 doi:10.20944/preprints202202.0017.v1

Article

Novel Protein-Protein Interactions Highlighting the Crosstalk between Hypoplastic Left-Heart Syndrome, Ciliopathies and Neurodevelopmental Delays

Kalyani B. Karunakaran[1], George C Gabriel[2], N. Balakrishnan[1], Cecilia W. Lo[2], and Madhavi K. Ganapathiraju*[3,4]

[1] Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore 560012, India
[2] Department of Developmental Biology, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA
[3] Department of Biomedical Informatics, School of Medicine, University of Pittsburgh, Pittsburgh, PA, USA
[4] Intelligent Systems Program, School of Computing and Information, University of Pittsburgh, Pittsburgh, PA, USA
* Correspondence: madhavi@pitt.edu;

Abstract: Hypoplastic left heart syndrome (HLHS) is a severe congenital heart disease (CHD) affecting 1 in 5,000 newborns. We constructed the interactome of 74 HLHS-associated genes identified from a large-scale mouse mutagenesis screen, augmenting it with 408 novel protein-protein interactions (PPIs) using our High-precision Protein-Protein Interaction Prediction (HiPPIP) model. The interactome is available on a webserver with advanced search capabilities (http://severus.dbmi.pitt.edu/wiki-HLHS). 364 genes including 73 novel interactors were differentially regulated in tissues/iPSC-derived cardiomyocytes of HLHS patients. Novel PPIs facilitated the identification of TOR signaling and endoplasmic reticulum stress modules. 60.5% of the interactome consisted of housekeeping genes that may harbor large-effect mutations and drive HLHS etiology but show limited transmission. Network proximity of diabetes, Alzheimer's disease, and liver carcinoma-associated genes to HLHS genes suggested a mechanistic basis for their comorbidity with HLHS. Interactome genes showed tissue-specificity for sites of extracardiac anomalies (placenta, liver and brain). The HLHS interactome shared significant overlaps with the interactomes of ciliopathy and microcephaly-associated genes, with the shared genes respectively enriched for genes involved in intellectual disability and/or developmental delay, and neuronal death pathways. This supported the increased burden of ciliopathy variants and prevalence of neurological abnormalities observed among HLHS patients with developmental delay and microcephaly respectively.

Keywords: protein-protein interactions; interactome; congenital heart disease; developmental disorder; hypoplastic left heart syndrome; web application

1. Introduction

Hypoplastic left heart syndrome (HLHS) is a severe form of congenital heart disease (CHD), which is one of the most common birth defects affecting ~1% of live births and a major driver of infant mortality resulting from congenital defects [1]. CHD constitutes structural abnormalities that can affect any cardiac structure including the atria, ventricles, aorta, and pulmonary artery or the valves connecting these chambers. Examples of CHD include atrial and ventricular septal defects, conotruncal defects affecting the ventricular septum and the outflow tract, complex CHD involving disturbance of left-right patterning (e.g. transposition of the great arteries), valvular defects including inflow (mitral and tricuspid) and outflow (aortic/pulmonic) valves.

CHD classified as left ventricular outflow tract obstructive (LVOTO) lesions comprise a constellation of structural heart defects involving obstruction of flow from the left ventricle (LV). Clinical studies have provided strong evidence of a shared genetic etiology for LVOTO lesions, such as hypoplastic left heart syndrome (HLHS), bicuspid aortic valve (BAV) and coarctation (CoA) [1]. HLHS is a complex CHD, constituting ~1.4 to 3.8% of the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2022 doi:10.20944/preprints202202.0017.v1

CHD cases and estimated to affect 1 in 5,000 newborns [2]. It is characterized by underdevelopment of the structures on the left side of the heart, namely, atresia or critical stenosis of the mitral or aortic valves and hypoplasia of the left ventricle, ascending aorta and aortic arch [2]. Until ~30 years ago, infants born with this condition would have died within the first few weeks of life; 23% of the deaths occurring in the first week of life due to cardiac abnormalities have been attributed to HLHS [2,3]. The incidence of HLHS during the fetal stage could be higher. Currently, surgical palliative techniques and improved postoperative care have significantly improved survival, with ~60-70% HLHS neonates surviving for at least 5 years following repair [4-6]. Mortality however is highest in the first year of life, with 30% dying or requiring heart transplant before one year old. Nevertheless, 90 percent of those surviving to one year will survive long-term up to 18 years old and beyond [7].

Brain comorbidities such as corpus callosum agenesis, holoprosencephaly, microcephaly and white matter injury have been identified in HLHS neonates, and cognitive, motor and behavioral adverse outcomes such as attention-deficit hyperactivity disorder, learning disabilities, and global developmental delay have been noted among HLHS survivors [8,9]. Given the high mortality and comorbidities associated with HLHS, there is critical need to investigate the molecular mechanism(s) of disease pathogenesis in HLHS, as only then can therapies be developed to improve outcome.

In humans, a genetic etiology for HLHS is demonstrated by high familial aggregation of HLHS with other LVOTO defects. Thus, using a statistical framework to calculate genetic effect size, >0.9 heritability was observed for HLHS and >0.7% for HLHS associated with other cardiovascular malformations (p-value < 1E-05) [10]. HLHS is also shown to have a complex multigenic etiology, with clinical studies suggesting a digenic etiology being the most likely [11]. Supporting such complex genetics, a large-scale mutagenesis screen in mice for mutations causing CHD recovered 8 mutant lines with HLHS. None shared any genes in common, and none showed Mendelian pattern of inheritance. Together these findings indicated HLHS has a multigenic etiology and is genetically heterogeneous. Interestingly, the recovery of the HLHS causing mutations in one HLHS mouse line, *Ohia*, confirmed a digenic etiology with mutations in two genes, *Sap130* and *Pcdha9*, shown to cause HLHS [12]. Further supporting a multigenic etiology is the finding that 5 of the 8 HLHS mutant mouse lines had two or more genes in 10 of 14 human chromosome linkage intervals associated with HLHS, with significant enrichment observed when two or more of the mouse HLHS associated genes were interrogated across these linkage intervals (OR 322.5; CI 24.9–4,177.2; $P = 5.6 \times 10^{-10}$) [13].

CHD-associated de novo mutations in histone-modifying genes such as KMT2D, CHD7, KDM5A, KDM5B, WDR5, RNF20, UBE2B and USP44 were identified in an exome sequencing study conducted with 60 HLHS cases and 264 controls by the Pediatric Cardiac Genomics Consortium (PCGC) [14]. Genome sequencing studies and genome-wide screening by comparative genomic hybridization have identified HLHS-associated variants in cardiomyopathy-associated genes such as MYBPC3, RYR2 and MYH6 [15], genes associated with mechanotransduction in cardiomyocytes such as VASP and TLN2 [16]. Other genes implicated in HLHS included RBFOX2 which mediates RNA metabolism [19], the cardiac transcription factor PROX1 [20], the endocytic receptor LRP2 [17] and the transcriptional regulator POGZ found in patients with HLHS and developmental delay [21]. However, despite the recovery of genes associated with HLHS, an integrative approach to elucidate their functional consequences is still lacking.

In the current study, we examined HLHS within the mechanistic framework of the protein-protein interaction (PPI) network or protein 'interactome'. Proteins fuel the cellular machinery, and their interactions reflect functions they subserve. This can be informative of disease mechanisms and may also help uncover higher-order relationships in the genetic architecture of complex disorders [22-24]. However, only ~145,000 PPIs (25%) out of the estimated ~600,000 PPI estimated to exist are known from public repositories such as HPRD [25] and BioGRID [26]. Detecting these PPIs using experimental techniques such

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2022 doi:10.20944/preprints202202.0017.v1

as the yeast two-hybrid system and co-immunoprecipitation is prohibitively time-consuming and expensive. Hence, we have developed a machine learning computational method to predict PPIs called HiPPIP (high-precision protein-protein interaction prediction. HiPPIP computes features of protein pairs such as cellular localization, molecular function, biological process membership, genomic location of the gene and gene expression in microarray experiments, and classifies the pairwise features as *interacting* or *non-interacting* based on a random forest model [22]. This method has been validated as accurate by computational evaluations [22] and experimental validations [22,27,28]. The novel PPIs predicted using HiPPIP have yielded discoveries with translational impact, including identifying the central role of cilia in CHD [12,22,29]. Here we constructed a 'HLHS interactome' with over 400 novel PPIs predicted by HiPPIP and over 1,400 known PPIs. We further developed a web resource with the novel PPIs on Wiki-HLHS, an interactive webserver for exploring novel interactions relevant to HLHS proteins or pathways of interest. We demonstrate the utility of the HLHS interactome for discovering higher-order genetic architecture of HLHS based on network analysis, functional enrichment, and transcriptome analyses.

2. Materials and Methods

Compilation of HLHS-associated genes and prediction of novel interactions

A list of 74 HLHS-associated genes was compiled from HLHS mutant mice from 8 independent mouse lines recovered from a large scale mouse mutagenesis screen [12,29]. This includes all homozygous mutations identified in the 8 HLHS mouse lines and heterozygous mutations also found in the HLHS human linkage intervals. Novel PPIs of the proteins encoded by these genes were predicted using the HiPPIP model that we developed [31]. Each HLHS protein (say N1) was paired with each of the other human proteins say, (M1, M2,…Mn), and each pair was evaluated with the HiPPIP model [31]. The predicted interactions of each of the HLHS proteins were extracted (namely, the pairs whose score is >0.5, a threshold which through computational evaluations and experimental validations was revealed to indicate interacting partners with high confidence). The interactome figures were created using Cytoscape [32].

Identification of network modules

Network modules among the HLHS proteins and their interactors were identified using Netbox [33]. Netbox reports modularity and a scaled modularity score, as compared with the modularity observed in 1000 random permutations of the subnetwork. Scaled modularity refers to the standard deviation difference between the observed subnetwork and the mean modularity of the random networks [34].

Functional enrichment analysis

Pathway associations of genes in the HLHS interactome were computed using IPA (Ingenuity Pathway Analysis) [35]. Statistical significance of the overlaps between genes in the HLHS interactome and pathways in the Ingenuity Knowledge Base (IKB) was computed with Fisher's exact test based on hypergeometric distribution. Biological process, cellular component and molecular function (Gene Ontology [36]), pathway (Reactome [37]), disease (OMIM [38] and DisGeNET [39]) and transcription factor target (MSigDB [40]) enrichments were computed using WebGestalt [41]. WebGestalt computes the distribution of genes belonging to a particular functional category in the input list and compares it with the background distribution of genes belonging to this functional category among all the genes that belongs to any functional category in the database selected by the user. Statistical significance of functional category enrichment is computed using Fisher's exact test, and corrected using the Benjamini-Hochberg method for multiple test adjustment. Annotations with FDR-corrected p-value < 0.05 were considered significant.

