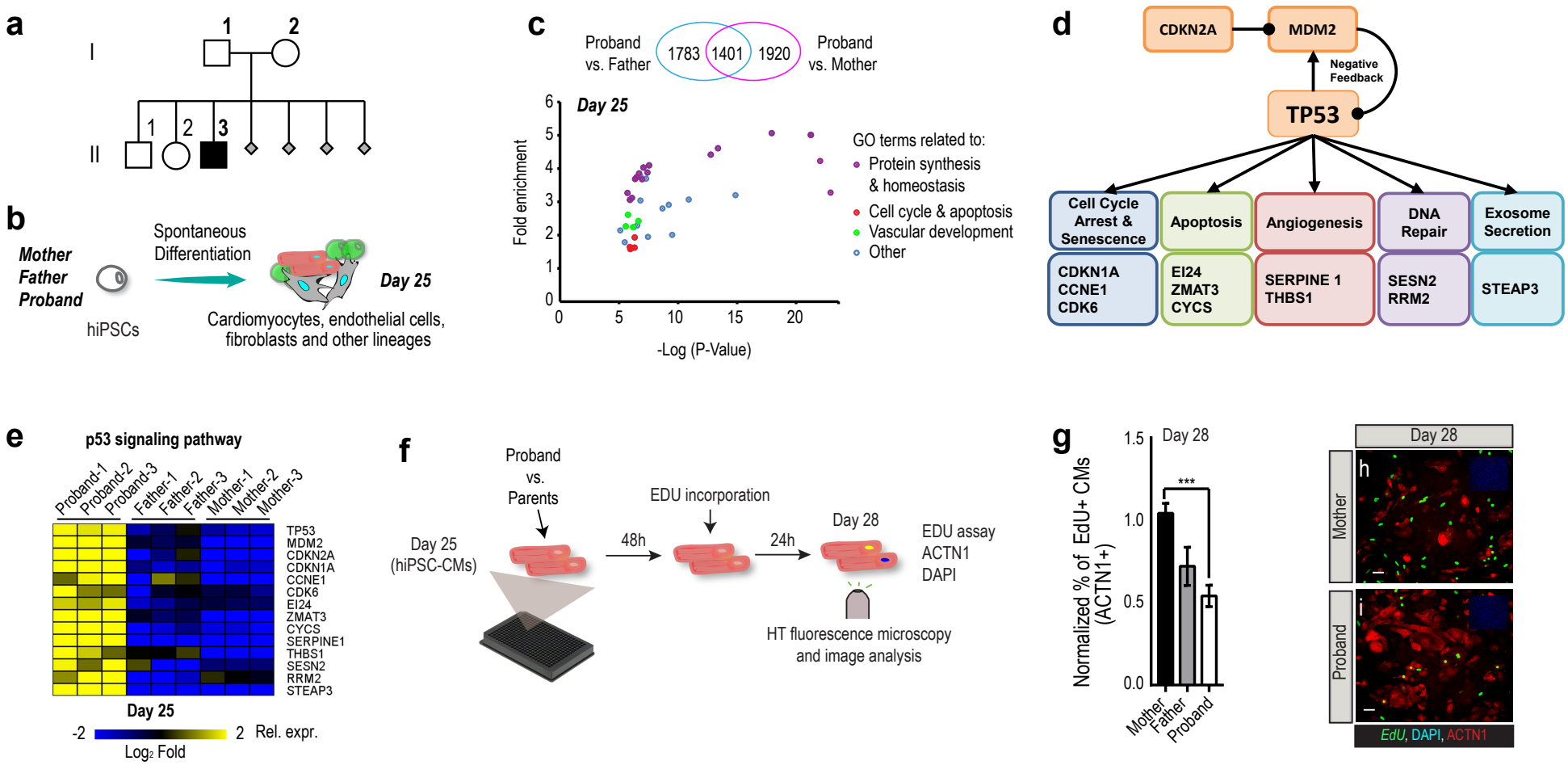

a
I
II
1
2
1
2
3
b
Mother
Father
Proband
hiPSCs
Spontaneous Differentiation
Day 25
Cardiomyocytes, endothelial cells, fibroblasts and other lineages
c
Proband vs. Father
1783
1401
1920
Proband vs. Mother
Day 25
Fold enrichment
-Log (P-Value)
GO terms related to:
Protein synthesis & homeostasis
Cell cycle & apoptosis
Vascular development
Other
d
CDKN2A
MDM2
Negative Feedback
TP53
Cell Cycle Arrest & Senescence
Apoptosis
Angiogenesis
DNA Repair
Exosome Secretion
CDKN1A
CCNE1
CDK6
EI24
ZMAT3
CYCS
SERPINE 1
THBS1
SESN2
RRM2
STEAP3
