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

1   **ABSTRACT**

2   Background: Congenital heart diseases, such as hypoplastic left heart syndrome (HLHS), are

3   considered to have complex genetic underpinnings that are poorly understood. Here, an

4   integrated multi-disciplinary approach was applied to identify novel genes and underlying

5   mechanisms associated with HLHS.

6   Methods: A family-based strategy was employed that coupled whole genome with RNA

7   sequencing of patient-derived induced pluripotent stem cells (iPSCs) from an HLHS proband-

8   parent trio to identify, prioritize and functionally evaluate candidate genes in model systems.

9   Results: Consistent with the hypoplastic phenotype, the proband's iPSCs had reduced

10 proliferation capacity. Mendelian inheritance modeling identified 10 genes with recessive rare

11 variants and altered expression compared to the parents' iPSCs. siRNA/RNAi-mediated

12 knockdown in generic human iPSC-derived cardiac progenitors and in the *in vivo Drosophila*

13 heart model revealed that LDL receptor related protein *LRP2* and apolipoprotein *APOB* are

14 required for robust hiPSC-derived cardiomyocyte proliferation and *Drosophila* heart function,

15 respectively, possibly involving an oligogenic mechanism via growth-promoting WNT and SHH

16 signaling. Burden analysis of rare damaging variants in the 2 genes and 80 interacting partners in

17 a cohort of 130 HLHS probands and 861 controls identified significant enrichment in *LRP2*

18 ($p<0.001$), a gene associated with poor clinical outcomes in 30% of cases.

19 Conclusions: Collectively, this cross-functional genetic approach to complex congenital heart

20 disease revealed disruption of LRP2 function as a novel genetic driver of HLHS, and hereby

21 established a scalable approach to decipher the oligogenic underpinnings of maladaptive left

22 heart development.

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

2   cardiomyocytes.

3

4   **Abbreviations**

5   HLHS = hypoplastic left heart syndrome

6   CHD = congenital heart disease

7   WGS = whole genome sequencing

8   SHH = sonic hedgehog

9   WNT = wingless/integrated

1   **INTRODUCTION**

2   Hypoplastic left heart syndrome (HLHS) is a congenital heart disease (CHD)

3   characterized by underdevelopment of the left ventricle, mitral and aortic valves, and aortic arch.

4   Variable phenotypic manifestations and familial inheritance patterns, together with the numerous

5   studies linking it to a diverse array of genes,[1-4] suggest that HLHS is genetically heterogeneous

6   and may have significant environmental contributors. In this scenario, synergistic combinations

7   of filtering and validation approaches are necessary to prioritize candidate genes and gene

8   variants that may affect cardiogenic pathways throughout the dynamic process of human heart

9   development.

10   Although the cellular mechanisms causing HLHS remain poorly characterized, a recent

11   study reported generation of the first animal model of HLHS. Based on a digenic mechanism,

12   mice deficient for HDAC-associated protein-encoding *Sap130* and protocadherin-coding *Pcdha9*

13   exhibited left ventricular (LV) hypoplasia that was likely due – at least in part – to defective

14   cardiomyocyte proliferation and differentiation, and increased cell death.[5] Similarly in humans,

15   Gaber et al[6] provide evidence that HLHS-LV samples have more DNA damage and senescence

16   with cell cycle arrest, and fewer cardiac progenitors and myocytes than controls. These

17   observations suggest that impaired cardiomyocyte proliferation could be a mechanism

18   contributing to HLHS pathogenesis, although pathogenic genes controlling this process in

19   humans remain to be identifed and validated. Therefore, new synergistic experimental

20   approaches that functionally validate cellular mechanisms of defective cardiogenesis are needed

21   to probe the oligogenic hypothesis of left-sided heart defects, such as HLHS.[7, 8]

4

1 Over the last decade, induced pluripotent stem cells (iPSCs) have provided a

2 revolutionary experimental tool to reveal aspects of the cellular manifestations associated with

3 disease pathogenesis.[9-11] Progress in next-generation sequencing technology allows rapid whole

4 genome DNA and RNA sequencing, thereby providing access to high-resolution and

5 personalized genetic information. However, the interpretation of patient-specific sequence

6 variants is often challenged by uncertainty in establishing a pathogenic link between biologically

7 relevant variant(s) and a complex disease.[12]

8 Testing numerous potentially disease-causing genetic variants from patient sequence

9 analysis in a mammalian *in vivo* model has been challenging because of high costs and low

10 throughput. *Drosophila* with its genetic tools has emerged as the low cost, high throughput

11 model of choice for human candidate disease gene testing, including neurological and cardiac

12 diseases.[13-15] *Drosophila* has been established as an efficient model system to identify key genes

13 and mechanisms critical for heart development and function that served as prototypes for

14 mammalian studies, due to high degree of conservation of genetic pathways and reduced genetic

15 complexity,[16] e.g the first cardiogenic transcription factor *Nkx2-5/tinman*[17, 18] that marked the

16 beginning of a molecular-genetic understanding of cardiac specification.[19-21]

17 For this study, we combined whole-genome sequencing (WGS), iPSC technology and

18 model system validation with a family-based approach to address potential HLHS disease

19 mechanisms. We found that loss-of-function of 2 genes, *LRP2*, coding for a multi-ligand

20 receptor, and *APOB*, a circulating lipoprotein ligand, resulted in reduced proliferation of hiPSCs

21 and functional defects of the *Drosophila* heart. Furthermore, system-based analysis suggested

22 underlying developmental (SHH and WNT) and growth-associated (TP53) pathways in HLHS.

23 Finally, *LRP2* was enriched in variant burden analysis, suggesting a novel link between LRP2 in

5

1 cardiac development and the etiology of HLHS. Our integrated multidisciplinary high-

2 throughput approach establishes a scalable and synergistic gene discovery platform in genetically

3 complex human heart diseases.

4 **METHODS**

5 **Study subjects**

6 Written informed consent was obtained for the index family and an HLHS cohort, under

7 a research protocol approved by the Mayo Clinic Institutional Review Board. Cardiac anatomy

8 was assessed by echocardiography. Candidate genes were identified and prioritized by WGS of

9 genomic DNA and RNA sequencing of patient-specific iPSC and cardiomyocytes. For variant

10 burden analysis, controls were obtained from the Mayo Clinic Center for Individualized

11 Medicine's Biobank. Methods for genomic analyses, RNA Sequencing, iPSC technology,

12 bioinformatics and statistics are described in the Online Appendix. Data are available in NCBI

13 SRA database (see below for SRA Accession IDs).

14

15 *Drosophila* **heart studies**

16 *Drosophila* orthologs were determined using the DIOPT database,[22] and RNAi lines

17 were obtained from the Vienna Resource Drosophila Center (VDRC) stock center and crossed to

18 Hand[4.2]-Gal4 (heart-specific driver) alone or in combination with one copy of the *tinman* loss-of-

19 function allele *tin*[EC40].[18] Fly hearts were filmed and analyzed according to standard protocols.[13]

20

21

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

2   hiPSC-derived cardiomyocytes were produced as previously described.[23, 24]

3   Cardiomyocytes were plated in 384 wells and transfected with siRNAs (see Online Appendix).

4   Two days after transfection, EdU was added to the media for 24 hours and fixed. EdU was

5   detected using Click-it Plus EdU Imaging Kit (Life Technologies). For differentiation

6   experiments, human cardiac mesoderm progenitors were fixed after 7days and stained with anti-

7   α-Actinin, and anti-Transgelin antibodies. For qRT-PCR experiments, total RNA was extracted

8   using TRIzol and chloroform. 1ug of RNA was converted in cDNA using QuantiTect Reverse

9   Transcription kit (Qiagen). Human primer sequences for qRT-PCR were obtained from Harward

10   Primer Bank. At least 3 independent biological replicates were performed for each experiment.

11

12   **Statistical analysis**

13   All statistical analysis for iPSC-derived cardiomyocytes were performed using GraphPad Prism

14   version 7.0 (GraphPad Software, San Diego CA, USA). Statistical significance was analyzed by

15   unpaired Student's *t*-test, and one-way ANOVA and shown as mean ± SEM. P-values were

16   considered significant when < 0.05.

