

RESEARCH ARTICLE

Open Access

# Genome-wide expression patterns in physiological cardiac hypertrophy

Ignat Drozdov<sup>1,2</sup>, Sophia Tsoka<sup>2</sup>, Christos A Ouzounis<sup>2,3</sup>, Ajay M Shah<sup>1\*</sup>

## Abstract

**Background:** Physiological left ventricular hypertrophy (LVH) involves complex cardiac remodeling that occurs as an adaptive response to chronic exercise. A stark clinical contrast exists between physiological LVH and pathological cardiac remodeling in response to diseases such as hypertension, but little is known about the precise molecular mechanisms driving physiological adaptation.

**Results:** In this study, the first large-scale analysis of publicly available genome-wide expression data of several *in vivo* murine models of physiological LVH was carried out using network analysis. On evaluating 3 million gene co-expression patterns across 141 relevant microarray experiments, it was found that physiological adaptation is an evolutionarily conserved processes involving preservation of the function of cytochrome c oxidase, induction of autophagy compatible with cell survival, and coordinated regulation of angiogenesis.

**Conclusion:** This analysis not only identifies known biological pathways involved in physiological LVH, but also offers novel insights into the molecular basis of this phenotype by identifying key networks of co-expressed genes, as well as their topological and functional properties, using relevant high-quality microarray experiments and network inference.

## Background

Physiological left ventricular hypertrophy (LVH) is a complex cardiac adaptive response to chronic exercise [1], sometimes referred to as the “athletic heart” [2]. It is characterized by an increase in left ventricular (LV) mass, wall thickness and chamber size, underpinned by profound molecular and biochemical changes, that allows the heart to efficiently provide an increased cardiac output during periods of exercise [1]. The physiological LVH state can typically be maintained for months or years without significant compromise of cardiac function. In contrast, pathological LVH occurring in response to chronic cardiac overload, imposed by diseases such as hypertension, is characterized by a progression to contractile dysfunction and heart failure and an increased long-term mortality [3]. Other differences between physiological and pathological LVH include the occurrence of significant fibrosis and capillary rarefaction in the

latter condition. Due to the stark clinical contrast between physiological and pathological LV remodeling, it is of importance to delineate the precise molecular mechanisms that drive these divergent responses to stress.

Some progress has been made in elucidating mechanisms of physiological hypertrophy through a number of genomic analyses and several reports implicate activation of the phosphoinositide-3-kinase (PI3K)/Akt pathway as an important component [1]. More recent studies offer the possibility to examine gene expression patterns in this phenotype more consistently and broadly [4,5]. However, restrictions still exist, primarily due to an innate heterogeneity of signaling cascades and limitations of conventional statistical methods to address higher order relationships between genes. Visualization and analysis of biological data as networks is a powerful explorative alternative with the capacity to accurately assess complex relationships and eliminate noise inherent to microarray experiments [6]. Although such methods have already been successful in defining miRNA signature in obesity and diabetes [7], discovering novel cancer-associated genes [8], and predicting

\* Correspondence: [ajay.shah@kcl.ac.uk](mailto:ajay.shah@kcl.ac.uk)

<sup>1</sup>King's College London (KCL) BHF Centre of Research Excellence - Cardiovascular Division - School of Medicine - James Black Centre - 125 Coldharbour Lane, London SE5 9NU - UK

Full list of author information is available at the end of the article

the involvement of genes in core biological processes [9], their use in cardiovascular biology has been limited [10].

Recent availability of comprehensive mouse cardiac hypertrophy microarray datasets, deposited in resources such as ArrayExpress [11] and Gene Expression Omnibus [12], makes it possible to investigate global molecular mechanisms of this phenotype. The inference of gene relevance networks by co-expression analysis is based on the hypothesis that genes encoding proteins participating in the same pathway or biological process may often be co-regulated under a large number of experimental conditions [13]. An important advantage of network analysis algorithms is their ability to exploit local structure between biologically related nodes, thus eliminating most of the inherent noise [6]. Additionally, confidence in network inference through co-expression analysis may be increased by an integrative approach that utilizes multiple datasets across a variety of experimental conditions and microarray platforms [14].