e
p53 signaling pathway
Proband-1
Proband-2
Proband-3
Father-1
Father-2
Father-3
Mother-1
Mother-2
Mother-3
TP53
MDM2
CDKN2A
CDKN1A
CCNE1
CDK6
EI24
ZMAT3
CYCS
SERPINE1
THBS1
SESN2
RRM2
STEAP3
Day 25
-2
2
Rel. expr.
Log2 Fold
f
Proband vs. Parents
Day 25 (hiPSC-CMs)
48h
EDU incorporation
24h
Day 28
EDU assay
ACTN1
DAPI
HT fluorescence microscopy and image analysis
g
Day 28
Normalized % of EdU+ CMs (ACTN1+)
***
Mother
Father
Proband
Day 28
Mother
Proband
h
i
EdU, DAPI, ACTN1

Figure 2

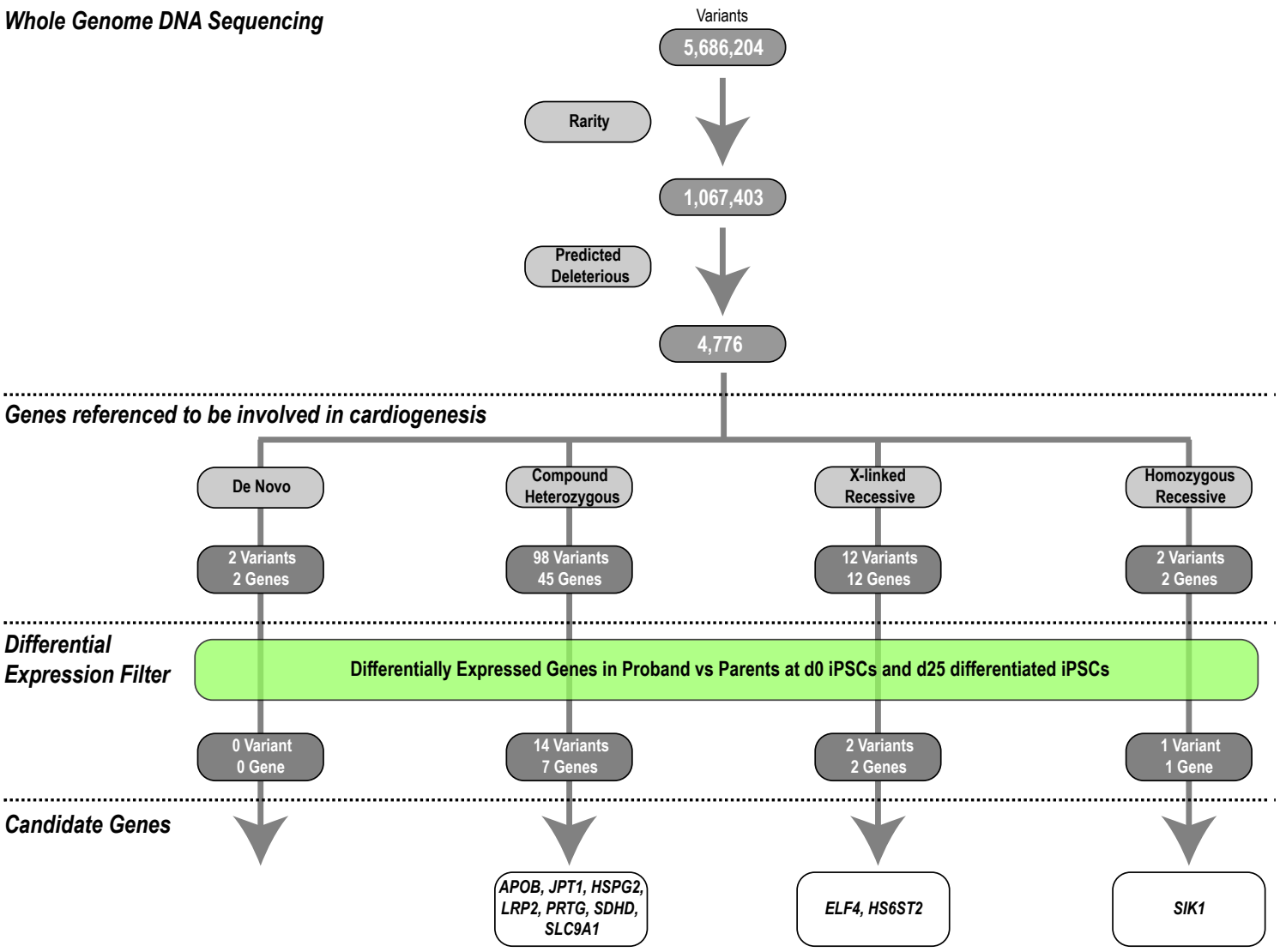
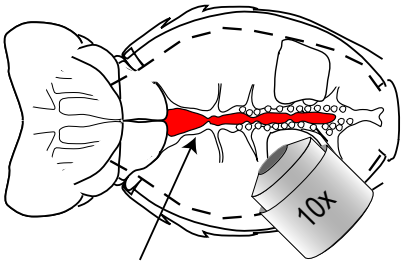