17

18   **RESULTS**

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

20   This study analyzed a family quintet comprised of unrelated parents and their three

21   offspring (Figure 1A). The male proband (II.3) was diagnosed with non-syndromic HLHS by

22   physical examination and echocardiography, which demonstrated aortic and mitral valve atresia,

23   virtual absence of the left ventricular cavity, and severe aortic arch hypoplasia. He was born

1   prematurely at 29 weeks gestation and underwent staged surgical palliation at 2 and 11 months

2   of age. Conversion to a fenestrated Fontan circulation at 3 years of age failed owing to systolic

3   and diastolic heart failure, necessitating early take-down. The patient subsequently died of multi-

4   organ system failure. Echocardiography revealed structurally and functionally normal hearts in

5   the proband's mother (I.2), father (I.1) and siblings (II.1 and II.2). Maternal history is notable for

6   4 miscarriages.

7   Patient-derived iPSCs are a valuable tool to investigate heart defects, such as those

8   observed in HLHS.[3, 25] In this study, iPSCs from the mother (I.2), father (I.1) and HLHS proband

9   (II.3) were generated [26] to investigate differences in transcriptional profiles potentially associated

10   with HLHS, cells from the proband-parent trio were differentiated to day 25, using a cardiogenic

11   differentiation protocol and processed for subsequent RNA sequencing (Figure 1B). In this *in*

12   *vitro* cellular context, bioinformatic analysis revealed that 5,104 differentially expressed

13   transcripts (DETs) in D25 differentiated samples between proband vs. mother/father (Online

14   Table 1, Benjamini-corrected $p<0.001$). We found that 1,401 DETs were concordantly

15   differentially expressed between proband and both parents (Figure 1C, Online Figure 1A-C,

16   Online Table 1 and 2). Consistent with previous observations in HLHS fetuses,[6] KEGG analysis

17   revealed TP53 pathway enrichment (Figure 1D), including cell cycle inhibition (Figure 1E),

18   suggesting that cell proliferation may be affected in proband cells.

19   To test this hypothesis, we measured cell cycle activity in proband and parent hiPSC-

20   derived cardiomyocytes (hiPSC-CMs) using an EdU-incorporation assay (Figure 1F; see Online

21   Figure 4 for fibroblasts). Indeed, proband hiPSC-CMs exhibited reduced number of EdU-

22   positive cells (Figure 1G-I). To evaluate TP53-dependency of cell cycle defects, we performed

23   siRNA-mediated TP53 or CDKN1A knockdown (KD)in hiPSC-CMs. We observed significantly

1   elevated EdU incorporation in proband and parents (Online Figure 1D-M). Collectively, our

2   findings suggest HLHS may involve TP53 pathway-dependent blockade of CM proliferation,

3   consistent with observations made in both HLHS fetuses[6] and a HLHS mouse model.[5]

4

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

6   Array comparative genome hybridization ruled out a chromosomal deletion or

7   duplication in the proband. WGS was performed to identify potentially pathogenic coding or

8   regulatory single nucleotide variants (SNVs) or insertion/deletions (INDELs). Variants were

9   filtered for rarity, predicted impact on protein structure or expression, and *de novo* or recessive

10   inheritance, yielding 114 variants in 61 genes (Figure 2, Online Table 3). We next prioritized

11   genes most likely to drive downstream pathways of dysregulated cardiogenesis in the HLHS

12   proband by cross-referencing these candidate genes with 3,816 DETs identified in

13   undifferentiated iPSC at day 0 (d0) (Online Table 4) and 5,104 DETs identified at day 25 (d25)

14   differentiated cell lineages (Online Table 1). Ten genes harboring recessive variants were

15   differentially expressed within the HLHS proband's iPSCs at d0 and d25: *HSPG2, APOB, LRP2,*

16   *PRTG, SLC9A1, SDHD, JPT1, ELF4, HS6ST2* and *SIK1* (Figure 2).

17

18   **Knockdown of candidate gene orthologs in *Drosophila* heart**

19   In order to determine whether these variants occured within genes that could be important

20   for cardiac differentiation *in vivo*, we took advantage of our established *Drosophila* heart

21   development model and functional analysis tools (Figure 3A).[14] We hypothesized that genes

22   critical for the *Drosophila* heart have conserved roles also in humans, as previously observed[19,]

23   [20, 27]. Predicted by DIOPT database[22] to have orthologs in *Drosophila* (Figure 3B), we analyzed 9

9

1   genes using heart-specific RNAi KD. By *in vivo* heart structure and function analysis,[13] we found

2   that KD of *LRP2* (*mgl*), and *APOB* (*apolpp*) resulted in arrhythmias (Figure 3C,D, E; Online

3   movies S1, S2, S3). KD of *JPT1 (CG1943)* resulted in dilation, suggesting a developmental or

4   remodeling defect (Figure 3F).

5   Since HLHS is likely oligogenic,[28, 29] some genes in HLHS might only become apparent

6   in combination with mutations in other cardiac-relevant genes. To test this, we examined the

7   candidates also in the heterozygous background for *tinman/NKX2-5*, which in humans is well-

8   known to contribute to a variety of CHD/HLHS manifestations.[1, 25, 30, 31] In this *in vivo* context,

9   heart-specific knockdown of 2 out of 9 genes, namely HSPG2/Perlecan (*trol*), involved in

10   extracellular matrix assembly,[32] and the SDHD (*SdhD*), exhibited a constricted phenotype

11   (Figure 3G,H). These findings demonstrate that our bioinformatic candidate gene prioritization

12   identified several candidates as cardiac relevant.

13

14   **APOB and LRP2 regulate proliferation in human iPSC-derived cardiomyocytes**

15   Decreased proliferation of left ventricular cardiomyocytes is a phenotypic hallmark of

16   HLHS;[5, 6] see Figure 1G), suggesting that cell cycle impairment likely contributes to the disease.

17   Thus, we asked whether siRNA-mediated candidate gene KD affects hiPSC-CM proliferation

18   (Figure 4A).[23] Indeed, siRNAs directed against all 10 candidate genes together caused a marked

19   reduction of EdU+ cardiomyocytes (ACTN1+) and cardiac nuclei numbers (Figure 4B-E),

20   concomitant with downregulation of cell cycle genes (Figure 4F). In addition, proliferation

21   inhibitors (*TP53*) and apoptosis markers were increased (Figure 4G). Subsequently, KD of the

22   individual candidates revealed that siAPOB and siLRP2 were each able to recapitulate the

23   reduced hiPSC-CM proliferation and increased apoptosis phenotype (Figure 4H-M).