Gene expression enrichment analysis

The enrichment of the HLHS interactome in genes expressed in specific tissues was computed using RNA-sequencing data from 53 postnatal human tissues extracted from

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2022 doi:10.20944/preprints202202.0017.v1

GTEx [42]. Two gene sets were compiled for the analysis. The first set contained genes showing high or medium expression (transcripts per million (TPM) ≥ 9) in 53 tissues, provided that they were not housekeeping genes, i.e. genes detected in all the tissues with transcripts per million ≥ 1, as identified in the Human Protein Atlas [43]. The second set contained all the genes that showed high or medium expression in the 53 tissues, irrespective of whether they were housekeeping genes or not. TPM is a metric for quantifying gene expression; it directly measures the relative abundance of transcripts. GMT files served as inputs for the gene over-representation analysis (GSEA) that was conducted based on hypergeometric distribution. Tissue-specificity of the genes in the HLHS interactome were checked using TissueEnrich [44]. The analysis was based on tissue-specific genes compiled from GTEx [42], Human Protein Atlas [43], and Mouse ENCODE [45]. This included 'tissue-enriched genes' with at least 5-folds higher mRNA levels in a particular tissue compared to all the other tissues, 'group-enriched genes' with at least 5-folds higher mRNA levels in a group of 2-7 tissues and 'tissue-enhanced genes' with at least 5-folds higher mRNA levels in a particular tissue compared to average levels in all tissues.

Network overlap analysis

Statistical significance of the overlaps between genes in the HLHS interactome and in the SARS-CoV-2-modulated host protein interactome, the ciliary interactome, the ciliopathy interactome and the microcephaly interactome was computed based on hypergeometric test.

3. Results

We compiled a list of 74 genes associated with HLHS that were previously identified from 8 independent mouse lines with HLHS [12,29]. The protein interactome of these HLHS-associated genes were assembled by collecting the known protein-protein interactions (PPIs) from the Human Protein Reference Database (HPRD) [25] and the Biological General Repository for Interaction Datasets (BioGRID) [26], and also predicted novel PPIs by applying the HiPPIP algorithm, described in our earlier work [22]. In this manner, we assembled the HLHS interactome with 1,496 previously known PPIs (blue) and 408 novel computationally predicted PPIs (red), which altogether connected 72 of the 74 HLHS-associated genes with 1,248 known interactors and 377 novel interactors (**Figure 1** and **Data File S1**). Among the 74 genes, only WFDC11 and XKR9 had neither known nor novel PPIs.

Wiki-HLHS: A web server of HLHS PPIs

To accelerate biomedical discovery, we made the HLHS interactome publicly accessible with the construction of a web application called *Wiki-HLHS* (http://severus.dbmi.pitt.edu/wiki-HLHS). This web server has advanced search capabilities, and for each pair of PPI, there is side-by-side comprehensive Gene Ontology (GO) annotations, and annotations related to diseases, drugs and pathways. Here, a user can query for results such as "show me PPIs where one protein is involved in HLHS and the other is involved in microcephaly", and then see the results with the functional details of the two proteins side-by-side. This pairwise retrieval of PPIs and their biomedical associations is a unique feature of this web application not available from any other PPI web database. The PPIs and their annotations also are indexed in major search engines like Google and Bing. A user can browse the genes in the HLHS interactome using *List view* of HLHS associated genes. Novel PPIs are shown in a different color in search results.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2022 doi:10.20944/preprints202202.0017.v1

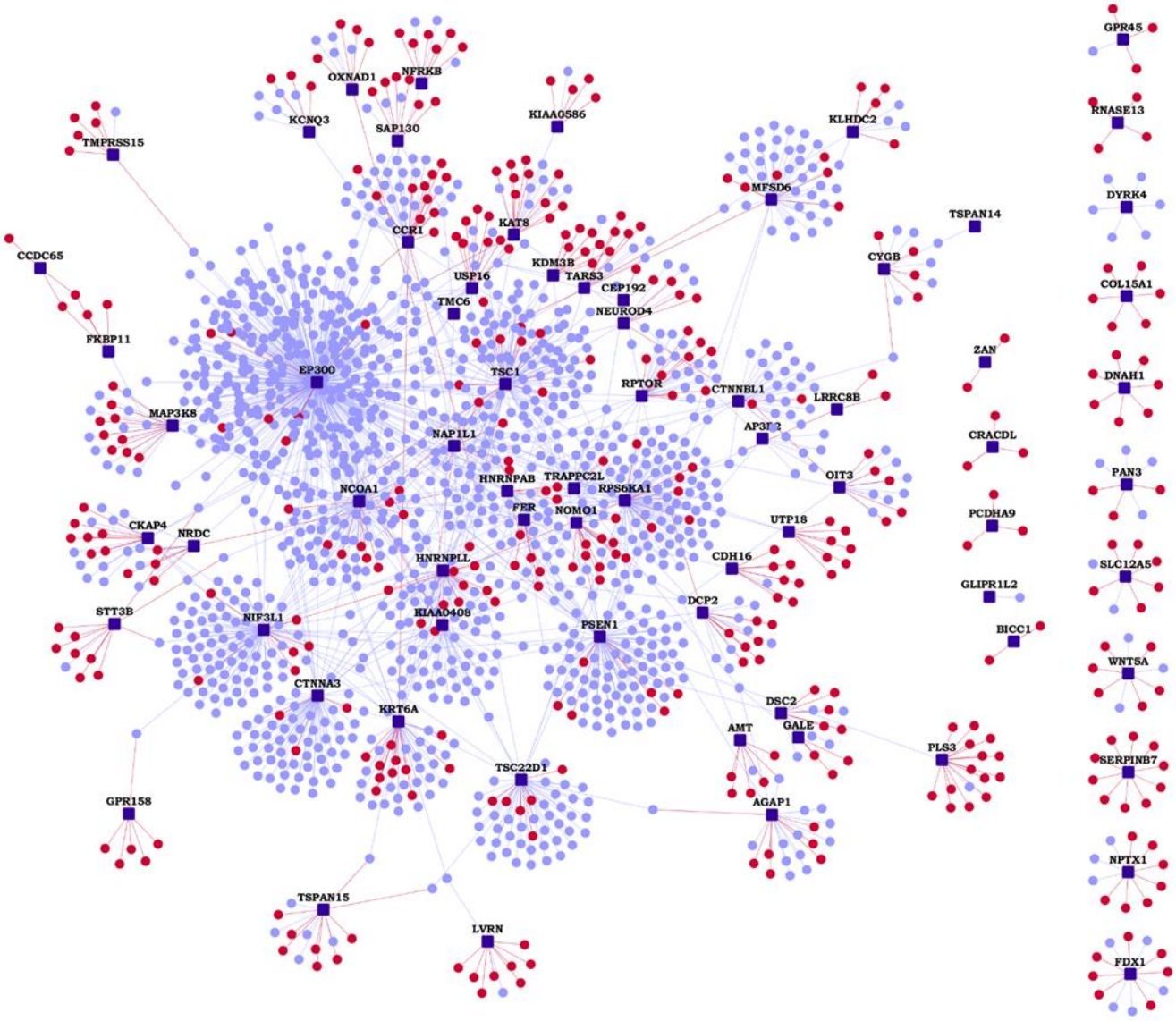

Figure 1. Hypoplastic left heart syndrome **(HLHS) protein-protein interactome:** Network view of the HLHS interactome is shown as a graph, where genes are shown as nodes and protein-protein interactions (PPIs) as edges connecting the nodes. HLHS-associated genes are shown as dark blue square-shaped nodes, novel interactors and known interactors as red and light blue colored circular nodes respectively. Red edges are the novel interactions, whereas blue edges are known interactions.

Identification of network modules from the HLHS interactome

We identified network modules in the HLHS interactome using Netbox [33], starting with the HLHS genes as core genes and adding nodes from the human interactome. The number of edges of node to core proteins is statistically significant compared to its degree of interactions in the human interactome. It includes all edges between these genes and the core genes, and identifies highly interconnected modules in this network. Netbox connected 143 proteins (48 HLHS candidate gene and 95 linker proteins) into 19 modules, of which 11 modules had 4 or more nodes each (**Data File S2**). Three modules had statistically significant enrichment of GO biological process terms: *TOR signaling* (p-value = 6.97E-04, odds ratio = 27.79), *response to endoplasmic reticulum (ER) stress* (p-value = 2.00E-03, odds ratio = 55.26) and *intracellular receptor signalling pathway* (p-value = 9.44E-14, odds ratio = 20.58) (**Figure 2**). The novel PPIs facilitated the identification of 2 functional modules that may be critical to HLHS pathology, namely, TOR signalling and ER stress [46,47].

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2022 doi:10.20944/preprints202202.0017.v1