In this study, a computational approach has been undertaken that identifies key expression patterns of physiological LVH using integrative analysis of 3 million gene co-expressions across 141 relevant microarray conditions. We included transcriptome data from studies in mouse models of physiological LVH induced by swimming exercise, cardiac-specific activation of Akt, and cardiac-specific activation of PI3K. This is the first study in cardiac hypertrophy at this scale and it may provide a basis for further understanding of both physiological and pathological LVH phenotypes.

## Results

### Generation of Microarray co-expression Networks

Gene expression profiles in heart tissue were investigated under normal conditions, during physiological (exercise) stress, and in two gene-modified models of physiological LVH involving cardiac activation of the PI3K/Akt pathway. To estimate the specificity of the hypertrophic gene signature, an additional dataset monitoring gene expression in healthy mouse organs was also used.

Four mouse microarray datasets totaling 141 arrays were obtained from ArrayExpress for further analysis (Table 1). The Akt dataset was generated using a tetracycline-regulated transgenic system with the capacity to conditionally switch a constitutively active form of the Akt1 protein kinase on or off in the adult heart [15]. This dataset consisted of normal heart tissue ( $n = 4$ ), short-term (2 weeks), activated Akt1 ( $n = 4$ ), and switched-off Akt1 (2 days following a 2 week activation,  $n = 4$ ). The PI3K dataset consisted of wild type hearts ( $n = 3$ ) and hearts with expression of dominant-negative PI3K ( $n = 3$ ) or constitutively active PI3K ( $n = 3$ ) [16]. The Swimming dataset, containing 30 arrays, monitored expression in mouse hearts under normal conditions, swimming (short- and long-term), and swimming followed by 1 week of rest [17]. Finally, the Normal dataset ( $n = 90$  arrays) monitored transcript expression in healthy mouse tissues including bladder, bone, spleen, stomach, and the heart [18].