10

1   In order to determine the impact of *APOB* and *LRP2* during the process of cardiogenesis,

2   both genes were simultaneously knocked down in 5 day old hiPSC-derived cardiac progenitors

3   and cardiac differentiation was assessed at day 12 (Figure 4N). Lineage quantification was

4   performed using ACTN1 as a marker of cardiomyocytes, and TAGLN as a marker of fibroblasts.

5   Notably, APOB and LRP2 KD reduced cardiomyocyte numbers and differentiation and

6   increased fibroblast density (Figure 4O-Y). Collectively, these data identify *APOB* and *LRP2* as

7   major regulators of cell cycle, apoptosis, and cardiogenesis in hiPSC-CMs, potentially

8   contributing to the developmental cardiac impairment in HLHS patients.

9

10   **Potential regulatory network of validated gene canditates**

11   In order to delineate how potentially pathogenic mutations in the proband's six genes

12   validated in model systems might affect signaling homeostasis in proband cells, we assembled a

13   gene regulatory network containing *APOB, HS6ST2, HSPG2, JPT1, LRP2, SLC9A1* and their

14   first neighbors (genetic and protein-protein interactions, using BioGRID data for *Drosophila*

15   *melanogaster*) (Figure 5A; Online Table 5). Strikingly, the resulting network (82 genes; Online

16   Figure 5) connected all 6 genes to WNT and SHH signaling cascades, both key regulators of

17   cardiac differentiation and proliferation.[33, 34] Consistent with these findings, RNA-seq analysis of

18   the family trio cells reveiled that negative regulator of SHH pathway, *PTCH1*, was upregulated,

19   while agonists of WNT signaling pathway, *WNT1/3a/8a/10b* and *FZD10*,[35] were downregulated,

20   suggesting these pathways are attenuated in the proband. In addition, *TP53* and *CDKN1A*

21   expression was upregulated in response to siAPOB or siLRP2 (Figure 4), and notably also in

22   proband cells (Figure 1). Importantly, siTP53/CDKN1A restored cell cycle activity in proband-

11

1   derived cardiomyocytes (Online Figure 1). Thus, a WNT/SHH/TP53-mediated impairment of

2   cardiac differentiation and proliferation may contribute to HLHS in the proband (Figure 5A).

3

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

5   To evaluate whether the proband's regulatory network of genes is generalizable to

6   HLHS, we determined the frequency of rare, predicted-damaging variants in these genes in a

7   cohort of 130 HLHS cases and 861 controls. Of the 82 genes, 65 had at least one rare variant

8   present in cases or controls, yet *LRP2* was the only gene with significant enrichment in cases

9   (Online Table 6). HLHS patients had a ~3-fold increase in the frequency of *LRP2* missense

10   variants compared to healthy controls (10% versus 3.4%; p=0.0008). Moreover, among the 13

11   patients (including the index case) who carried a *LRP2* variant, 4 developed early heart failure

12   (Figure 5B, Online Table 7). Comparison of *LRP2* variants identified in cases and controls did

13   not reveal significant differences in their population-based frequencies, predicted-damaging

14   scores, or presence in binding domains (data not shown). Of note, 13 of the 130 HLHS cases

15   (including the index family proband) possessed <80% of ancestral Caucasian alleles, while all

16   controls posessed ≥80%. Four of the 13 cases had rare, predicted-damaging missense variants in

17   *LRP2* however all variants assessed were required to be rare in all racial populations. To

18   eliminate the potentially confounding variable of race a Caucasian-only sub-analysis was

19   performed, resulting in a less significant p-value for rare, predicted-damaging missense variants

20   (7.7% versus 3.4%; p=0.05). However, removal of the predicted-damaging restriction on rare

21   *LRP2* variants among Caucasians revealed significant enrichment in cases (p=0.0035), most

22   notably in missense and intronic variants (p=0.0178 and 0.0082, respectively) Online Table 8).

23   Location of variants within active histone marks or transcription factor binding sites was not

12

1   different between cases and controls. The biological significance of rare *LRP2* variants in HLHS

2   and potential race-specific differences require further study.

3

4   **DISCUSSION**

5   **Integrated Multidisciplinary Disease Gene Discovery and Modeling**

6   Unraveling the molecular-genetic etiology of HLHS pathogenesis is crucial for (1) our

7   ability to provide reliable prenatal diagnostics to families and (2) to develop novel approaches to

8   to treat the disease. As an important step towards these goals, our integrated multidisciplinary

9   approach identifies rare variants in *LRP2* as a molecular signature enriched in HLHS patients,

10   consistent with our mechanistic analysis in model systems.

11   For this study, we used a powerful combination of high-throughput DNA/RNA patient

12   sequencing coupled with high-throughput functional screening in model systems enabling to

13   probe gene function (alone or in combination) on a wide array of cellular processes that are

14   deployed during heart formation. For validation in model systems, we have established an

15   integrated multi-site and multidisciplinary pipeline that systematically evaluates the functional

16   role of genes presenting rare and deleterious variants in HLHS patients in hiPSC and *Drosophila*

17   heart models. As a main objective – identify genes and potential mechanisms associated with

18   HLHS - our pioneer study highlights rare, predicted-damaging *LRP2* missense variants as 3-fold

19   enriched in 130 HLHS patients compared to 861 controls. Validation of this gene in hiPSC and

20   *Drosophila* heart models demonstrates a requirement in cardiac proliferation, differentiation, and

21   rhythm. Mutations in *LRP2* have been previously associated with cardiac defects in mouse[36] and

22   in Donnai-Barrow Syndrome in humans.[36, 37] However, *LRP2* has not previously been linked to

23   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

14

1   combination of disease genes, and such a combination of alleles in turn could affect several,

2   successive steps of heart development. This makes it extremely difficult to model the disease by

3   single gene mutations. Current hypotheses of the etiology of HLHS include changes in cell cycle

4   progression of myocytes, as well as altered blood flow ("no flow – no grow") as a consequence

5   of defects of valves or the outflow tract.

6   Interestingly, the only mouse HLHS model to date, a digenic mutant for sap130 and

7   pcdha9,[5] has a penetrance of less than thirty percent, indicating a profound role for subtle

8   differences between genetic backgrounds. This mouse model suggests a separate mechanism

9   with pcdha9 affecting aortic growth, whereas sap130 can exert a more severe HLHS-like

10   phenotype, which might reflect a modular etiology of HLHS that separates valve and ventricular

11   defects.

12   The gene network analysis that we have conducted points to the possibility that several of

13   the prioritized candidate genes identified in the index patient can have a modulatory impact on

14   WNT and/or SHH via a diverse set of mechanisms.[33] There is evidence that the three of the

15   candidates with fly heart phenotypes - Trol/HSPG2, Mgl/LRP2 and Apolpp/APOB - can alter

16   WNT and SHH signaling,[43, 44] consistent with both pathways being attenuated in the proband

17   iPSCs. We hypothesize that a collective of likely hypomorphic mutations affects heart

18   development leading to HLHS. Impaired ventricular growth could in addition be caused by

19   changes in the p53 pathway, and our analysis of iPSC-derived cardiomyocytes suggests that p53

20   indeed depends on APOB and LRP2 levels. Such a multi-hit model of HLHS caused by sub-

21   threshold hypomorphic alleles represents an attractive explanation of the disease.