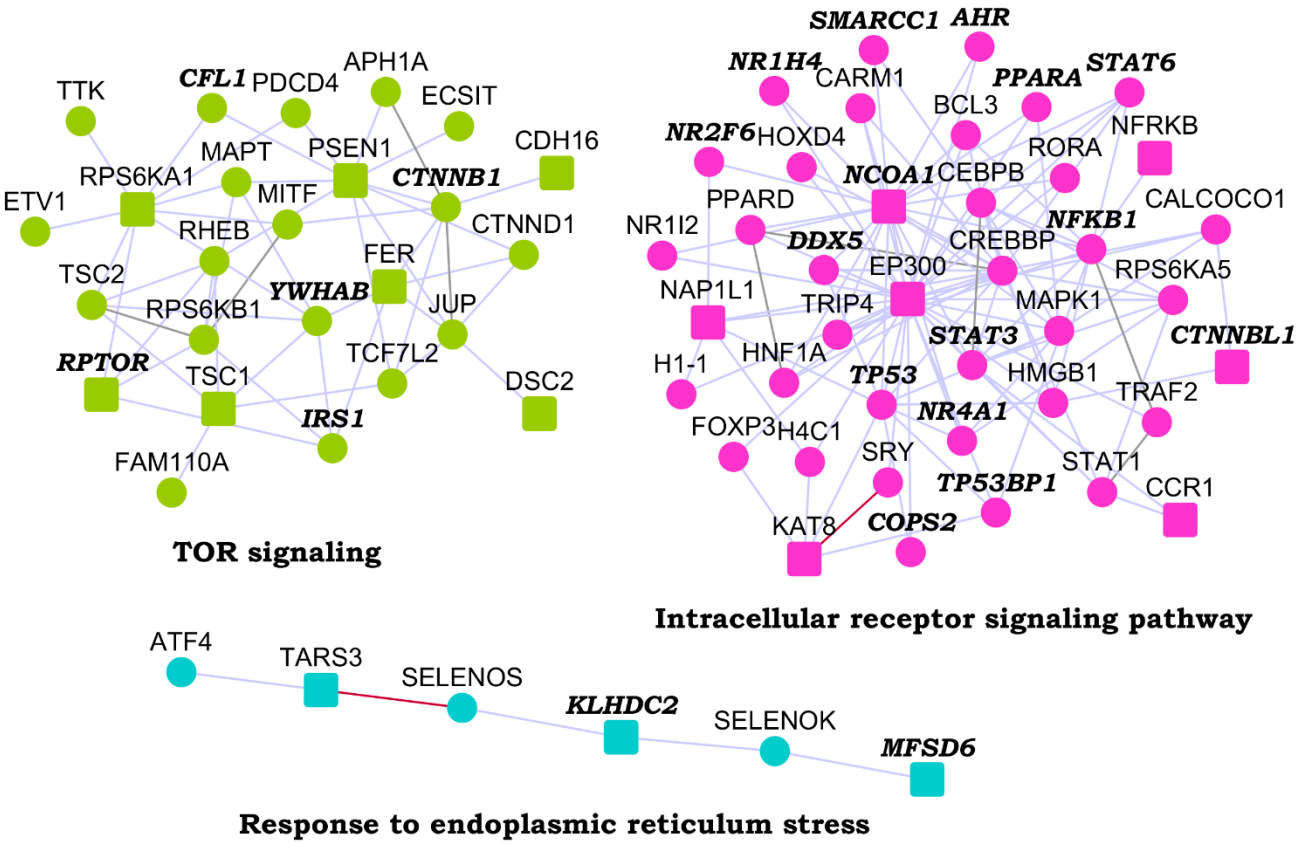

Figure 2. Modules identified from network of the hypoplastic left heart syndrome (HLHS) interactome: Three modules that were enriched in specific GO biological processes are shown. Within each module, nodes with bold italicized labels depict genes with at least one transcriptomic evidence relevant to HLHS. HLHS-associated genes are shown as square-shaped nodes and novel interactors and known interactors are shown as circular nodes. Red edges are the novel interactions, whereas blue edges are known interactions.

Functional enrichment for human diseases in the HLHS interactome

We compiled the list of pathways for proteins in the HLHS interactome that are associated with the Ingenuity Pathway Analysis suite [35]. Selected pathways that are significantly associated with HLHS are shown in **Figure 3** (complete list in **Data File S3**). The Gene Ontology (GO) terms and diseases from OMIM and DisGeNET that are significantly associated with the HLHS interactome at p-value < 0.05 are shown in **Data Files S4-S8**. Examination of OMIM related genes (**Figure 4A**; **Supplemental Data File S7**) showed enrichment associated with *non-insulin-dependent diabetes mellitus* (p-value = 1.44E-07, odds ratio = 14.49) and *insulin-dependent diabetes mellitus* (p-value = 0.02, odds ratio = 15.57) in the HLHS interactome, indicating a mechanistic link between HLHS and disease processes related to energy metabolism. Nine diabetes-associated genes that had direct interactions with 10 HLHS candidates were responsible for this enrichment, including the novel interaction of the diabetes-associated IRS1 with the HLHS candidate NRDC (**Figure 5A**; **Supplemental Data File S7**). Supporting this association, 8.5% of infants born to diabetic mothers have been shown to have congenital heart defects including HLHS, double-outlet right ventricle, truncus arteriosus, transposition of the great arteries and ventricular septal defects [48]. Interestingly, also significantly enriched are genes associated with Alzheimer's disease (AD) (p-value = 1.84E-05, odds ratio=21.23) (**Figure 4A**; **Supplemental Data File S7**), with 5 genes associated with AD exhibiting direct interactions with 9 HLHS candidate genes (**Figure 5B**), supporting recent study showing increased risk of dementia among patients with congenital heart disease [49]. Finally, examination for enrichment in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2022 doi:10.20944/preprints202202.0017.v1

DisGeNET showed marked enrichment for many different types of cancer, with mammary neoplasms, adenocarcinoma and liver carcinoma being the top 3 diseases recovered from DisGeNET (**Figure 4B**; **Supplemental Data File S8**).

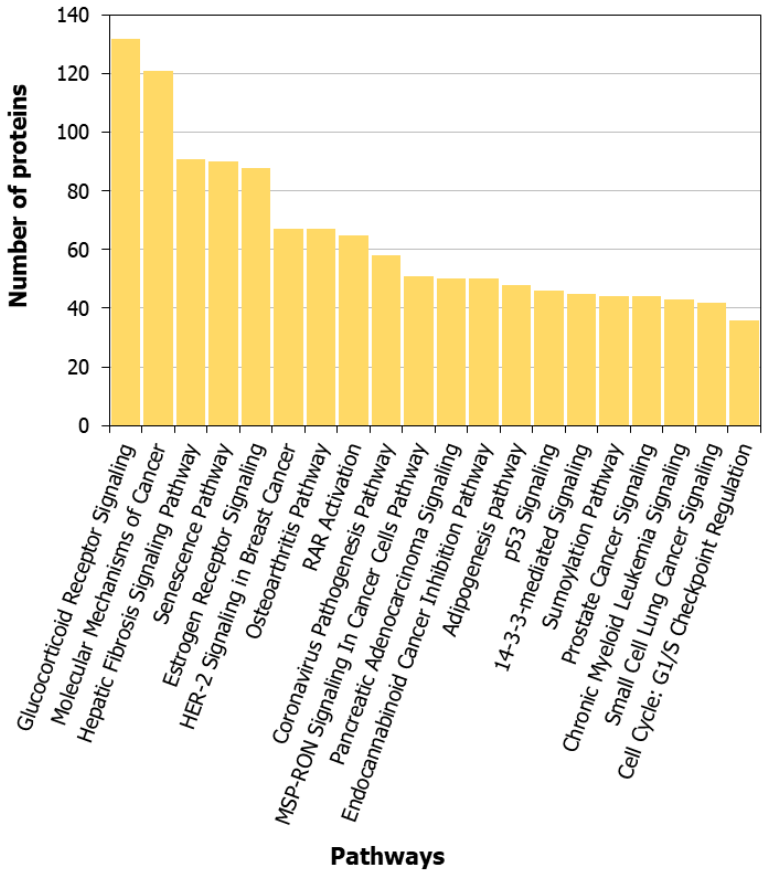

Figure 3. Pathways associated with the hypoplastic left heart **syndrome (HLHS) interactome:** The number of proteins from the HLHS interactome that are involved in the top-30 pathways most significantly associated with the interactome are shown.

A previous autopsy study showed 43% of HLHS patients have hepatic necrosis [50], and more recent studies have indicated a high incidence of hepatocellular carcinoma among patients having had the Fontan procedure, a third stage surgical palliation that all HLHS patients must undergo [51]. 67 genes associated with liver carcinoma exhibit direct interactions with 28 HLHS candidate genes, including 7 novel interactions (listed in the format HLHS candidate gene – liver carcinoma-associated gene: NRDC – IRS1, RPS6KA1 – CXCL12, NIF3L1 – CXCL12, STT3B – GNMT, ZAN – TFPI2, COL15A1 – PDGFB and MFSD6 – STAT1). We noted that of the top ten diseases recovered from DisGeNET, the top nine are cancer related, but the tenth is "insulin resistance", further supporting a link between HLHS and diabetes.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2022 doi:10.20944/preprints202202.0017.v1

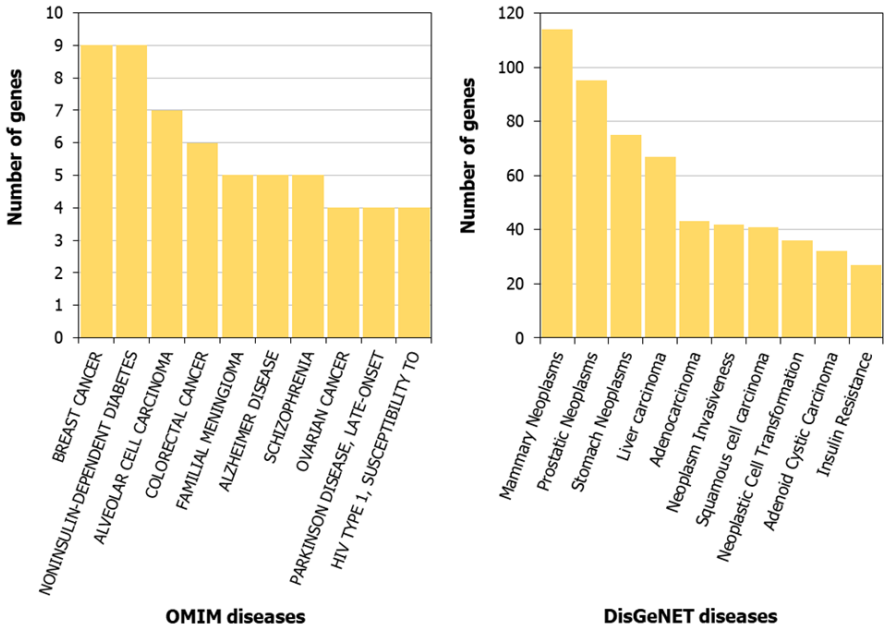

Figure 4. Diseases associated with the hypoplastic left heart syndrome (HLHS) interactome: The number of genes from the HLHS interactome that are involved in the top-10 (A) OMIM diseases and (B) DisGeNET diseases most significantly associated with the interactome are shown.

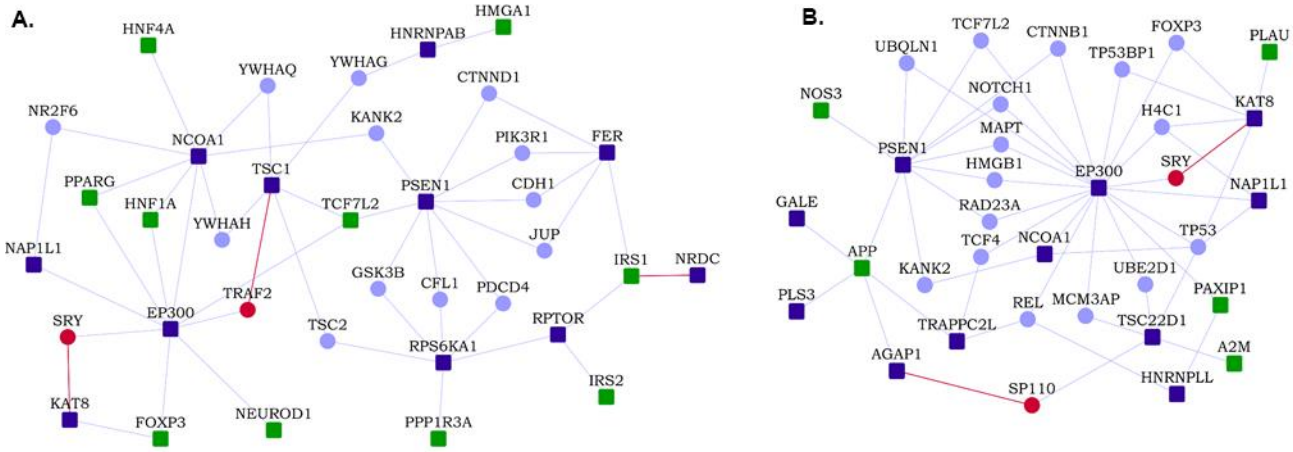

Figure 5. Network proximity of other disease-associated genes to genes associated with hypoplastic left heart syndrome (HLHS): Dark blue square-shaped nodes are HLHS-associated genes and green square-shaped nodes are diabetes-associated genes in (A) and Alzheimer's disease-associated genes in (B). Light blue nodes are known interactors and red nodes are novel interactors. Red edges are the novel interactions, whereas blue edges are known interactions.