Whole Genome DNA Sequencing
Variants
5,686,204
Rarity
1,067,403
Predicted Deleterious
4,776
Genes referenced to be involved in cardiogenesis
De Novo
Compound Heterozygous
X-linked Recessive
Homozygous Recessive
2 Variants 2 Genes
98 Variants 45 Genes
12 Variants 12 Genes
2 Variants 2 Genes
Differential Expression Filter
Differentially Expressed Genes in Proband vs Parents at d0 iPSCs and d25 differentiated iPSCs
0 Variant 0 Gene
14 Variants 7 Genes
2 Variants 2 Genes
1 Variant 1 Gene
Candidate Genes
APOB, JPT1, HSPG2, LRP2, PRTG, SDHD, SLC9A1
ELF4, HS6ST2
SIK1

Figure 3

a

Loss of function strategy:
Heart-specific (*Hand4.2 driver*) RNAi
of 5H genes in post-mitotic CMs



3 week old female
Drosophila melanogaster

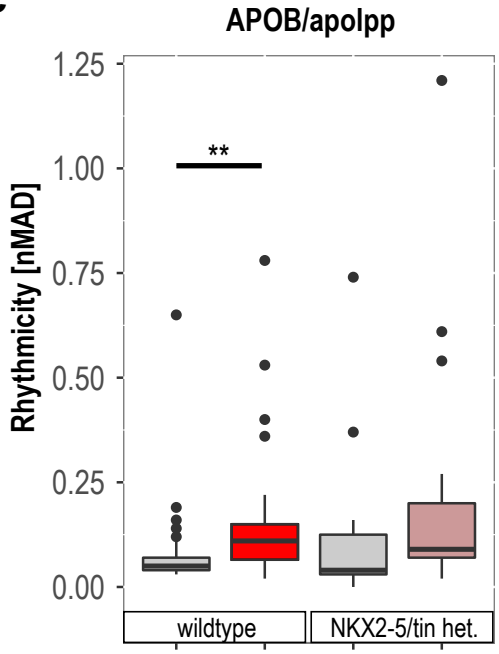
Kinetic imaging (30s@140fps) followed by automated and quantitative assessment of heart size and function