15

1   In summary, this integrated multidisciplinary strategy of functional genomics using

2   patient-specific iPSC combined with *in vivo* and human cellular model systems of functional

3   validation postulates novel genetic pathways and potential polygenic interactions underlying

4   HLHS. These complementory experimental approaches enabled the deconvolution of the

5   complex polygenic risk factors underlying HLHS genotype/phenotype and establish the novel

6   groundwork for definitive mechanistic studies of interacting risk factors that contribute to

7   arrested cardiac development and adverse outcomes. This scalable approach promises more

8   efficient discovery of novel genes associated more broadly with HLHS as multiple HLHS

9   families can now be multiplexed in future studies.

10

11 **Study Limitations**

12   HLHS candidate gene selection was based on in silico predictive algorithms to filter for

13 functional coding and regulatory variants. Mendelian inheritance modeling, a strategy used to

14 identify major-effect variants, did not include consideration of incompletely penetrant

15 heterozygous alleles inherited from an unaffected parent. Differential gene expression, which

16 was functionally validated as a powerful filter for candidate variant prioritization, excluded

17 functional variants that do not alter gene expression. Not all human genes are conserved in

18 *Drosophila*, but >70% of disease-causing human genes have a fly ortholog. While structural

19 differences exist between hiPSC-CM, *Drosophila* heart and human newborn cardiomyocytes, our

20 combinatorial approach allows to uncover gene networks and interactions that is not feasible in

21 mammalian model systems.

22

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).

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

1 R.B. and A.R.C.: Development, Aging and Regeneration Program, Sanford-Burnham-Prebys

2 Medical Discovery Institute, 10901 North Torrey Pines Road, La Jolla, CA92037, USA. Phone:

3 858-795-5295. Email: rolf@SBPdiscovery.org.

4

1   **REFERENCES**

2

3   1.   Elliott DA, Kirk EP, Yeoh T, Chandar S, McKenzie F, Taylor P, Grossfeld P, Fatkin D, Jones O,

4   Hayes P, Feneley M and Harvey RP. Cardiac homeobox gene NKX2-5 mutations and congenital heart

5   disease: associations with atrial septal defect and hypoplastic left heart syndrome. *J Am Coll Cardiol*.

6   2003;41:2072-2076.

7   2.   Lascone M, Ciccone R, Galletti L, Marchetti D, Seddio F, Lincesso AR, Pezzoli L, Vetro A,

8   Barachetti D, Boni L, Federici D, Soto AM, Comas JV, Ferrazzi P and Zuffardi O. Identification of de

9   novo mutations and rare variants in hypoplastic left heart syndrome. *Clin Genet*. 2012;81:542-554.

10   3.   Theis JL, Hrstka SC, Evans JM, O'Byrne MM, de Andrade M, O'Leary PW, Nelson TJ and Olson

11   TM. Compound heterozygous NOTCH1 mutations underlie impaired cardiogenesis in a patient with

12   hypoplastic left heart syndrome. *Hum Genet*. 2015;134:1003-1011.

13   4.   Theis JL, Zimmermann MT, Evans JM, Eckloff BW, Wieben ED, Qureshi MY, O'Leary PW and

14   Olson TM. Recessive MYH6 mutations in hypoplastic left heart with reduced ejection fraction. *Circ*

15   *Cardiovasc Genet*. 2015;8:564-571.

16   5.   Liu X, Yagi H, Saeed S, Bais AS, Gabriel GC, Chen Z, Peterson KA, Li Y, Schwartz MC,

17   Reynolds WT, Saydmohammed M, Gibbs B, Wu Y, Devine W, Chatterjee B, Klena NT, Kostka D, de

18   Mesy Bentley KL, Ganapathiraju MK, Dexheimer P, Leatherbury L, Khalifa O, Bhagat A, Zahid M, Pu

19   W, Watkins S, Grossfeld P, Murray SA, Porter Jr GA, Tsang M, Martin LJ, Benson DW, Aronow BJ and

20   Lo CW. The complex genetics of hypoplastic left heart syndrome. *Nat Genet*. 2017;49:1152.

21   6.   Gaber N, Gagliardi M, Patel P, Kinnear C, Zhang C, Chitayat D, Shannon P, Jaeggi E, Tabori U,

22   Keller G and Mital S. Fetal reprogramming and senescence in hypoplastic left heart syndrome and in

23   human pluripotent stem cells during cardiac differentiation. *Am J Pathol*. 2013;183:720-734.

19

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.

20

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

2   2010.

3   17.   Azpiazu N and Frasch M. Tinman and bagpipe: two homeo box genes that determine cell fates in

4   the dorsal mesoderm of Drosophila. *Genes Dev*. 1993;7:1325-1340.

5   18.   Bodmer R. The gene tinman is required for specification of the heart and visceral muscles in

6   Drosophila. *Development*. 1993;118:719-729.

7   19.   Bodmer R. Heart development in Drosophila and its relationship to vertebrates. *Trends*

8   *Cardiovasc Med*. 1995;5:21-28.

9   20.   Cripps RM and Olson EN. Control of cardiac development by an evolutionarily conserved

10   transcriptional network. *Dev Biol*. 2002;246:14-28.

11   21.   Benson DW, Silberbach GM, Kavanaugh-McHugh A, Cottrill C, Zhang Y, Riggs S, Smalls O,

12   Johnson MC, Watson MS, Seidman JG, Seidman CE, Plowden J and Kugler JD. Mutations in the cardiac

13   transcription factor NKX2.5 affect diverse cardiac developmental pathways. *J Clin Invest*.

14   1999;104:1567-1573.

15   22.   Hu Y, Flockhart I, Vinayagam A, Bergwitz C, Berger B, Perrimon N and Mohr SE. An

16   integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC*

17   *Bioinformatics*. 2011;12:357.

18   23.   Cunningham TJ, Yu MS, McKeithan WL, Spiering S, Carrette F, Huang C-T, Bushway PJ,

19   Tierney M, Albini S, Giacca M, Mano M, Puri PL, Sacco A, Ruiz-Lozano P, Riou J-F, Umbhauer M,

20   Duester G, Mercola M and Colas AR. Id genes are essential for early heart formation. *Genes Dev*.

21   2017;31:1325-1338.

1   24.   Yu MS, Spiering S and Colas AR. Generation of first heart field-like cardiac progenitors and

2   ventricular-like cardiomyocytes from human pluripotent stem cells. *J Vis Exp*. 2018:e57688.

3   25.   Hrstka SCL, Li X, Nelson TJ and Wanek Program Genetics Pipeline G. NOTCH1-dependent

4   nitric oxide signaling deficiency in Hypoplastic Left Heart Syndrome revealed through patient-specific

5   phenotypes detected in bioengineered cardiogenesis. *Stem Cells*. 2017;35:1106-1119.

6   26.   Takahashi K and Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and

7   adult fibroblast cultures by defined factors. *Cell*. 2006;126:663-676.

8   27.   Qian L, Wythe JD, Liu J, Cartry J, Vogler G, Mohapatra B, Otway RT, Huang Y, King IN,

9   Maillet M, Zheng Y, Crawley T, Taghli-Lamallem O, Semsarian C, Dunwoodie S, Winlaw D, Harvey

10   RP, Fatkin D, Towbin JA, Molkentin JD, Srivastava D, Ocorr K, Bruneau BG and Bodmer R.

11   Tinman/Nkx2-5 acts via miR-1 and upstream of Cdc42 to regulate heart function across species. *J Cell

12   Biol*. 2011;193:1181-1196.

13   28.   Blue GM, Kirk EP, Giannoulatou E, Sholler GF, Dunwoodie SL, Harvey RP and Winlaw DS.