GO Biological Process Enrichment and Overlap with HLHS Transcriptome Datasets

Among GO biological process (**Supplemental Data File S4**), the most significantly enriched in the HLHS interactome was *covalent chromatin modification* (p-value < 1E-15, odds ratio = 2.41). this observation was corroborated by the finding that the most significantly enriched GO subcellular locations included *transcription factor complex* (p-value < 1E-06, odds ratio = 3.14) and *nuclear chromatin* (p-value < 1E-06, odds ratio = 2.82), and among the molecular functions (**Supplemental Data File S6**), *DNA-binding transcription activator activity, RNA polymerase II-specific* (p-value < 1E-14, odds ratio = 3.33) and *transcription coactivator activity* (p-value < 1E-14, odds ratio = 3.38). Motivated by the enrichment of transcriptional regulatory processes in our interactome and several previous studies suggesting transcriptomic changes associated with HLHS [17,52,53], we further investigated the overlap of the HLHS interactome with four HLHS-related transcriptomic datasets. We

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2022 doi:10.20944/preprints202202.0017.v1

studied whether the genes in the HLHS interactome were differentially expressed or alternatively spliced in four different RNAseq datasets comprising either tissues or cardiomyocytes derived from induced pluripotent stem cells (iPSCs) (**Supplemental Data File S9**).

Our analysis identified 73 novel interactors (19%) (**Table 1**) and 364 overall genes (21%) in the HLHS interactome had one or more transcriptomic association (**Supplemental Data File S9**). each of these four RNAseq datasets showed considerable, albeit statistically non-significant, overlap with the interactome. The datasets include differential expression in cardiomyocytes differentiated from iPSCs of 5 HLHS patients versus 2 controls (GSE92447 [53]), yielding 131 genes present in the interactome (odds ratio = 1.09), HLHS-right ventricle versus control-left ventricle/control-right ventricle (GSE23959 [52]), yielding 7 genes in the interactome (odds ratio = 2.43), iPSC-derived cardiomyocytes at 25 days from 1 HLHS proband versus parents yielding 131 genes (odds ratio = 1.01) [17], and genes affected by alternative splicing in HLHS-right ventricle versus control-right ventricle/control-left ventricle (GSE23959 [52]), yielding 136 overlapping genes (odds ratio = 1.02). Though these overlaps are not statistically significant at the systems level, the individual genes and their transcriptomic evidence may provide biologically relevant information about the etiology of HLHS.

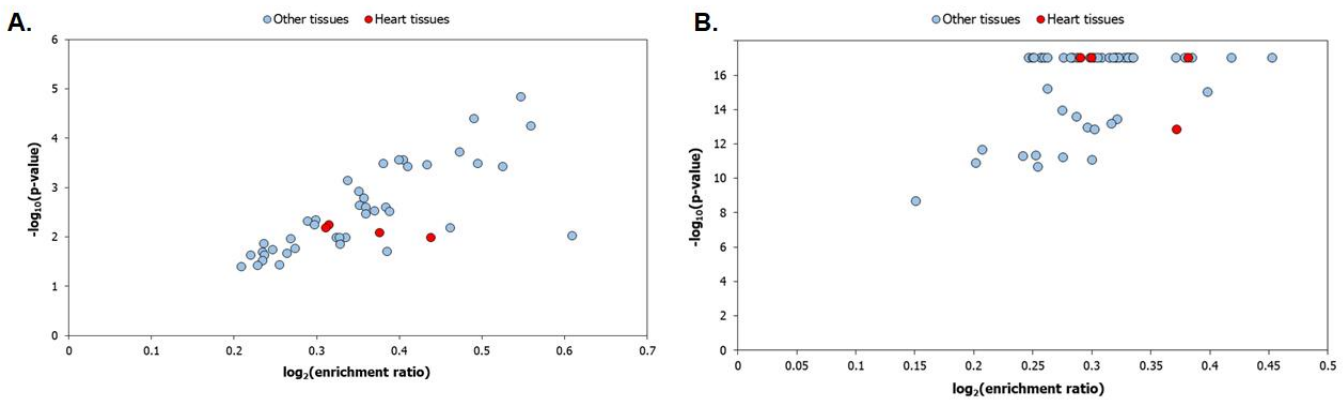

Figure 6. Tissue enrichment of the genes in the hypoplastic left heart syndrome (HLHS) interactome: The tissue enrichment patterns of the HLHS interactome were identified using the gene expression profiles of 53 postnatal human tissues extracted from GTEx. Enrichment was assessed by considering (A) any gene that showed high/medium expression in the tissues (transcripts per million (TPM) ≥ 9) and (B) any gene that showed high/medium expression in the tissues, except for housekeeping genes (detected in all the tissues with TPM ≥ 1). Statistical significance of tissue enrichment was computed using Fisher's exact test, and corrected using the Benjamini-Hochberg method for multiple test adjustment. It can be observed that the HLHS interactome genes showed a higher statistical significance of enrichment in several tissues (including the heart-related tissues shown as red data points) when housekeeping genes were considered as shown in (A) compared to when housekeeping genes were excluded as shown in (B), indicating that housekeeping genes could be overrepresented in the interactome.

We studied the tissue-specific expression of the HLHS interactome genes using RNAseq data of 53 postnatal human tissues obtained from GTEx [42], with and without the inclusion of housekeeping genes from Human Protein Atlas [43]. An expression of more than 9 transcripts per million (TPM) is considered high/medium expression. 9,634 genes detected in all the tissues with TPM ≥ 1 were considered as housekeeping genes. Statistical significance of the enrichment was computed using Fisher's exact test, and corrected using the Benjamini-Hochberg multiple test adjustment. This analysis showed The HLHS interactome genes were significantly enriched in several tissues (**Figure 6A**) – including in heart-related tissues such as the atrial appendage, coronary artery, aorta and the left ventricle – compared to when housekeeping genes were excluded as shown in

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2022 doi:10.20944/preprints202202.0017.v1

Figure 6B. This could indicate that a large number of genes in the HLHS interactome were housekeeping genes. In line with this, we found that 60.5% (1028 genes) of the interactome was comprised of housekeeping genes, a highly statistically significant over-enrichment (p-value = 2.03E-10) of 1.14 fold compared to expectations (906 genes). We also noted that the left ventricle showed a lower statistical significance of enrichment compared with the other three heart-related tissues (atrial appendage, coronary artery and aorta).

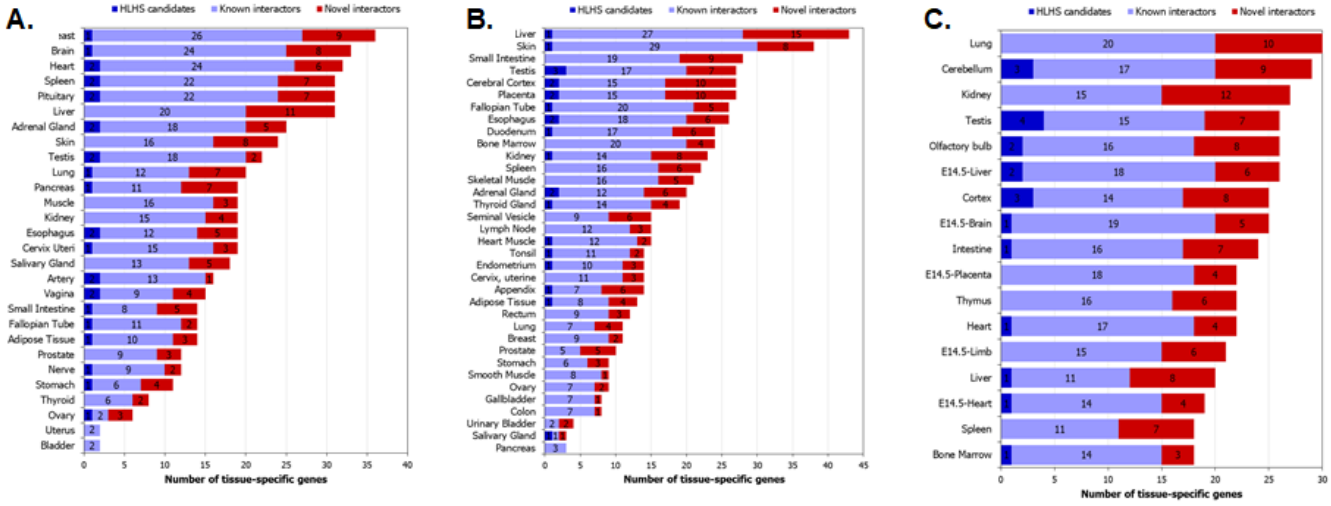

Figure 7. Tissue-specificity of hypoplastic **left heart syndrome (HLHS) interactome genes:** The graphs show the number of genes from the interactome that exhibit tissue-specificity according to data from (A) GTEx, (B) Human Protein Atlas and (C) mouse ENCODE. The genes show at least 5-folds higher expression in a tissue ('tissue-enriched') or a group of 2-7 tissues compared to all the other tissues ('group-enriched').

We further employed the TissueEnrich tool to examine the tissue-specificity of the genes in the HLHS interactome based on expression data from GTEx [42], Human Protein Atlas [43] and Mouse ENCODE [45] (**Figure 7A-C**). Genes with an expression level greater than 1 TPM (transcripts per million) and relative expression at least 5-fold higher in a particular tissue (tissue-enriched) or a group of 2-7 tissues (group-enriched) were considered [54]. As expected from an interactome showing an over-enrichment of housekeeping genes, the HLHS interactome did not show any statistically significant tissue-specific enrichment. However, it was noteworthy that 6 tissues – placenta, skin, liver, lung, brain and testis – showed large overlaps with the HLHS interactome according to data from at least 2 of the databases and appeared among their lists of top-ten tissues (in terms of the number of tissue-specific genes found in the interactome, and not the statistical significance of this overlap) (**Figure 7A-C**). 10 HLHS candidates had novel PPIs with 11 heart-specific proteins across the 3 databases (HLHS candidates are shown in bold): **NFRKB**-OPCML, **NEUROD4**-IL23A, **TSC1**-PAEP, **OIT3**-PLA2G12B, **CDH16**-CDH5, **SLC12A5**-JPH2, **SLC12A5**-MYL9, **DSC2**-FHOD3, **GALE**-PLA2G5, **PLS3**-NRSN1 and **TSPAN15**-ADAMDEC1.