b

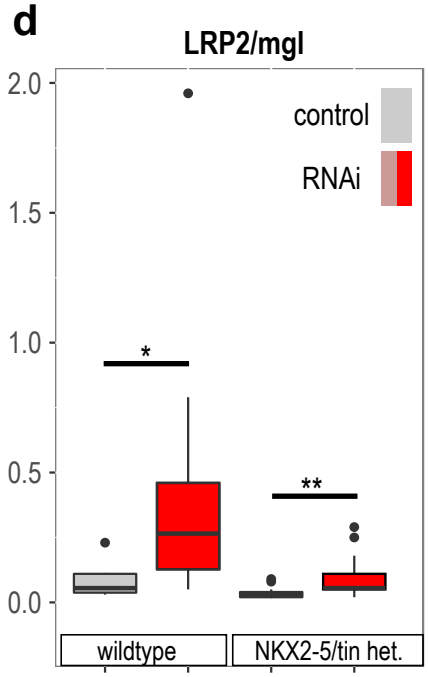
human gene	DIOPT* score	fly gene	heart phenotypes	NKX2-5 / tin interaction phenotype
HSPG2	9	*trol*	-/-	constriction
APOB	6	*apolpp*	arrhythmic	-/-
LRP2	11	*mgl*	arrhythmic	-/-
SLC9A1	5	*Nhe2*	-/-	-/-
SDHD	12	*SdhD*	-/-	constriction
JPT1	3	*Jupiter*	-/-	-/-
JPT1	2	*CG1943*	dilation	-/-
ELF4	5	*Eip74EF*	-/-	-/-
HS6ST2	9	*Hs6st*	-/-	-/-
SIK1	3	*Sik2*	-/-	-/-