14   Advances in the genetics of congenital heart disease: A clinician's guide. *J Am Coll Cardiol*.

15   2017;69:859-870.

16   29.   Gelb BD and Chung WK. Complex genetics and the etiology of human congenital heart disease.

17   *Cold Spring Harb Perspect Med*. 2014;4.

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

19   31.   Kobayashi J, Yoshida M, Tarui S, Hirata M, Nagai Y, Kasahara S, Naruse K, Ito H, Sano S and

20   Oh H. Directed differentiation of patient-specific induced pluripotent stem cells identifies the

21   transcriptional repression and epigenetic modification of NKX2-5, HAND1, and NOTCH1 in Hypoplastic

22   Left Heart Syndrome. *PLoS One*. 2014;9:e102796.

22

1 32.   Sasse P, Malan D, Fleischmann M, Roell W, Gustafsson E, Bostani T, Fan Y, Kolbe T, Breitbach

2 M, Addicks K, Welz A, Brem G, Hescheler J, Aszodi A, Costell M, Bloch W and Fleischmann BK.

3 Perlecan is critical for heart stability. *Cardiovasc Res*. 2008;80:435-444.

4 33.   Briggs LE, Burns TA, Lockhart MM, Phelps AL, Van den Hoff MJB and Wessels A. Wnt/β-

5 catenin and sonic hedgehog pathways interact in the regulation of the development of the dorsal

6 mesenchymal protrusion. *Dev Dyn*. 2016;245:103-113.

7 34.   Gessert S and Kühl M. The multiple phases and faces of Wnt signaling during cardiac

8 differentiation and development. *Circ Res*. 2010;107:186-199.

9 35.   Dawson K, Aflaki M and Nattel S. Role of the Wnt-Frizzled system in cardiac pathophysiology: a

10 rapidly developing, poorly understood area with enormous potential. *J Physiol*. 2013;591:1409-1432.

11 36.   Baardman ME, Zwier MV, Wisse LJ, Gittenberger-de Groot AC, Kerstjens-Frederikse WS,

12 Hofstra RM, Jurdzinski A, Hierck BP, Jongbloed MR, Berger RM, Plosch T and DeRuiter MC. Common

13 arterial trunk and ventricular non-compaction in Lrp2 knockout mice indicate a crucial role of LRP2 in

14 cardiac development. *Dis Model Mech*. 2016;9:413-425.

15 37.   Kantarci S, Ragge NK, Thomas NS, Robinson DO, Noonan KM, Russell MK, Donnai D,

16 Raymond FL, Walsh CA, Donahoe PK and Pober BR. Donnai-Barrow syndrome (DBS/FOAR) in a child

17 with a homozygous LRP2 mutation due to complete chromosome 2 paternal isodisomy. *Am J Med Genet*

18 *A*. 2008;146A:1842-1847.

19 38.   Cammarato A, Ahrens CH, Alayari NN, Qeli E, Rucker J, Reedy MC, Zmasek CM, Gucek M,

20 Cole RN, Van Eyk JE, Bodmer R, O'Rourke B, Bernstein SI and Foster DB. A mighty small heart: the

21 cardiac proteome of adult Drosophila melanogaster. *PLoS One*. 2011;6:e18497.

1   39.   Ocorr K, Reeves NL, Wessells RJ, Fink M, Chen HS, Akasaka T, Yasuda S, Metzger JM, Giles

2   W, Posakony JW and Bodmer R. KCNQ potassium channel mutations cause cardiac arrhythmias in

3   Drosophila that mimic the effects of aging. *Proc Natl Acad Sci U S A*. 2007;104:3943-8.

4   40.   Diop SB and Bodmer R. Gaining insights into diabetic cardiomyopathy from Drosophila. *Trends

5   Endocrinol Metab*. 2015;26:618-627.

6   41.   Yagi H, Liu X, Gabriel GC, Wu Y, Peterson K, Murray SA, Aronow BJ, Martin LJ, Benson DW

7   and Lo CW. The genetic landscape of Hypoplastic Left Heart Syndrome. *Pediatr Cardiol*. 2018;39:1069-

8   1081.

9   42.   Pierpont Mary E, Brueckner M, Chung Wendy K, Garg V, Lacro Ronald V, McGuire Amy L,

10 Mital S, Priest James R, Pu William T, Roberts A, Ware Stephanie M, Gelb Bruce D, Russell Mark W

11 and null n. Genetic basis for congenital heart disease: Revisited: A scientific statement from the American

12 Heart Association. *Circulation*. 2018;138:e653-e711.

13 43.   Christ A, Christa A, Klippert J, Eule JC, Bachmann S, Wallace Valerie A, Hammes A and

14 Willnow Thomas E. LRP2 acts as SHH clearance receptor to protect the retinal margin from mitogenic

15 stimuli. *Dev Cell*. 2015;35:36-48.

16 44.   Datta MW, Hernandez AM, Schlicht MJ, Kahler AJ, DeGueme AM, Dhir R, Shah RB, Farach-

17 Carson C, Barrett A and Datta S. Perlecan, a candidate gene for the CAPB locus, regulates prostate cancer

18 cell growth via the Sonic Hedgehog pathway. *Molecular Cancer*. 2006;5:9.

19

20

1   **Figure Legend 1. Family-based iPSC characterization for HLHS.** (A) Pedigree of family 5H:

2   proband with HLHS (black symbol), relatives without CHD (white symbols), miscarriages (gray

3   diamonds). (B) Schematic for family-based iPSC production and characterization. (C) Whole

4   genome RNA sequencing identifies 1401 concordantly DETs between proband and parents. (D)

5   KEGG pathway analysis shows enrichment of DETs in TP53 pathway. (E) Heatmap of p53

6   signaling pathway associated genes in probands vs parents. (F) Schematic describing EdU

7   incorporation assay in hiPSC-CMs. (G) Graph representing quantification of EdU+

8   cardiomyocytes in HLHS family-derived iPSC-CMs. ***$p<0.001$ One-way ANOVA. (H,I)

9   Representative images of iPSC-CMs derived from mother (H) and proband (I), stained for EdU,

10   ACTN1 and DAPI. Scale bar: 50 µm.

11

1   **Figure Legend 2. Identification of HLHS candidate genes from whole genome and RNA**

2   **sequencing.** An iterative, family-based variant filtering approach based on rarity, functional

3   impact, and mode of inheritance yielded 61 candidate genes. RNA sequencing data from d0 iPSC

4   and d25 differentiated iPSC were used to filter for transcriptional differences yielding 10

5   candidate genes.

6

7

1   **Figure Legend 3. Phenotypic assessment of HLHS candidate genes in *Drosophila* adult**

2   **hearts.** (**a**) Schematic of the *Drosophila* adult heart assay. (**b**) Human candidate genes and

3   corresponding *Drosophila* ortholog as determined by DIOPT score (*confidence score: number

4   of databases reporting orthology). Listed are heart phenotypes upon knockdown in wildtype or

5   *NKX2-5*/*tin*$^{+/-}$ heterozygous background. (**c-e**) RNAi-induced arrhythmicity and M-modes

6   observed with *APOB* and *LRP2* knockdown. (**f-h**) Heart size (EDD: enddiastolic diameter)

7   alterations upon RNAi-knockdown of *JPT1*, *HSPG2* and *SDHD* (also in NKX2-5/tin

8   heterozygous background). Wilcoxon rank sum test: *p<0.05, **p<0.01, **p<0.001.