HLHS and developmental delay

Cilia are dynamic projections on cellular surfaces, which detect a wide variety of cues from the environment and transduce signals into the cell to regulate physiological and developmental processes. A genetic screen for recessive CHD-associated mutations had highlighted the role of cilia-transduced cell signalling in CHD [29]. Thirty four of the 61 CHD-associated genes recovered in his screen were cilia-related. To examine for possible ciliary connection to HLHS, we computed the overlap of the HLHS interactome with the interactome of ciliary proteins containing a total of 1,665 proteins and 1,776 PPIs [55]. The interactomes shared a highly statistically significant overlap (p-value = 3.97E-25) of 284

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2022 doi:10.20944/preprints202202.0017.v1

genes, and 30% of the overlapping genes had HLHS-related transcriptomic evidence. The Reactome pathways *gene expression*, *SUMOylation* and *cell cycle* were enriched among the shared genes by 2.4 folds, 6.4 folds and 3.4 folds, respectively.

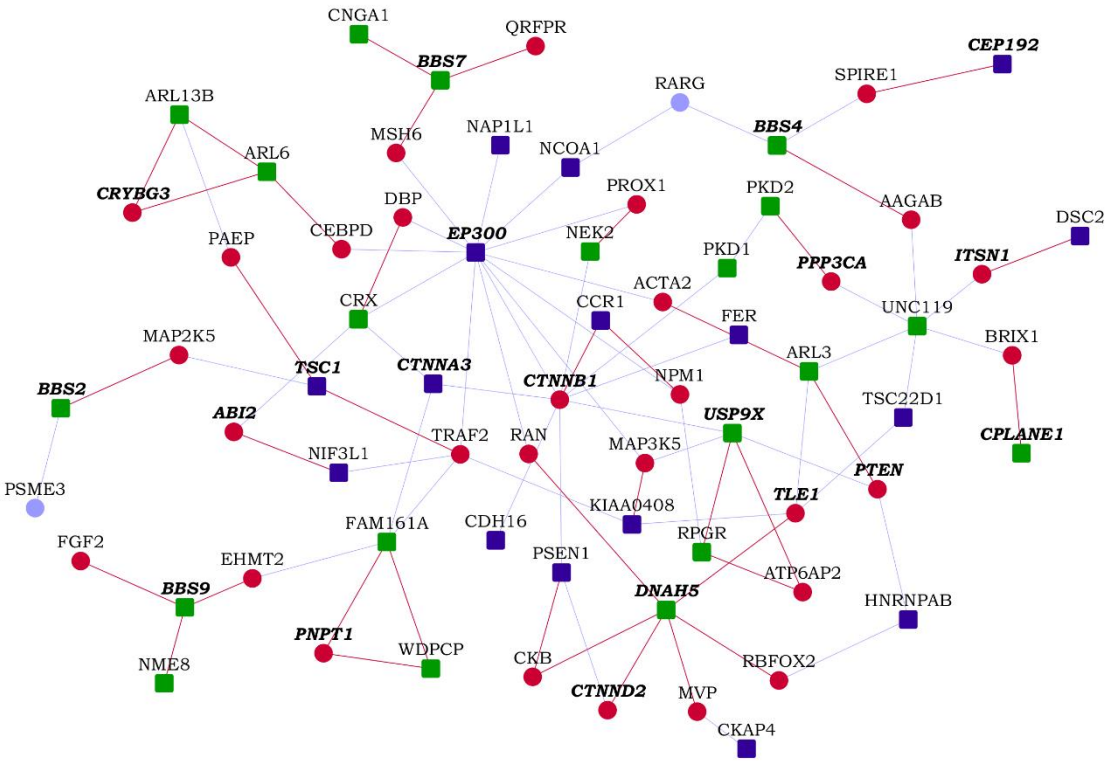

Figure 8. A partial network view of novel protein-protein interactions (PPIs) interconnecting hypoplastic left heart syndrome (HLHS) genes with ciliopathy-associated genes: Genes are shown as nodes and PPIs as edges. As the integrated HLHS and ciliopathy interactome is very large, only a partial view incorporating genes that are associated with intellectual disability and/or developmental delay (ID/DD) and the novel interactors of HLHS-associated genes/ciliopathy-associated genes are shown. Legend – square-shaped dark blue nodes: HLHS-associated genes; square-shaped green nodes: ciliopathy-associated genes; nodes with bold and italicized labels: ID/DD-associated genes; red nodes/edges: novel interactors/interactions; light blue nodes/edges: known interactors/interactions.

Next, we collected a list of 187 genes that have been implicated in 35 ciliopathies from Reiter et al. [56], and assembled its interactome containing 2,486 proteins and 3,022 interactions. 28% of the HLHS interactome overlapped with this interactome (p-value = 1.18E-58; odds-ratio 2) (**Figure 8** and **Data File S10**). 67 HLHS-associated genes, 157 ciliopathy-associated genes and 3 genes associated with both (CCDC65, KIAA0586 and DNAH1) were connected via 841 intermediate interactors. 8 direct known interactions were found between HLHS candidates and ciliopathy-associated genes (HLHS candidates are shown in bold): **TSC22D1**-UNC119, **RPTOR**-CILK1, **MFSD6**-TMEM237, **EP300**-CRX, **CTNNA3**-CRX, **CTNNA3**-FAM161A, **NIF3L1**-NME7 and **TSC1**-GLIS2) and one direct novel interaction (**RPTOR**-CCDC40). We identified the top-30 GO biological processes that were significantly associated with the HLHS and the ciliopathy interactomes and computed the number of genes that were exclusively found in the HLHS/ciliopathy interactomes or shared between the two interactomes in each of these processes (**Figures S1-S2**). *Regulation of DNA-binding transcription factor activity* was enriched 4 folds among the shared genes between the two interactomes (p-value = 3.96E-12). Speculating that transcription factor (TF) activity could be a major factor in the cross talk between HLHS and ciliopathy, we sought to identify the TFs whose target genes were significantly enriched among the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2022 doi:10.20944/preprints202202.0017.v1

genes shared by the HLHS and ciliopathy interactomes. The enrichment analysis was performed using WebGestalt [41] and based on curated TF-target gene sets in MSigDB [40]. The targets of CREBP1 and ALX4 showed significant over-enrichment (p-values of 2.57E-04 and 1.18E-03) of 8.68 folds and 38.85 folds (5 genes and 2 genes) compared to expectations. The targets of CREBP1 found from among the genes shared between the HLHS and ciliopathy interactomes were EP300, PRNP, SMAD3, SUMO1 and TBX6, while the targets of ALX4 were JUN and TCF7L2.

HLHS patients having developmental delay as a comorbidity have been shown to have a higher burden of ciliopathy variants compared to HLHS patients without developmental delay with summative C-score of 4.05 versus 2.02 (p-value of the observed difference < 0.01). Summative C-score is a standardized value used to assess the level of gene disruption in a condition; in this specific study, the C-score of 4.07 was identified as a threshold at which 50% of pathogenic variants and 3% of benign variants were retained [57]. This prompted us to compare the enrichment of genes implicated in developmental delay in the HLHS and ciliopathy interactomes, and specifically among the genes uniquely found in the HLHS/ciliopathy interactomes and those shared between the HLHS and ciliopathy interactomes. 703 genes harbouring loss-of-function and missense variants linked to intellectual disability and/or developmental delay (ID/DD) were collected from the Developmental Brain Disorder Gene Database [58]. These ID/DD genes showed an overlap of higher significance with the HLHS interactome (p-value = 3.77E-08, odds ratio = 1.67) compared with the ciliopathy interactome (p-value = 7.61E-03, odds ratio = 1.23). Additionally, significant enrichment for ID/DD was shown by genes uniquely found in the HLHS interactome (p-value = 7.16E-06, odds ratio = 1.65) and genes shared by the HLHS and ciliopathy interactomes (p-value = 1.93E-03, odds ratio = 1.73) (**Figure 8**), but not by the genes uniquely found in the ciliopathy interactome. The HLHS candidate KIAA0586 was associated with ciliopathy as well as ID/DD. Eight other HLHS candidates were linked to ID/DD (i.e. they also harboured ID/DD-associated variants), namely, TSC1, KDM3B, CEP192, TRAPPC2L, EP300, CTNNA3, KCNQ3 and AP3B2. We identified 19 novel PPIs of HLHS candidates with ID/DD genes (HLHS candidates are shown in bold): **SERPINB7**-OGDH, **SERPINB7**-PIGN, **NOMO1**-NDE1, **NOMO1**-KCNA1, **NCOA1**-PPP1CB, **FER**-HSD17B4, **CKAP4**-POLR3B, **AMT**-IMPDH2, **RPS6KA1**-ASCC3, **STT3B**-SMARCC1, **DSC2**-ITSN1, **AGAP1**-SLC19A3, **OXNAD1**-SLC6A1, **MFSD6**-HIBCH, **KAT8**-PRRT2, **HNRNPAB**-SYNCRIP, **NIF3L1**-ABI2, **PLS3**-CUL4B and **CCR1**-CTNNB1.

These results implicate the cilium as a potential focal point for examining HLHS etiology and its comorbid relationships with ciliopathy, intellectual disability and developmental delay.

HLHS and microcephaly

Severe neurological outcomes such as seizure activity, ischemia, and haemorrhage in HLHS patients is more prevalent with neonatal microcephaly than without (43% versus 4%, p-value = 0.02); the prevalence is 33% in HLHS patients with fetal microcephaly (p-value = 0.06) [59]. This prompted us to examine their interconnections. 84 genes associated with microcephaly were collected from the MONARCH database [60] and the microcephaly interactome containing 1,867 proteins and 2,081 interactions was assembled. 62 HLHS candidates were connected to 77 microcephaly genes via 652 intermediate interactors. 5 direct known interactions were found between HLHS candidates and microcephaly-associated genes (HLHS candidates are shown in bold): **TSC1**-POGZ, **TSC1**-CDK6, **CTNNBL1**-STAMBP, **TRAPPC2L**-TRAPPC6B and **TSC22D1**-QARS1. 24% of the HLHS interactome overlapped with 22% of the microcephaly interactome (405 genes), a highly statistically significant overlap (p-value = 2.39E-65) with an enrichment ratio of 2.31-fold compared to expectations (176 genes) (**Figure 9** and **Data File S11**). We identified the top-30 GO biological processes that were significantly associated with the HLHS and the microcephaly interactomes and computed the number of genes that were exclusively found in the HLHS/microcephaly interactomes or shared between the two interactomes in each of these processes (**Figures S3-S4**). *Neuron death* was enriched 4-fold (32 genes) among the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 February 2022 doi:10.20944/preprints202202.0017.v1

genes common to the HLHS and microcephaly interactomes (p-value = 5.13E-10). It was also interesting to note that neuron death, neurodegeneration and other related processes and diseases appeared as enriched terms in the HLHS interactome across several functional categories. We extracted a total of 95 genes from the HLHS interactome belonging to these categories, specifically, 5 genes linked to Alzheimer's disease (OMIM ID: 104300) and 4 genes in late-onset Parkinson's disease (OMIM ID: 168600), 5 genes linked to neurofibrillary degeneration (DisGeNET ID: C0085400) and 17 genes in neurodegenerative disorders (DisGeNET ID: C0524851) and 79 genes linked to neuron death (GO ID: 0070997). Two pathways showed high statistical significance and enrichment ratio among these 95 genes, namely, *constitutive signaling by AKT1 E17K in cancer* (p-value = 5.94E-07, odds ratio = 32.83) and *intrinsic pathway for apoptosis* (p-value = 6.94E-07, odds ratio = 21.32 folds). 10 novel interactors of HLHS candidates were found among these 95 genes (ITSN1, RHOA, NQO1, CTNNB1, TRAF2, MAP3K5, PRPH, UBQLN2, UNC5B and FUS). Five of these novel interactors (ITSN1, RHOA, TRAF2, CTNNB1 and UNC5B) seemed to be involved in *death receptor signaling/apoptosis*, *axon guidance/EPH-Ephrin signaling and/or developmental biology*, and their novel PPIs with HLHS candidates were as follows (HLHS candidates are shown in bold): **DSC2**-ITSN1, **AMT**-RHOA, **CCR1**-CTNNB1, **TSC1**-TRAF2 and **OIT3**-UNC5B.