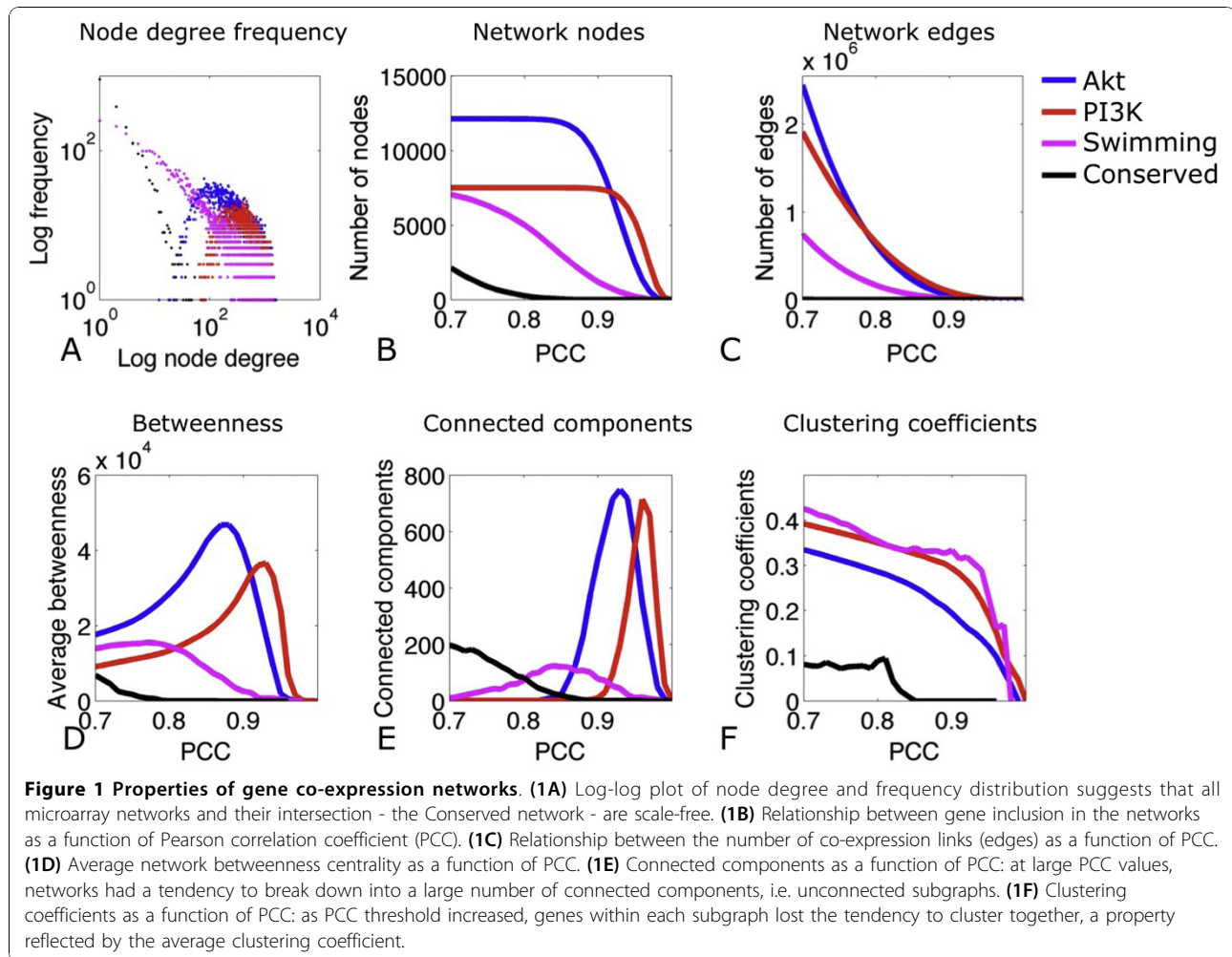
After pre-processing (see **Methods**), pair wise gene expression similarities were measured using the Pearson Correlation Coefficient (PCC). Co-expression networks were “undirected” and, at  $PCC \geq 0.70$ , obeyed a power law, suggesting a scale-free architecture dominated by a number of highly connected “hub genes” (Figure 1A). The PCC threshold was set to 0.70 on the basis of the following evidence: (i) gene correlation profiles with PCC over 0.60 were demonstrated to be more biologically relevant [19] and (ii) similar studies of human gene co-expression landscape [20] have employed comparable threshold criteria. Additionally, below this cut-off all networks were excessively large (average node degree  $>500$ ), suggesting a presence of false-positive edges. However, a more stringent PCC threshold was avoided, as further filtering has been implemented by selecting gene pairs that were correlated across all three datasets. Finally, the “data driven cut-off” approach (as performed by [19]) was not deemed appropriate as it is intended primarily for comparison of multiple networks derived from differential phenotypes.

At  $PCC \geq 0.70$  it was noted that an increase of this cut-off value removed weakly connected links from all

**Table 1 Summary of microarray datasets included in the analysis**

Dataset Name	Experimental Condition	ArrayExpress Accession	Platform	No. of Probes	Normalized No. of Probes
Akt	Wild Type, Short-term Akt1 induction ( $n = 12$ )	E-GEOD-3383 [15]	Mouse430A_2	~34,000	12127
PI3K	Wild Type, caPI3K, dnPI3K ( $n = 9$ )	E-GEOD-558 [16]	MG-U74A	~12,000	7511
Swimming	Wild Type, Short-term exercise, long-term exercise ( $n = 30$ )	E-GEOD-77 [17]	MG-U74A	~12,000	7511
Normal	Healthy tissue ( $n = 90$ )	E-GEOD-97	MG-U74A	~12,000	7511

caPI3K = constitutively active form of PI3K, dn = dominant-negative form of PI3K.