9

10

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): *$p<0.05$, **$p<0.01$, ***$p<0.001$,****$p<0.0001$.

28

1 **Figure Legend 5. Potential role for SHH, WNT and LRP2 in HLHS.** (**a**) A gene network

2 integrating family-centric HLHS candidate genes with heart development. ORANGE – genes

3 with cardiac phenotypes in iPSC/*Drosophila* assays. YELLOW – other candidate genes tested.

4 RED – Genes up-regulated in proband iPSCs vs. parents. BLUE – Genes downregulated down-

5 regulated in proband iPSCs vs. parents. (**b**) Cohort-wide analysis of LRP2 variants shows

6 significant enrichment for SNVs in HLHS patients compared to control populations. Variants

7 (blue / magenta) are found throughout LRP2 protein.
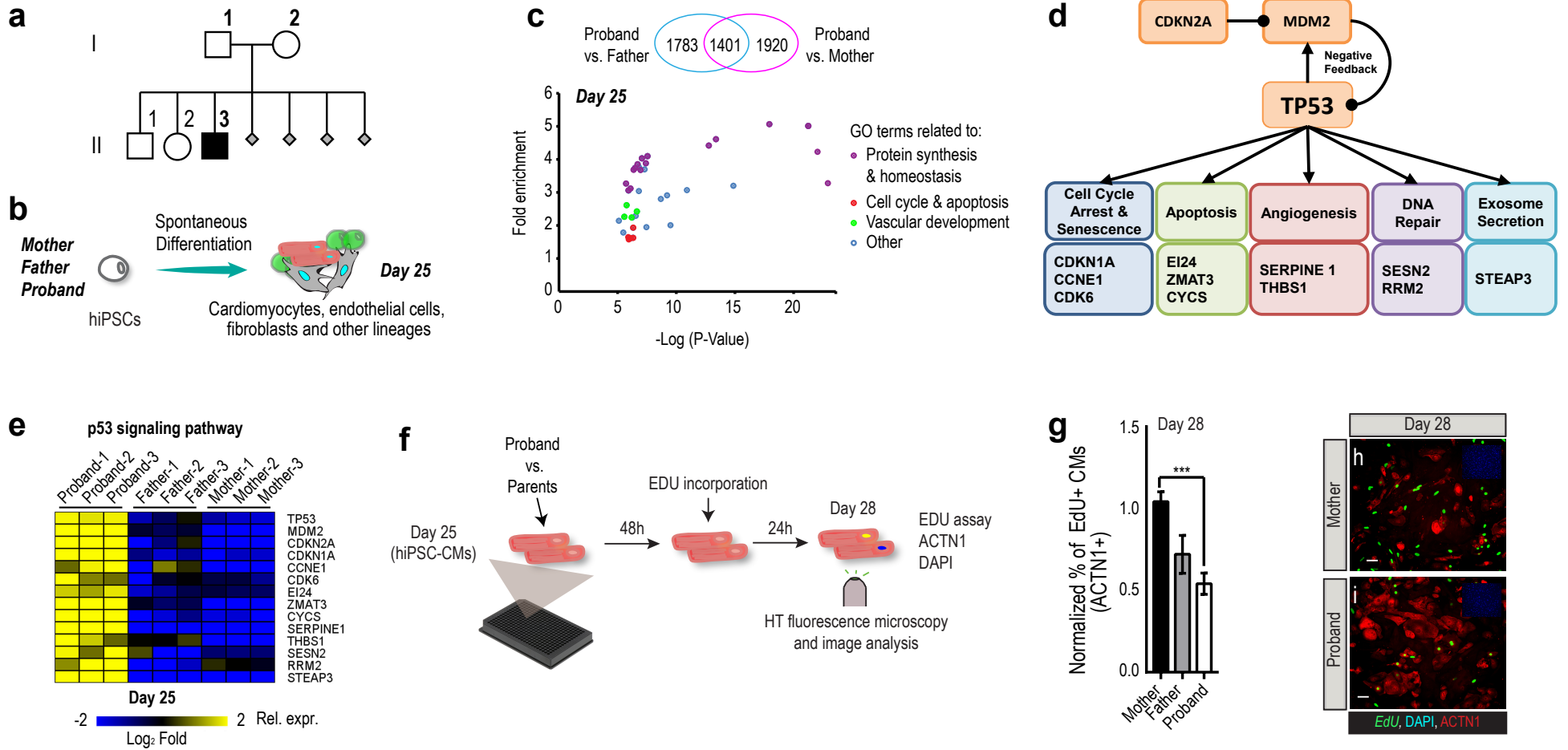
8