Wilcoxon rank sum test, $p < 0.05$

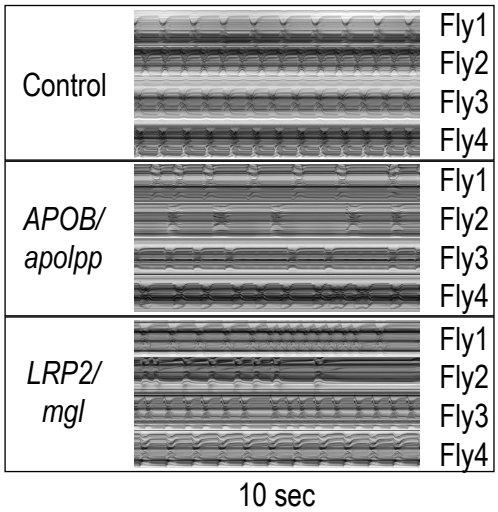
c



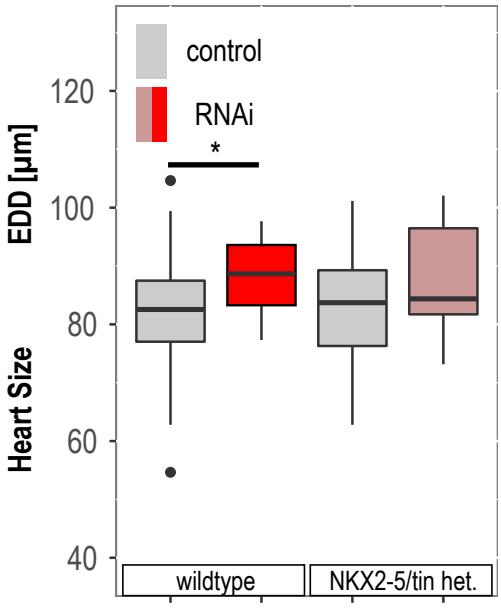
d



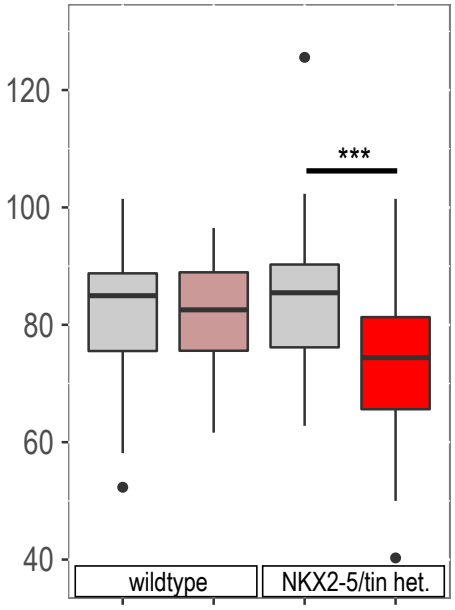
e



f



g