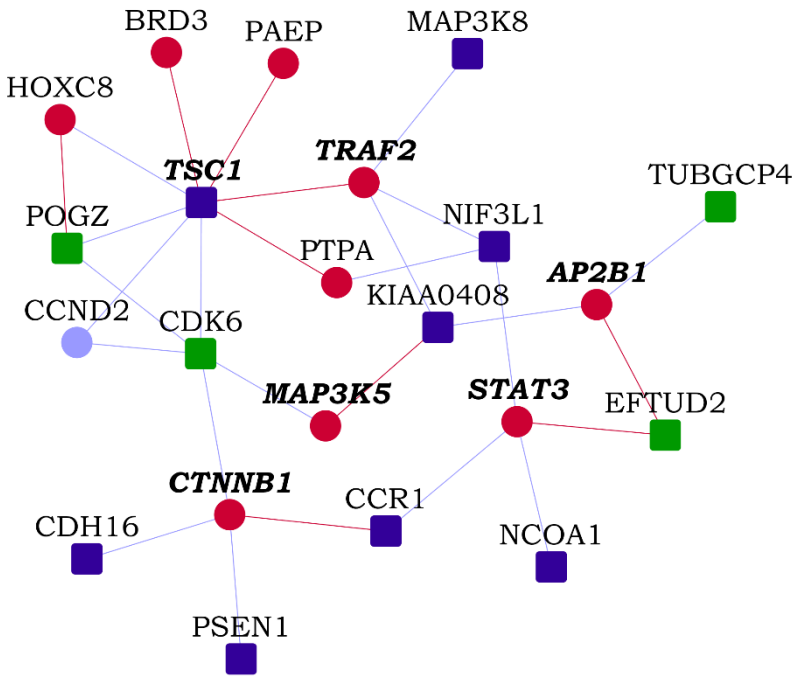

Figure 9. A partial network view of novel protein-protein interactions (PPIs) interconnecting hypoplastic left heart syndrome (HLHS) genes with microcephaly-associated genes: Genes are shown as nodes and PPIs as edges. As the integrated HLHS and microcephaly interactome is very large, only a partial view incorporating genes that are involved in neuronal death processes and the novel interactors of HLHS-associated genes/microcephaly-associated genes are shown. Legend – square-shaped dark blue nodes: HLHS-associated genes; square-shaped green nodes: microcephaly-associated genes; nodes with bold and italicized labels: genes involved in neuron death; red nodes/edges: novel interactors/interactions; light blue nodes/edges: known interactors/interactions.

Genes in the HLHS interactome linked to neuronal death processes may serve as potential candidates for examining the genetic basis of microcephaly in HLHS patients and the increased prevalence of poor neurological outcomes in these patients. The intriguing link between neurodegenerative processes and HLHS is another result that warrants closer inspection especially in light of the recent finding that adults with congenital heart

 doi:10.20944/preprints202202.0017.v1

disease show an increased risk of dementia and early onset dementia, particularly amongst patients with complex lesions [49].

Discussion

In this study, we adopted a protein interactome analysis approach to study HLHS-associated genes. Interactome analysis framework postulates that diseases develop when PPIs are perturbed by genetic mutations or aberrant expression of genes/proteins, ultimately leading to disrupted cellular functions [61]. Extensive inter-connectivity and intra-connectivity of the network components in the PPI network suggest that the effects of such perturbations may spread to other proteins, encoded by genes that do not harbor any disease-associated alterations, through the network of their interactions, posing deeper implications for disease development [61]. In this study, we assembled the HLHS interactome by supplementing previously known protein PPIs with computationally predicted PPIs, that are deemed accurate, and provided valuable insights into etiology through network and enrichment analysis.

We make the PPIs, including the novel PPIs, available on a searchable webserver to enable biologists to study the PPI of their interest (http://severus.dbmi.pitt.edu/wiki-HLHS). Our website provides advanced search capabilities which allows a user to ask questions that will help generate testable hypotheses around individual PPIs. The full-text of the PPIs and their annotations will also be indexed in internet search engines, so that biologists searching in Google, Bing, etc. will find this content. System-level analysis of the interactome with transcriptomic or proteomic data may help to identify its functional landscape. Investigation of individual PPIs will accelerate the understanding of disease biology by several years.

More than 60% of the HLHS interactome, including 51% (38) of the HLHS-associated genes (or 'core genes' used for interactome construction), was composed of genes that are constitutively expressed in all the tissues (i.e. housekeeping genes). It has been reported that none of the HLHS-associated mutations harboured by the core genes was shared among the HLHS mutant mouse lines [12]. The preponderance of housekeeping genes among the core genes as well as the HLHS interactome as a whole could explain this genetic heterogeneity. The transmission of mutations in housekeeping genes may be stymied due to their roles in sustaining essential cellular functions, whose perturbation may result in lethality or reduction in reproductive fitness [1]. Further interactome-based investigations driven by this observation such as the vulnerability to network perturbations, and compensatory mechanisms counteracting them, may provide interesting insights.

The HLHS interactome was not enriched for specific expression in any tissues, as can be expected from the over-enrichment of housekeeping genes in the interactome. However, it was noteworthy that 3 out of the 5 tissues that contained the greatest number of tissue-specific interactome genes were also documented as sites of extracardiac anomalies in HLHS, namely, the placenta, liver and brain. Specifically, placental abnormalities have been noted in pregnancies that involved fetal HLHS [62]. Increased occurrence of hepatic necrosis has been noted among patients with infantile coarctation of the aorta and HLHS compared with patients having other cardiac defects (38% versus 6%) [63]. The prevalence of brain abnormalities among HLHS neonates and survivors is well-documented [8,9].

We showed that the interactome of ciliopathy-associated genes shared a significant overlap with the HLHS interactome and that transcription regulation may be over-enriched among these common genes. The targets of transcription factors CREBP1 and ALX4 were identified to be significantly enriched among the shared genes. CREBP1 (also known an ATF2) has been shown to regulate the expression of 5 genes, namely, EP300 (an HLHS-associated gene sharing a direct interaction with the ciliopathy gene CRX), SUMO1 (a shared interactor having a novel PPI with the ciliopathy-associated gene MAK and a known PPI with the HLHS-associated gene NCOA1) and 3 known shared interactors of

 doi:10.20944/preprints202202.0017.v1

HLHS and ciliopathy genes (PRNP, SMAD3 and TBX6). ATF2 is involved in cardiomyocyte differentiation [64]. Future studies could concentrate on the role played by ATF2 and its targets in the shared etiology of ciliopathies and HLHS.

We also showed the preferential enrichment of genes involved in intellectual disability and/or developmental delay (ID/DD) among genes unique to the HLHS interactome and genes shared between the HLHS and ciliopathy interactomes (in comparison with genes unique to the ciliopathy interactome). This finding was in line with the observation of increased ciliopathy variant-burden among HLHS patients with developmental delay [57]. Additionally, we provided a list of 19 direct novel PPIs between HLHS-associated and ID/DD genes that may be biochemically validated. For example, OGDH is an ID/DD gene that is critical to the tricarboxylic acid cycle and found in the mitochondrial matrix. Loss of OGDH has been shown to lead to neurodegeneration [65]. This gene shows high expression in the left ventricle and in brain regions such as olfactory bulb, hippocampus, cerebellum and pons [66]. This evidence support OGDH as a potential candidate for future studies on the comorbidity of HLHS and ID/DD.

We predicted 5 direct novel PPIs between HLHS and microcephaly associated genes. Also, genes associated with HLHS and microcephaly share several common interactors that are significantly enriched for neuronal death pathways. This suggests a mechanistic basis for their comorbidity and the increased prevalence of neurological abnormalities among HLHS patients with microcephaly [59]. The overrepresentation of neurodegenerative disease associated genes and processes in the HLHS interactome should be investigated, with a focus on the potentially pleiotropic roles of the AKT1-mediated pathways and the intrinsic apoptotic pathway in HLHS and neurodegeneration. The ten direct PPIs between HLHS and diabetes associated genes can be used to examine their joint genetic basis and the increased risk of developing HLHS seen among infants born to diabetic mothers [48].

In summary, our study provides evidence for the utility of the HLHS interactome in investigating various HLHS comorbidities and the functional consequences of genes harbouring HLHS associated mutations. These results will directly inform and catalyze future investigations on the molecular basis of HLHS and biomedical studies seeking to improve clinical interventions in HLHS.

Conclusions

Knowledge on the exact mechanistic basis of HLHS is limited despite a steady increase in the generation of CHD and HLHS related data. In this scenario, the HLHS interactome will serve as a functional landscape to integrate and analyze publicly available HLHS related multi-omics data and generate new hypotheses that will allow biologists to prioritize pathways and drugs for experimental testing and the developmental of new avenues for therapeutic interventions. To facilitate analysis by both computational and biomedical scientists, the HLHS interactome is being released via an interactive webserver called Wiki-HLHS.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1: Figure S1: Top-30 Gene Ontology Biological Processes associated with the hypoplastic left heart syndrome (HLHS) interactome in relation with the ciliopathy interactome, Figure S2: Top-30 Gene Ontology Biological Processes associated with the ciliopathy interactome in relation with the hypoplastic left heart syndrome (HLHS) interactome, Figure S3: Top-30 Gene Ontology Biological Processes associated with the hypoplastic left heart syndrome (HLHS) interactome in relation with the microcephaly interactome, Figure S4: Top-30 Gene Ontology Biological Processes associated with the microcephaly interactome in relation with the hypoplastic left heart syndrome (HLHS) interactome, Data File S1: List of genes from the HLHS interactome, with their labels (HLHS candidate genes, known interactors and novel interactors), and the list of interactions, with their labels (known and novel interactions), Data File S2: List of modules detected in the HLHS interactome, Data File S3: List of all the pathways associated with at least one of the HLHS interactome genes, along with their statistical significance of association (with Bonferroni correction), Data File S4: List of all the Gene Ontology Biological Process terms significantly associated with the HLHS interactome, Data File S5: List of all

 doi:10.20944/preprints202202.0017.v1

the Gene Ontology Cellular Component terms significantly associated with the HLHS interactome, Data File S6: List of all the Gene Ontology Molecular Function terms significantly associated with the HLHS interactome, Data File S7: List of all the OMIM diseases significantly associated with the HLHS interactome, Data File S8: List of all the DisGeNET diseases significantly associated with the HLHS interactome, Data File S9: Complete list of HLHS-related biological evidence of genes in the HLHS protein interactome, Data File S10: 473 genes shared between the hypoplastic left heart syndrome (HLHS) interactome and the ciliopathy interactome, and Data File S11: 405 genes shared between the hypoplastic left heart syndrome (HLHS) interactome and the microcephaly interactome.