Figure 1

a

b

Mother
Father
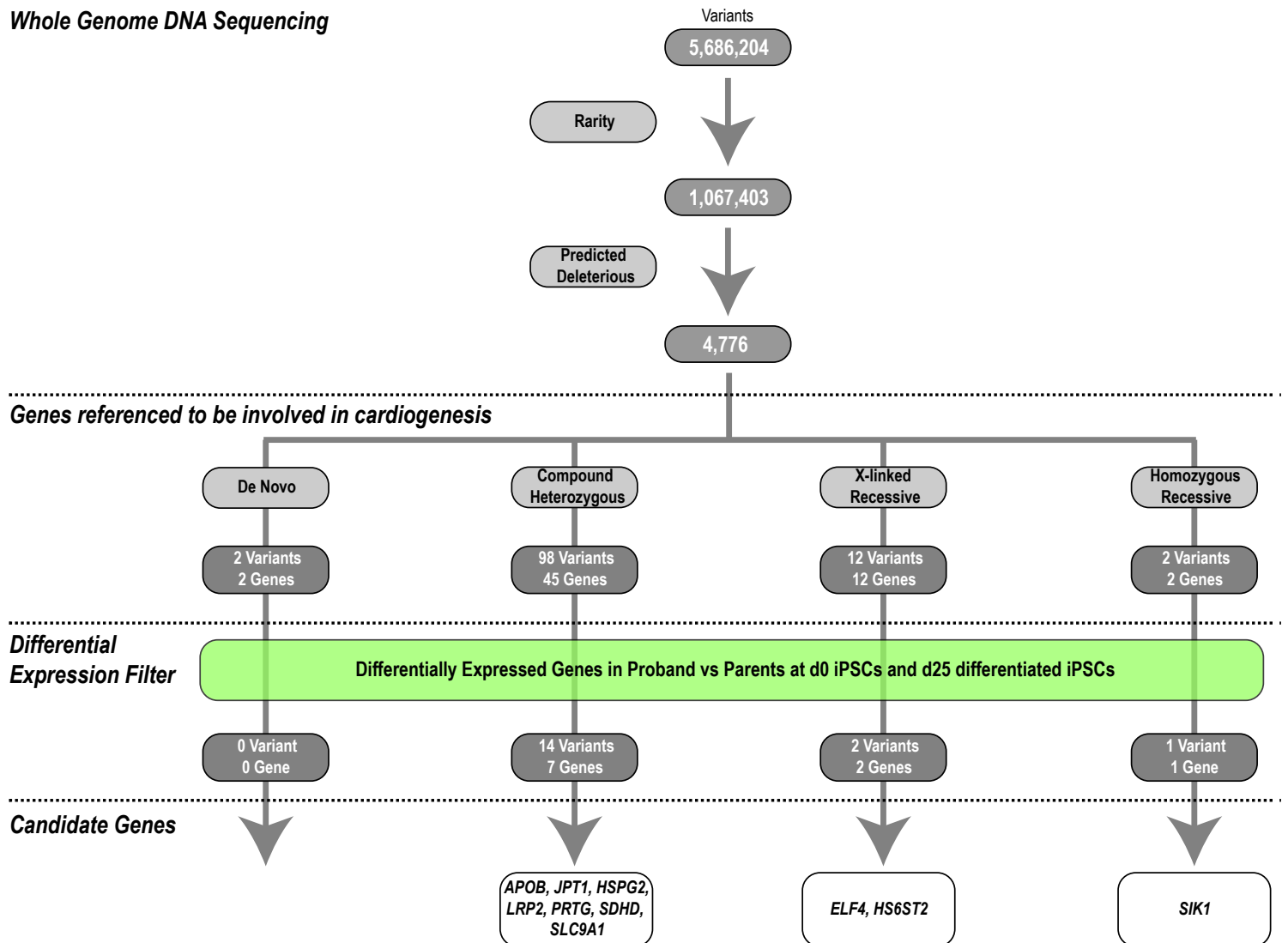*Proband*

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

TP53
MDM2
CDKN2A
CDKN1A
CCNE1
CDK6
EI24
ZMAT3
CYCS
SERPINE1
THBS1
SESN2
RRM2
STEAP3

Day 25

-2   2   Rel. expr.
Log$_2$ 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

h Mother — Day 28

i Proband

EdU, DAPI, ACTN1

Figure 2

Figure 3

a

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

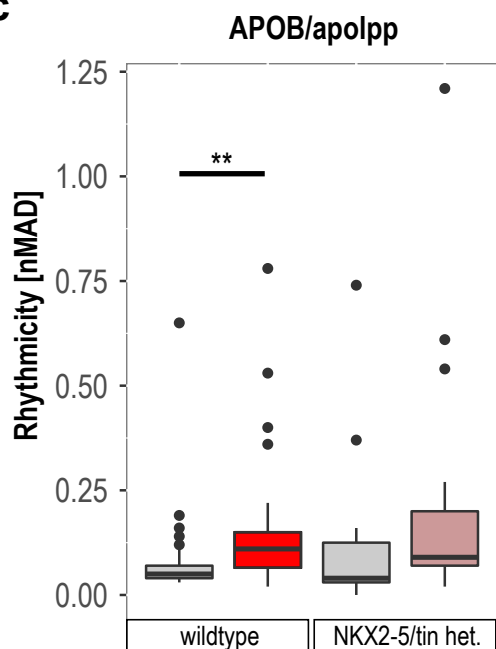3 week old female *Drosophila melanogaster*

10x
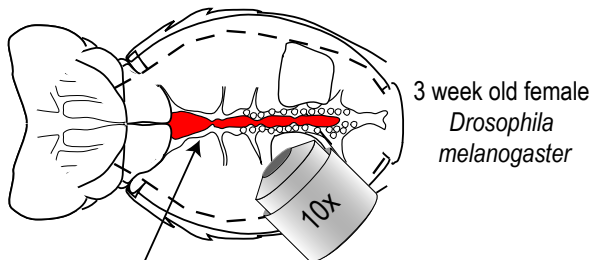
Kinetic imaging (30s@140fps) followed by automated and quantitative assessment of heart size and function