h

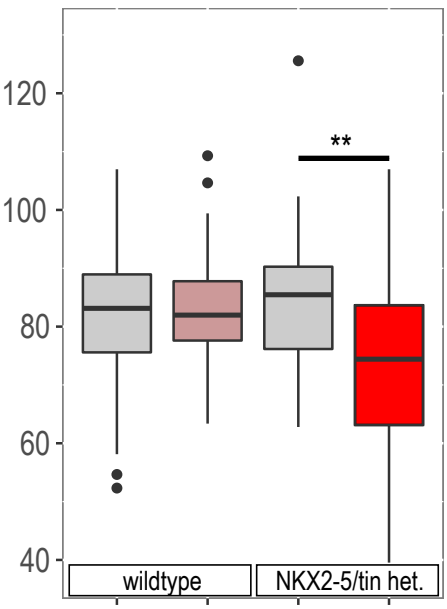


Figure 4

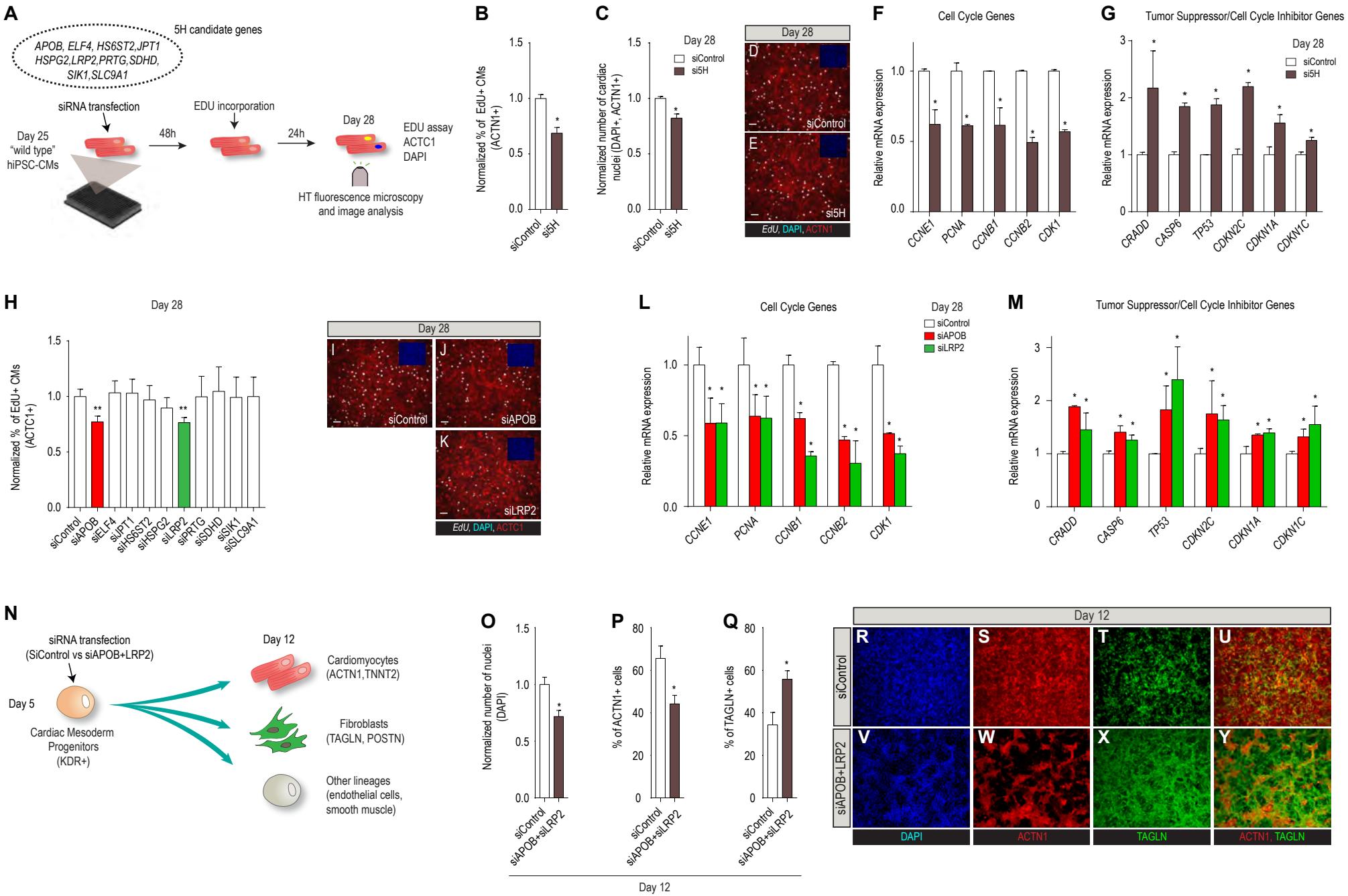
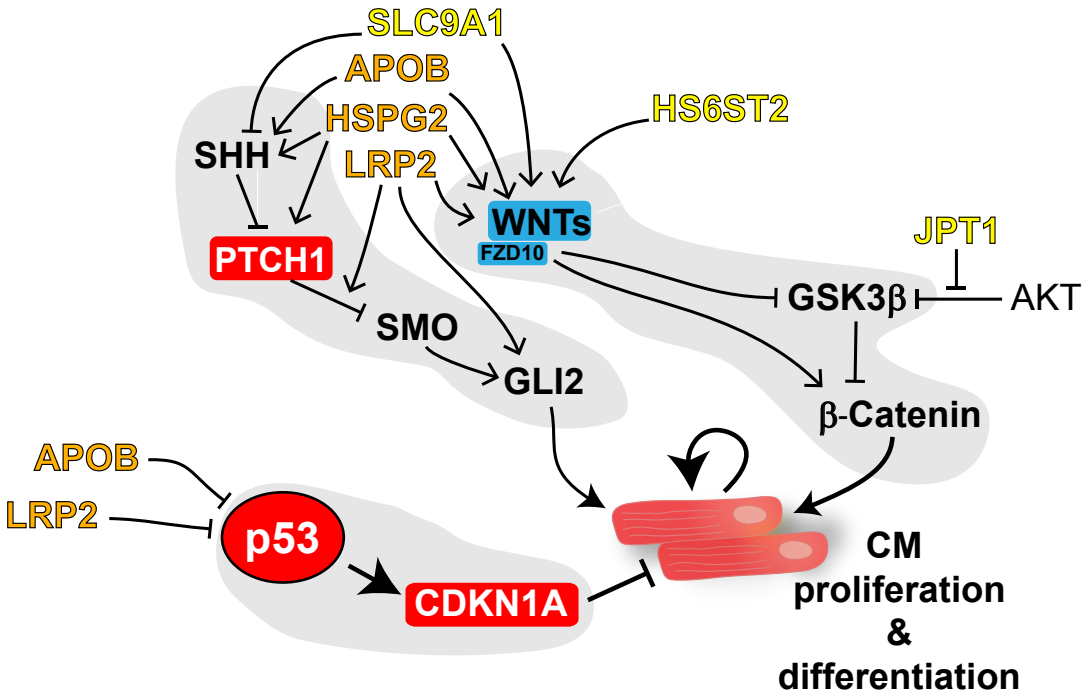
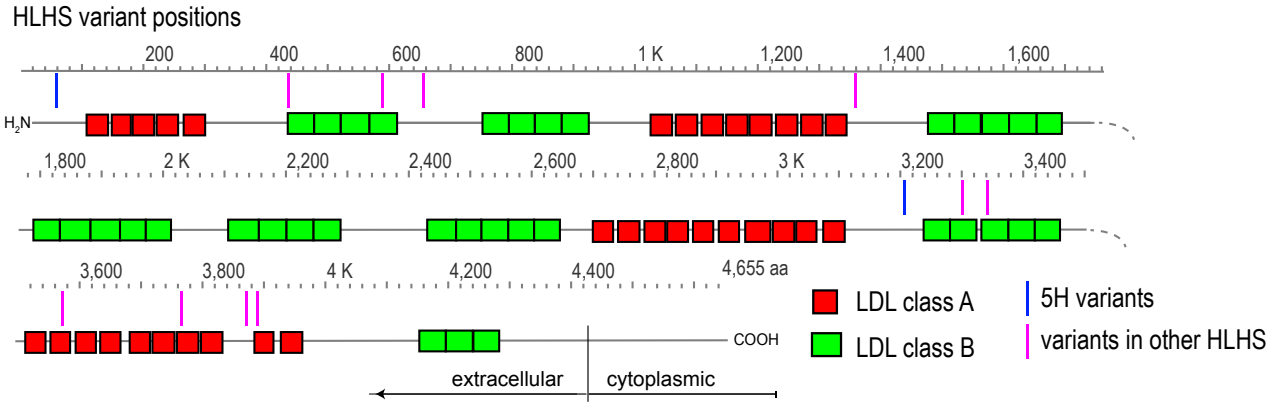