Author Contributions: Conceptualization, CWL and MG; methodology KBK, NB, MG; software, and validation, MG; resources, MG, NB and CWL; data curation, KBK and MG; writing—original draft preparation, KBK; writing—review and editing, GCG, NB, CWL, MG and KBK; supervision, MG; project administration, MG and CWL. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by NIH grant HL14278 (CWL).

Data Availability Statement: On journal website and at http://severus.dbmi.pitt.edu/wiki-HLHS

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zaidi, S.; Brueckner, M. Genetics and genomics of congenital heart disease. *Circulation research* **2017**, *120*, 923-940.
2. Gobergs, R.; Salputra, E.; Lubaua, I. Hypoplastic left heart syndrome: a review. *Acta medica Lituanica* **2016**, *23*, 86.
3. Šamánek, M.; Slavík, Z.; Zbořilová, B.; Hroboňová, V.; Voříšková, M.; Škovránek, J. Prevalence, treatment, and outcome of heart disease in live-born children: a prospective analysis of 91,823 live-born children. *Pediatric cardiology* **1989**, *10*, 205-211.
4. Hamzah, M.; Othman, H.F.; Baloglu, O.; Aly, H. Outcomes of hypoplastic left heart syndrome: analysis of National Inpatient Sample Database 1998–2004 versus 2005–2014. *European journal of pediatrics* **2020**, *179*, 309-316.
5. d'Udekem, Y.; Iyengar, A.J.; Galati, J.C.; Forsdick, V.; Weintraub, R.G.; Wheaton, G.R.; Bullock, A.; Justo, R.N.; Grigg, L.E.; Sholler, G.F. Redefining expectations of long-term survival after the Fontan procedure: twenty-five years of follow-up from the entire population of Australia and New Zealand. *Circulation* **2014**, *130*, S32-S38.
6. Alsoufi, B.; Mori, M.; Gillespie, S.; Schlosser, B.; Slesnick, T.; Kogon, B.; Kim, D.; Sachdeva, R.; Kanter, K. Impact of patient characteristics and anatomy on results of Norwood operation for hypoplastic left heart syndrome. *The Annals of thoracic surgery* **2015**, *100*, 591-598.
7. Siffel, C.; Riehle-Colarusso, T.; Oster, M.E.; Correa, A. Survival of children with hypoplastic left heart syndrome. *Pediatrics* **2015**, *136*, e864-e870.
8. Marino, B.S.; Lipkin, P.H.; Newburger, J.W.; Peacock, G.; Gerdes, M.; Gaynor, J.W.; Mussatto, K.A.; Uzark, K.; Goldberg, C.S.; Johnson Jr, W.H. Neurodevelopmental outcomes in children with congenital heart disease: evaluation and management: a scientific statement from the American Heart Association. *Circulation* **2012**, *126*, 1143-1172.
9. Hinton, R.B.; Andelfinger, G.; Sekar, P.; Hinton, A.C.; Gendron, R.L.; Michelfelder, E.C.; Robitaille, Y.; Benson, D.W. Prenatal head growth and white matter injury in hypoplastic left heart syndrome. *Pediatric research* **2008**, *64*, 364-369.
10. Hinton, R.B.; Martin, L.J.; Tabangin, M.E.; Mazwi, M.L.; Cripe, L.H.; Benson, D.W. Hypoplastic left heart syndrome is heritable. *Journal of the American College of Cardiology* **2007**, *50*, 1590-1595.
11. McBride, K.L.; Pignatelli, R.; Lewin, M.; Ho, T.; Fernbach, S.; Menesses, A.; Lam, W.; Leal, S.M.; Kaplan, N.; Schliekelman, P. Inheritance analysis of congenital left ventricular outflow tract obstruction malformations: segregation, multiplex relative risk, and heritability. *American journal of medical genetics Part A* **2005**, *134*, 180-186.
12. Liu, X.; Yagi, H.; Saeed, S.; Bais, A.S.; Gabriel, G.C.; Chen, Z.; Peterson, K.A.; Li, Y.; Schwartz, M.C.; Reynolds, W.T. The complex genetics of hypoplastic left heart syndrome. *Nature genetics* **2017**, *49*, 1152.

13. McBride, K.L.; Zender, G.A.; Fitzgerald-Butt, S.M.; Koehler, D.; Menesses-Diaz, A.; Fernbach, S.; Lee, K.; Towbin, J.A.; Leal, S.; Belmont, J.W. Linkage analysis of left ventricular outflow tract malformations (aortic valve stenosis, coarctation of the aorta, and hypoplastic left heart syndrome). *European journal of human genetics* **2009**, *17*, 811-819.
14. Zaidi, S.; Choi, M.; Wakimoto, H.; Ma, L.; Jiang, J.; Overton, J.D.; Romano-Adesman, A.; Bjornson, R.D.; Breitbart, R.E.; Brown, K.K. De novo mutations in histone-modifying genes in congenital heart disease. *Nature* **2013**, *498*, 220-223.
15. Theis, J.L.; Hu, J.J.; Sundsbak, R.S.; Evans, J.M.; Bamlet, W.R.; Qureshi, M.Y.; O'Leary, P.W.; Olson, T.M. Genetic Association between Hypoplastic Left Heart Syndrome and Cardiomyopathies. *Circulation: Genomic and Precision Medicine* **2021**, *14*, e003126.
16. Reuter, M.S.; Chaturvedi, R.R.; Liston, E.; Manshaei, R.; Aul, R.B.; Bowdin, S.; Cohn, I.; Curtis, M.; Dhir, P.; Hayeems, R.Z. The Cardiac Genome Clinic: implementing genome sequencing in pediatric heart disease. *Genetics in Medicine* **2020**, *22*, 1015-1024.
17. Theis, J.L.; Vogler, G.; Missinato, M.A.; Li, X.; Nielsen, T.; Zeng, X.-X.I.; Martinez-Fernandez, A.; Walls, S.M.; Kervadec, A.; Kezos, J.N. Patient-specific genomics and cross-species functional analysis implicate LRP2 in hypoplastic left heart syndrome. *Elife* **2020**, *9*, e59554.
18. Iascone, M.; Ciccone, R.; Galletti, L.; Marchetti, D.; Seddio, F.; Lincesso, A.; Pezzoli, L.; Vetro, A.; Barachetti, D.; Boni, L. Identification of de novo mutations and rare variants in hypoplastic left heart syndrome. *Clinical genetics* **2012**, *81*, 542-554.
19. Verma, S.K.; Deshmukh, V.; Nutter, C.A.; Jaworski, E.; Jin, W.; Wadhwa, L.; Abata, J.; Ricci, M.; Lincoln, J.; Martin, J.F. Rbfox2 function in RNA metabolism is impaired in hypoplastic left heart syndrome patient hearts. *Scientific reports* **2016**, *6*, 1-10.
20. Gill, H.K.; Parsons, S.R.; Spalluto, C.; Davies, A.F.; Knorz, V.J.; Burlinson, C.E.; Ng, B.L.; Carter, N.P.; Ogilvie, C.M.; Wilson, D.I. Separation of the PROX1 gene from upstream conserved elements in a complex inversion/translocation patient with hypoplastic left heart. *European journal of human genetics* **2009**, *17*, 1423-1431.
21. Homsy, J.; Zaidi, S.; Shen, Y.; Ware, J.S.; Samocha, K.E.; Karczewski, K.J.; DePalma, S.R.; McKean, D.; Wakimoto, H.; Gorham, J. De novo mutations in congenital heart disease with neurodevelopmental and other congenital anomalies. *Science* **2015**, *350*, 1262-1266.
22. Ganapathiraju, M.K.; Thahir, M.; Handen, A.; Sarkar, S.N.; Sweet, R.A.; Nimgaonkar, V.L.; Loscher, C.E.; Bauer, E.M.; Chaparala, S. Schizophrenia interactome with 504 novel protein–protein interactions. *NPJ schizophrenia* **2016**, *2*, 1-10.
23. Lim, J.; Hao, T.; Shaw, C.; Patel, A.J.; Szabó, G.; Rual, J.-F.; Fisk, C.J.; Li, N.; Smolyar, A.; Hill, D.E. A protein–protein interaction network for human inherited ataxias and disorders of Purkinje cell degeneration. *Cell* **2006**, *125*, 801-814.
24. Sakai, Y.; Shaw, C.A.; Dawson, B.C.; Dugas, D.V.; Al-Mohtaseb, Z.; Hill, D.E.; Zoghbi, H.Y. Protein interactome reveals converging molecular pathways among autism disorders. *Science translational medicine* **2011**, *3*, 86ra49-86ra49.
25. Keshava Prasad, T.; Goel, R.; Kandasamy, K.; Keerthikumar, S.; Kumar, S.; Mathivanan, S.; Telikicherla, D.; Raju, R.; Shafreen, B.; Venugopal, A. Human protein reference database—2009 update. *Nucleic acids research* **2008**, *37*, D767-D772.
26. Stark, C.; Breitkreutz, B.-J.; Reguly, T.; Boucher, L.; Breitkreutz, A.; Tyers, M. BioGRID: a general repository for interaction datasets. *Nucleic acids research* **2006**, *34*, D535-D539.
27. Zhu, J.; Zhang, Y.; Ghosh, A.; Cuevas, R.A.; Forero, A.; Dhar, J.; Ibsen, M.S.; Schmid-Burgk, J.L.; Schmidt, T.; Ganapathiraju, M.K. Antiviral activity of human OASL protein is mediated by enhancing signaling of the RIG-I RNA sensor. *Immunity* **2014**, *40*, 936-948.
28. Karunakaran, K.B.; Yanamala, N.; Boyce, G.; Becich, M.J.; Ganapathiraju, M.K. Malignant Pleural Mesothelioma Interactome with 364 Novel Protein-Protein Interactions. *Cancers* **2021**, *13*, 1660.
29. Li, Y.; Klena, N.T.; Gabriel, G.C.; Liu, X.; Kim, A.J.; Lemke, K.; Chen, Y.; Chatterjee, B.; Devine, W.; Damerla, R.R. Global genetic analysis in mice unveils central role for cilia in congenital heart disease. *Nature* **2015**, *521*, 520-524.