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 APOB/apolpp

d LRP2/mgl

e M-modes

f JPT1 / CG1943

g HSPG2 / trol

h SDHD / SdhD
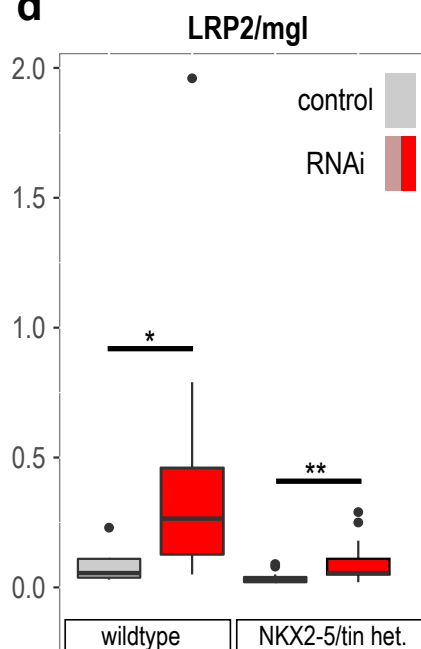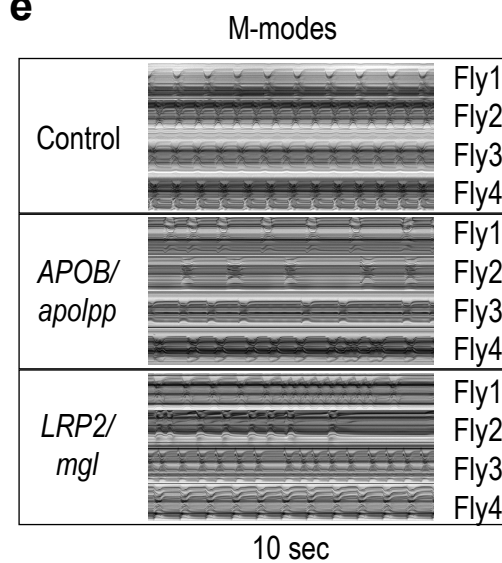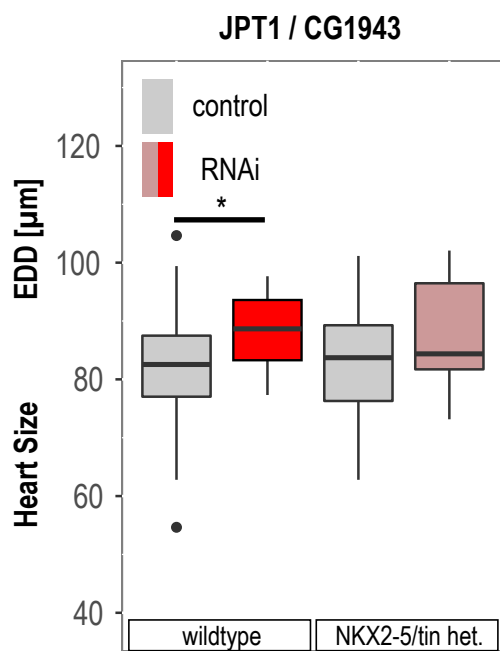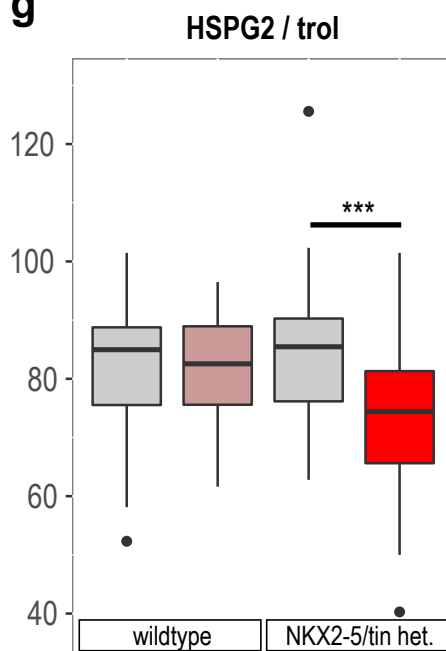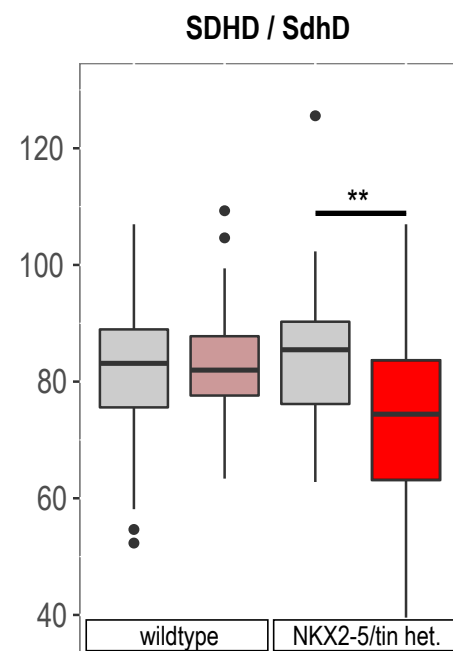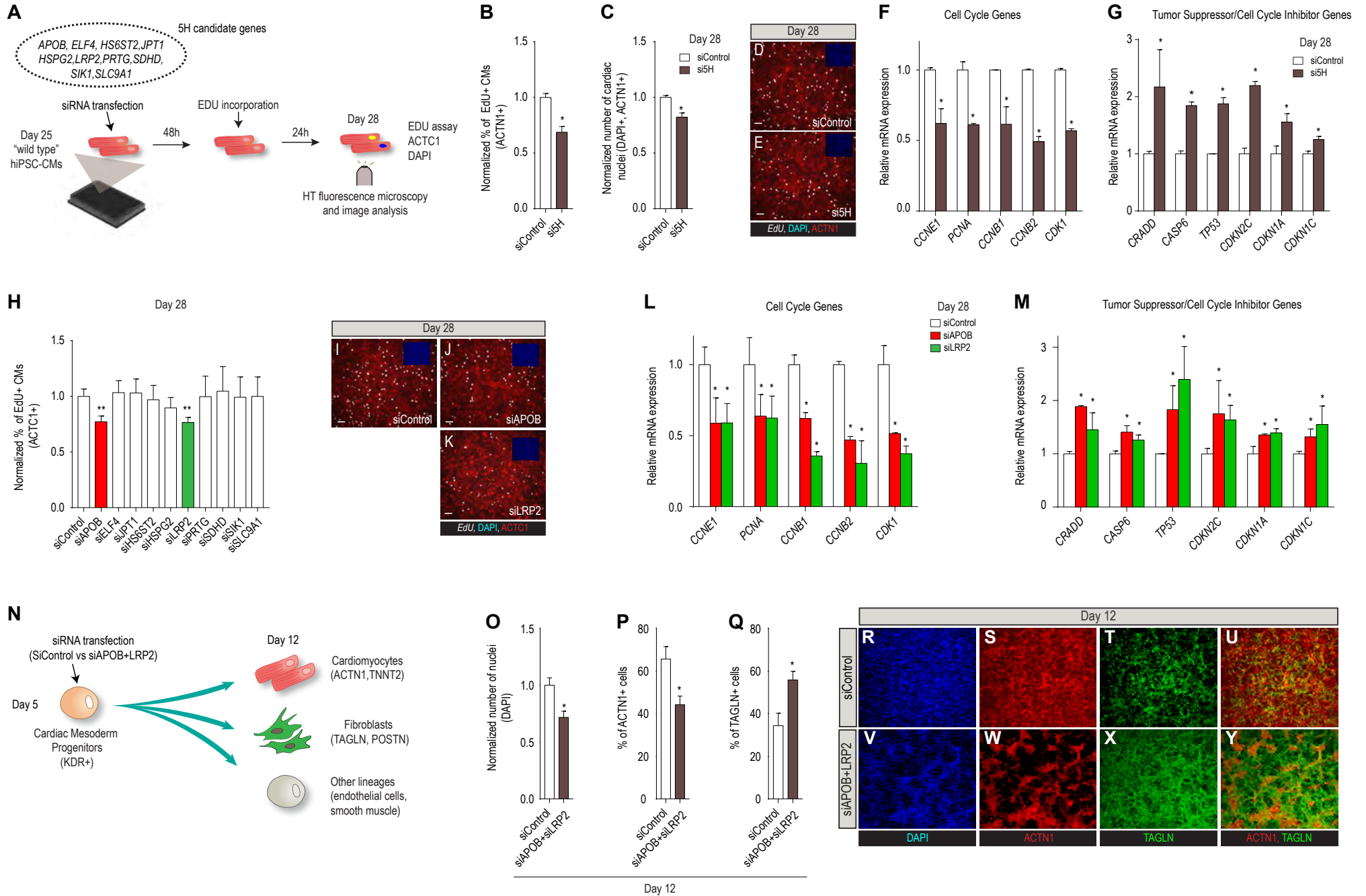
Figure 4

Figure 5

A
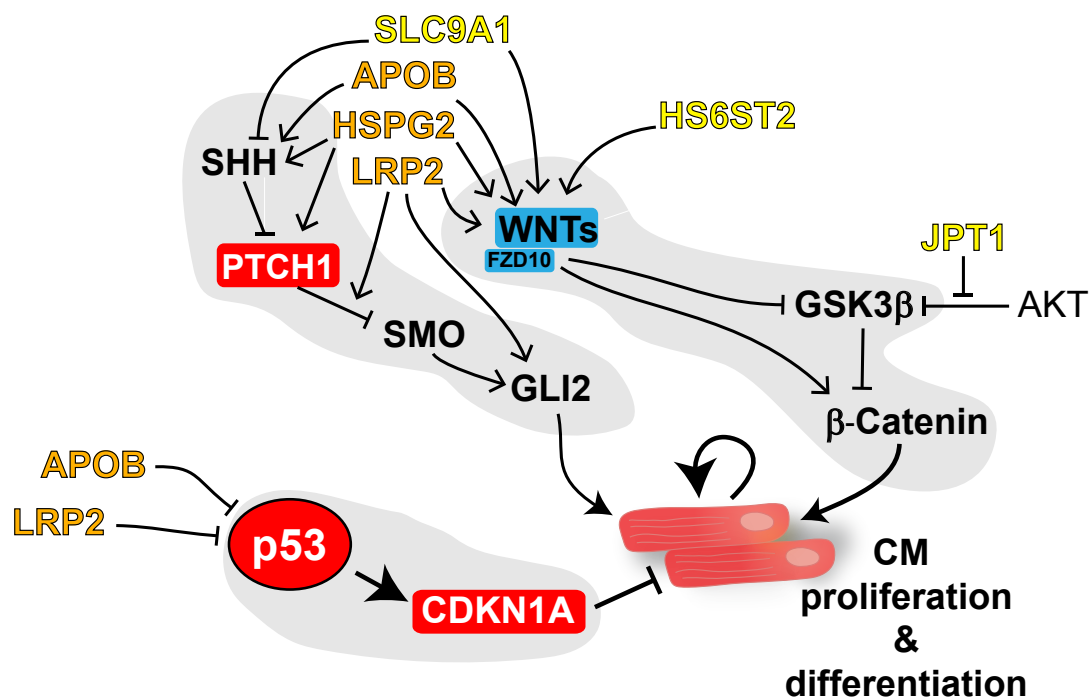
B

		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