A
5H candidate genes
APOB, ELF4, HS6ST2,JPT1
HSPG2,LRP2,PRTG,SDHD,
SIK1,SLC9A1
siRNA transfection
EDU incorporation
Day 25
"wild type"
hiPSC-CMs
48h
24h
Day 28
EDU assay
ACTC1
DAPI
HT fluorescence microscopy
and image analysis
B
Normalized % of EdU+ CMs (ACTN1+)
siControl
si5H
C
Day 28
Normalized number of cardiac nuclei (DAPI+, ACTN1+)
siControl
si5H
D
E
siControl
si5H
EdU, DAPI, ACTN1
F
Cell Cycle Genes
Relative mRNA expression
CCNE1
PCNA
CCNB1
CCNB2
CDK1
G
Tumor Suppressor/Cell Cycle Inhibitor Genes
Day 28
siControl
si5H
Relative mRNA expression
CRADD
CASP6
TP53
CDKN2C
CDKN1A
CDKN1C
H
Day 28
Normalized % of EdU+ CMs (ACTC1+)
siControl
siAPOB
siELF4
siJPT1
siHS6ST2
siHSPG2
siLRP2
siPRTG
siSDHD
siSIK1
siSLC9A1
I
J
K
siControl
siAPOB
siLRP2
EdU, DAPI, ACTC1
L
Cell Cycle Genes
Day 28
siControl
siAPOB
siLRP2
Relative mRNA expression
CCNE1
PCNA
CCNB1
CCNB2
CDK1
M
Tumor Suppressor/Cell Cycle Inhibitor Genes
Relative mRNA expression
CRADD
CASP6
TP53
CDKN2C
CDKN1A
CDKN1C
N
siRNA transfection
(SiControl vs siAPOB+LRP2)
Day 5
Cardiac Mesoderm
Progenitors
(KDR+)
Day 12
Cardiomyocytes
(ACTN1,TNNT2)
Fibroblasts
(TAGLN, POSTN)
Other lineages
(endothelial cells,
smooth muscle)
O
Normalized number of nuclei (DAPI)
siControl
siAPOB+siLRP2
P
% of ACTN1+ cells
siControl
siAPOB+siLRP2
Q
% of TAGLN+ cells
siControl
siAPOB+siLRP2
Day 12
Day 12
siControl
siAPOB+LRP2
R
S
T
U
V
W
X
Y
DAPI
ACTN1
TAGLN
ACTN1, TAGLN

Figure 5

A



B

		HLHS (n=130)		Controls (n=861)		
		Total SNVs	*Total Subjects*	*Total SNVs*	*Total Subjects*	*P-value*
LRP2	*Missense with CADD > 24*	*14*	*13 (10.0%)*	*29*	*29 (3.4%)*	*0.0008*