30. Orii, N.; Ganapathiraju, M.K. Wiki-pi: a web-server of annotated human protein-protein interactions to aid in discovery of protein function. *PloS one* **2012**, *7*, e49029.
31. Ganapathiraju, M.K.; Karunakaran, K.B.; Correa-Menendez, J. Predicted protein interactions of IFITMs may shed light on mechanisms of Zika virus-induced microcephaly and host invasion. *F1000Res* **2016**, *5*, 1919, doi:10.12688/f1000research.9364.2.
32. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* **2003**, *13*, 2498-2504.
33. Cerami, E.; Demir, E.; Schultz, N.; Taylor, B.S.; Sander, C. Automated network analysis identifies core pathways in glioblastoma. *PloS one* **2010**, *5*.
34. Wang, Z.; Zhang, J. In search of the biological significance of modular structures in protein networks. *PLoS computational biology* **2007**, *3*.
35. Kramer, A.; Green, J.; Pollard, J., Jr.; Tugendreich, S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **2014**, *30*, 523-530, doi:10.1093/bioinformatics/btt703.
36. Consortium, G.O. The Gene Ontology (GO) database and informatics resource. *Nucleic acids research* **2004**, *32*, D258-D261.
37. Croft, D.; Mundo, A.F.; Haw, R.; Milacic, M.; Weiser, J.; Wu, G.; Caudy, M.; Garapati, P.; Gillespie, M.; Kamdar, M.R. The Reactome pathway knowledgebase. *Nucleic acids research* **2014**, *42*, D472-D477.
38. Hamosh, A.; Scott, A.F.; Amberger, J.S.; Bocchini, C.A.; McKusick, V.A. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. *Nucleic acids research* **2005**, *33*, D514-D517.
39. Piñero, J.; Bravo, À.; Queralt-Rosinach, N.; Gutiérrez-Sacristán, A.; Deu-Pons, J.; Centeno, E.; García-García, J.; Sanz, F.; Furlong, L.I. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic acids research* **2016**, gkw943.
40. Liberzon, A.; Subramanian, A.; Pinchback, R.; Thorvaldsdóttir, H.; Tamayo, P.; Mesirov, J.P. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **2011**, *27*, 1739-1740.
41. Liao, Y.; Wang, J.; Jaehnig, E.J.; Shi, Z.; Zhang, B. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic acids research* **2019**.
42. Consortium, G. The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science* **2015**, *348*, 648-660.
43. Uhlén, M.; Fagerberg, L.; Hallström, B.M.; Lindskog, C.; Oksvold, P.; Mardinoglu, A.; Sivertsson, Å.; Kampf, C.; Sjöstedt, E.; Asplund, A. Tissue-based map of the human proteome. *Science* **2015**, *347*.
44. Jain, A.; Tuteja, G. TissueEnrich: Tissue-specific gene enrichment analysis. *Bioinformatics* **2019**, *35*, 1966-1967, doi:10.1093/bioinformatics/bty890.
45. Stamatoyannopoulos, J.A.; Snyder, M.; Hardison, R.; Ren, B.; Gingeras, T.; Gilbert, D.M.; Groudine, M.; Bender, M.; Kaul, R.; Canfield, T. An encyclopedia of mouse DNA elements (Mouse ENCODE). *Genome biology* **2012**, *13*, 1-5.
46. Gaber, N.; Gagliardi, M.; Patel, P.; Kinnear, C.; Zhang, C.; Chitayat, D.; Shannon, P.; Jaeggi, E.; Tabori, U.; Keller, G. Fetal reprogramming and senescence in hypoplastic left heart syndrome and in human pluripotent stem cells during cardiac differentiation. *The American journal of pathology* **2013**, *183*, 720-734.
47. Xu, X.; Jin, K.; Bais, A.S.; Zhu, W.; Yagi, H.; Feinstein, T.N.; Nguyen, P.; Criscione, J.; Liu, X.; Beutner, G. iPSC modeling shows uncompensated mitochondrial mediated oxidative stress underlies early heart failure in hypoplastic left heart syndrome. *bioRxiv* **2021**.
48. Becerra, J.E.; Khoury, M.J.; Cordero, J.F.; Erickson, J.D. Diabetes mellitus during pregnancy and the risks for specific birth defects: a population-based case-control study. *Pediatrics* **1990**, *85*, 1-9.
49. Bagge, C.N.; Henderson, V.W.; Laursen, H.B.; Adelborg, K.; Olsen, M.; Madsen, N.L. Risk of dementia in adults with congenital heart disease: population-based cohort study. *Circulation* **2018**, *137*, 1912-1920.

50. Komatsu, H.; Inui, A.; Kishiki, K.; Kawai, H.; Yoshio, S.; Osawa, Y.; Kanto, T.; Fujisawa, T. Liver disease secondary to congenital heart disease in children. *Expert review of gastroenterology & hepatology* **2019**, *13*, 651-666.
51. Kogiso, T.; Tokushige, K. Fontan-associated liver disease and hepatocellular carcinoma in adults. *Scientific Reports* **2020**, *10*, 1-14.
52. Ricci, M.; Xu, Y.; Hammond, H.L.; Willoughby, D.A.; Nathanson, L.; Rodriguez, M.M.; Vatta, M.; Lipshultz, S.E.; Lincoln, J. Myocardial alternative RNA splicing and gene expression profiling in early stage hypoplastic left heart syndrome. *PloS one* **2012**, *7*, e29784.
53. Yang, C.; Xu, Y.; Yu, M.; Lee, D.; Alharti, S.; Hellen, N.; Ahmad Shaik, N.; Banaganapalli, B.; Sheikh Ali Mohamoud, H.; Elango, R. Induced pluripotent stem cell modelling of HLHS underlines the contribution of dysfunctional NOTCH signalling to impaired cardiogenesis. *Human molecular genetics* **2017**, *26*, 3031-3045.
54. Fagerberg, L.; Hallström, B.M.; Oksvold, P.; Kampf, C.; Djureinovic, D.; Odeberg, J.; Habuka, M.; Tahmasebpoor, S.; Danielsson, A.; Edlund, K. Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Molecular & Cellular Proteomics* **2014**, *13*, 397-406.
55. Karunakaran, K.B.; Chaparala, S.; Lo, C.W.; Ganapathiraju, M.K. cilia interactome with predicted protein–protein interactions reveals connections to Alzheimer's disease, aging and other neuropsychiatric processes. *Scientific reports* **2020**, *10*, 1-16.
56. Reiter, J.F.; Leroux, M.R. Genes and molecular pathways underpinning ciliopathies. *Nature reviews Molecular cell biology* **2017**, *18*, 533-547.
57. Geddes, G.C.; Stamm, K.; Mitchell, M.; Mussatto, K.A.; Tomita-Mitchell, A. Ciliopathy variant burden and developmental delay in children with hypoplastic left heart syndrome. *Genetics in Medicine* **2017**, *19*, 711-714.
58. Gonzalez-Mantilla, A.J.; Moreno-De-Luca, A.; Ledbetter, D.H.; Martin, C.L. A cross-disorder method to identify novel candidate genes for developmental brain disorders. *JAMA psychiatry* **2016**, *73*, 275-283.
59. Hangge, P.T.; Cnota, J.F.; Woo, J.G.; Hinton, A.C.; Divanovic, A.A.; Manning, P.B.; Ittenbach, R.F.; Hinton, R.B. Microcephaly is associated with early adverse neurologic outcomes in hypoplastic left heart syndrome. *Pediatric research* **2013**, *74*, 61-67.
60. Cacheiro, P.; Haendel, M.A.; Smedley, D.; Consortium, I.M.P.; Initiative, M. New models for human disease from the International Mouse Phenotyping Consortium. *Mammalian Genome* **2019**, *30*, 143-150.
61. Barabási, A.-L.; Gulbahce, N.; Loscalzo, J. Network medicine: a network-based approach to human disease. *Nature reviews genetics* **2011**, *12*, 56-68.
62. Jones, H.N.; Olbrych, S.K.; Smith, K.L.; Cnota, J.F.; Habli, M.; Ramos-Gonzales, O.; Owens, K.J.; Hinton, A.C.; Polzin, W.J.; Muglia, L.J. Hypoplastic left heart syndrome is associated with structural and vascular placental abnormalities and leptin dysregulation. *Placenta* **2015**, *36*, 1078-1086.
63. Weinberg, A.G.; Bolande, R.P. The Liver in Congenital Heart Disease: Effects of Infantile Coarctation of the Aorta and the Hypoplastic Left Heart Syndrome in Infancy. *American Journal of Diseases of Children* **1970**, *119*, 390-394.
64. Monzen, K.; Hiroi, Y.; Kudoh, S.; Akazawa, H.; Oka, T.; Takimoto, E.; Hayashi, D.; Hosoda, T.; Kawabata, M.; Miyazono, K. Smads, TAK1, and their common target ATF-2 play a critical role in cardiomyocyte differentiation. *Journal of Cell Biology* **2001**, *153*, 687-698.
65. Yoon, W.H.; Sandoval, H.; Nagarkar-Jaiswal, S.; Jaiswal, M.; Yamamoto, S.; Haelterman, N.A.; Putluri, N.; Putluri, V.; Sreekumar, A.; Tos, T. Loss of nardilysin, a mitochondrial co-chaperone for α-ketoglutarate dehydrogenase, promotes mTORC1 activation and neurodegeneration. *Neuron* **2017**, *93*, 115-131.
66. Sadakata, T.; Furuichi, T. Identification and mRNA expression of Ogdh, QP-C, and two predicted genes in the postnatal mouse brain. *Neuroscience letters* **2006**, *405*, 217-222.
67. Hoffman, M.S. *The World almanac: and book of facts 1990*; 1989.

 doi:10.20944/preprints202202.0017.v1

68. Fruitman, D.S. Hypoplastic left heart syndrome: Prognosis and management options. *Paediatrics & child health* **2000**, *5*, 219-225.

69. Sharma, A.; Garcia Jr, G.; Wang, Y.; Plummer, J.T.; Morizono, K.; Arumugaswami, V.; Svendsen, C.N. Human iPSC-derived cardiomyocytes are susceptible to SARS-CoV-2 infection. *Cell Reports Medicine* **2020**, *1*, 100052.

70. Perez-Bermejo, J.A.; Kang, S.; Rockwood, S.J.; Simoneau, C.R.; Joy, D.A.; Silva, A.C.; Ramadoss, G.N.; Flanigan, W.R.; Fozouni, P.; Li, H. SARS-CoV-2 infection of human iPSC–derived cardiac cells reflects cytopathic features in hearts of patients with COVID-19. *Science translational medicine* **2021**, *13*.