networks while maintaining a constant number of genes (Figure 1B, C). Average betweenness centrality (see **Methods**) of all networks increased with PCC values, suggesting that applying a threshold to each network removed peripheral nodes and edges, leaving critical hubs intact (Figure 1D). Additionally, increasing the PCC threshold resulted in fragmentation of networks into a large number of structured subgraphs, reflected in the number of connected components and clustering

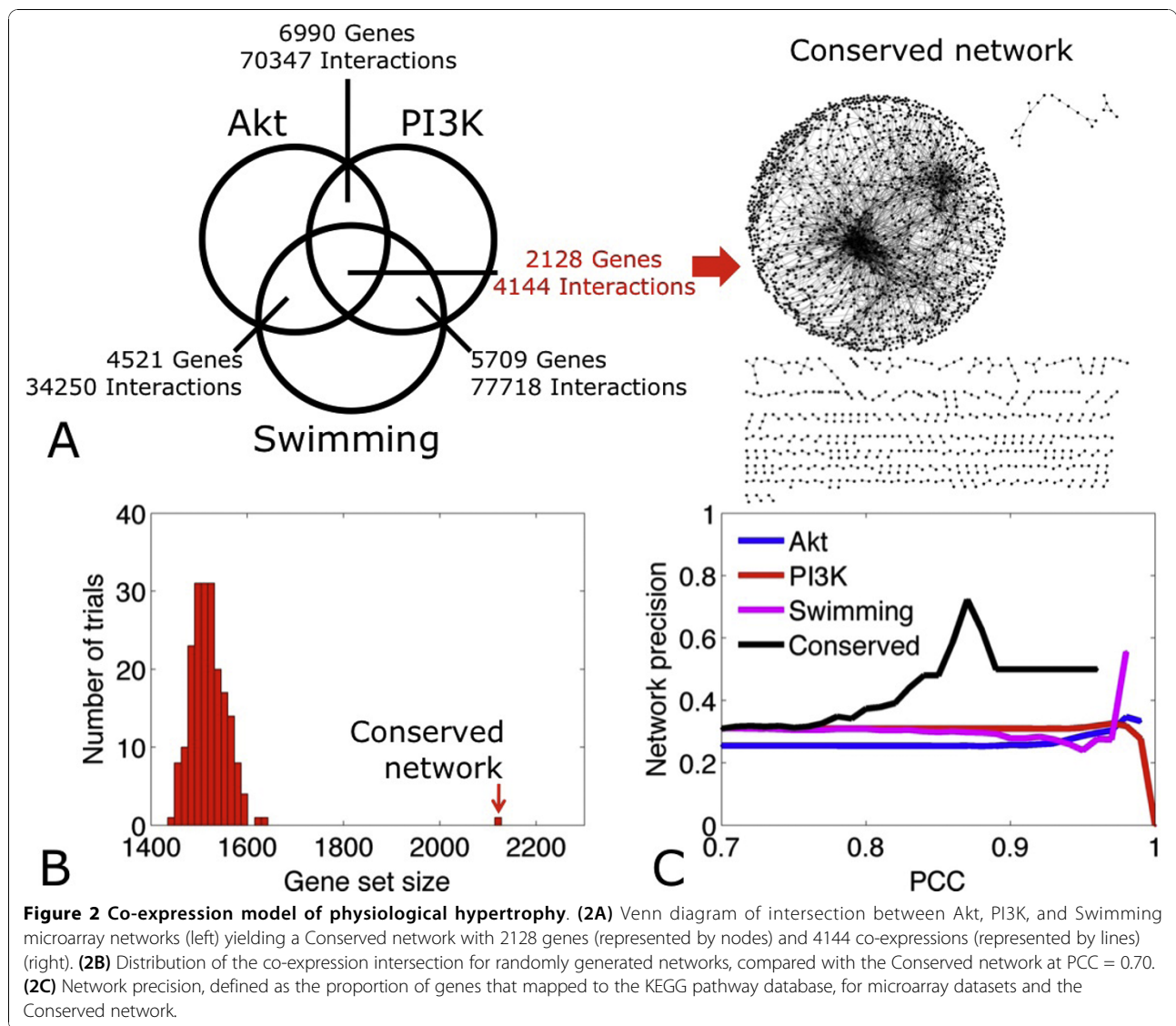
coefficients (Figure 1E, F). Overall, networks derived from hypertrophic tissues were highly structured (Table 2), characterized by nodes with multiple connections (average degree range: 211.3 to 508.6), small network diameters (range: 3.0 to 5.9) and relatively high clustering coefficients (range: 0.35 to 0.40).

**Table 2 Network Statistics at PCC = 0.70 for mouse microarray networks and the Conserved network**

Dataset	Number of Genes	Number of Interactions	Avg. Degree	Network Diameter	Clustering Coefficient
Akt	12127	2446804	403.5	3.3	0.35
PI3K	7511	1910082	508.6	3.0	0.39
Swimming	7047	744666	211.3	5.9	0.40
Conserved	2128	4144	3.9	11.8	0.10
Normal	3983	91544	45.9	4.58	0.53

**Co-expression model of Physiological cardiac hypertrophy**

Due to the large number of genes and co-expression links observed in this analysis, some observations could be due to experimental artifacts and thus of questionable biological relevance. The recurrence of a co-expression link in all three microarray datasets was considered to increase the reliability of the inference. At PCC = 0.70, the Akt and PI3K networks shared 6990 genes and 70347 interactions, the PI3K and Swimming networks shared 5709 genes and 77718 interactions, and the Akt and Swimming networks shared 4521 genes and 34250 interactions. There were 2128 genes and 4144 interactions common to all three networks, which formed a consensus 'Conserved' gene co-expression network (Figure 2A, Additional files 1 and 2).



Similarly to the Akt, PI3K, and Swimming networks, the Conserved network was scale-free.

To evaluate the statistical significance of the Conserved network, three randomized networks were generated. Randomization was performed by shuffling edges of the Akt, PI3K, and Swimming networks 4× (number of edges) times, while preserving the node degrees of the original networks [21]. This procedure was repeated 200 times (see **Methods**). The simulation showed that on average, the three random networks shared 1519 co-expressed genes (standard deviation = 35) and that at most their intersection contained 1641 genes (Figure 2B). These results indicated that identification of 2128 genes in the Conserved network is statistically significant (z-score = 17.1).

Phenotype specificity of the Conserved network was estimated by comparing it to gene co-expressions

inferred from the Normal mouse transcriptome (3983 genes, 91544 interactions;  $PCC \geq 0.70$ ) [18]. It was hypothesized that the extent of conserved nodes and edges between two networks may correspond to molecular mechanisms shared by the LVH phenotype and cells under basal conditions. Interestingly, it was determined that the Conserved and Normal networks shared only 88 genes and 57 co-expressions, confirming that the Conserved network may reflect LVH-specific cardiac response.

To gauge the extent of validated molecular pathways in all co-expression networks, all genes were mapped to the KEGG pathway database [22]. Genes with annotations in KEGG pathways were considered to be true positives and network precision (specificity) was estimated as the proportion of true positive genes to the overall number of genes in a network (true positive

cases plus presumed false positives). At PCC = 0.70, network precision for the Akt, PI3K, Swimming, and Conserved networks approached 31%. Interestingly, it was noted that while increasing PCC threshold had no apparent effect on specificity of individual microarray networks, specificity of the Conserved network increased up to two-fold with PCC values (Figure 2C), suggesting that gene pairs with high PCCs in the Conserved network are likely to be well-annotated molecular entities.

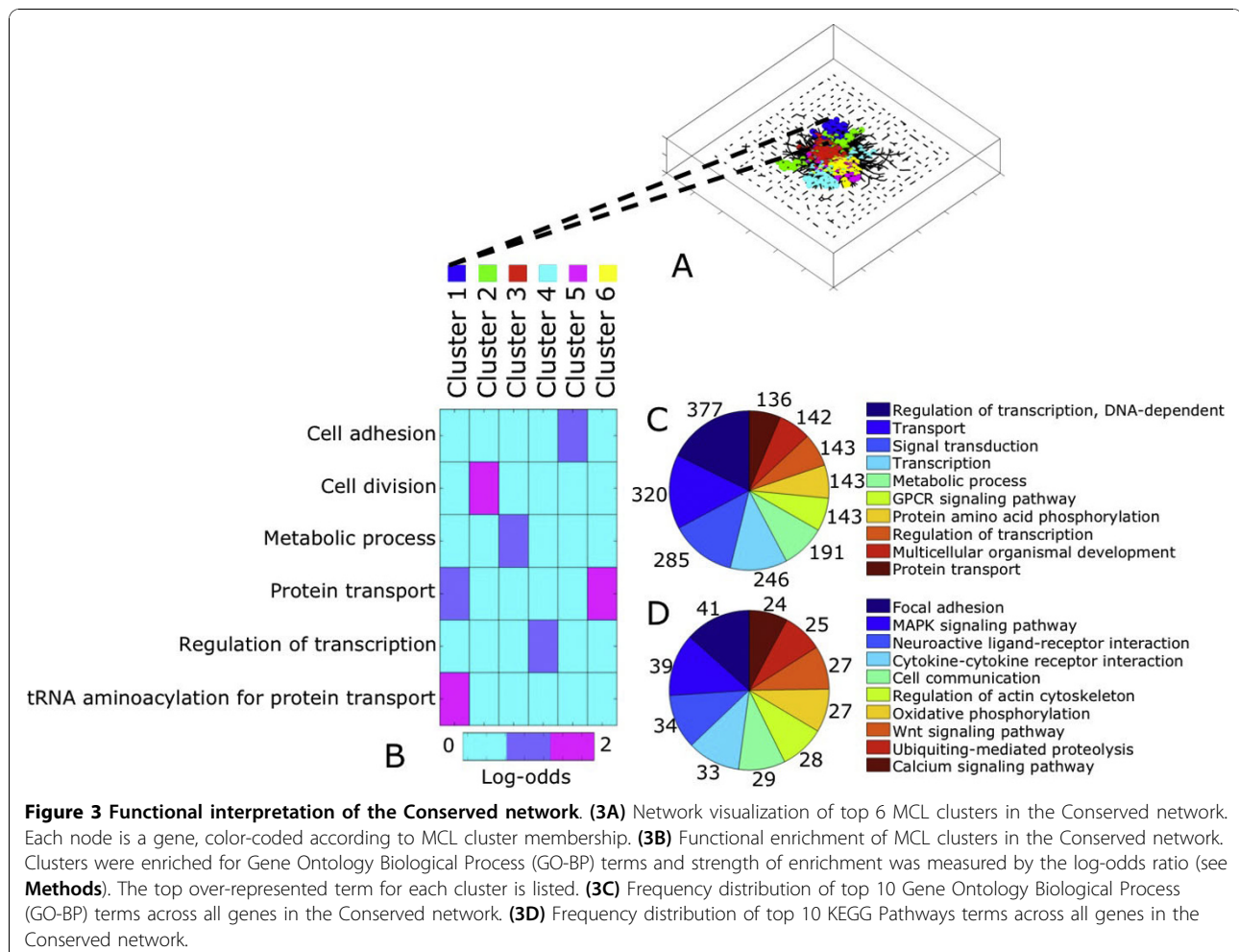
All genes in the Conserved network were processed further by MCL clustering (inflation parameter = 1.3) (see **Methods**). There were 302 clusters, of which six contained >40 genes. The largest cluster consisted of 245 genes (Figure 3A). Enrichment of each MCL cluster for GO Biological Process (BP) terms (see **Methods**) identified processes such as 'tRNA aminoacylation for protein transport' (Cluster 1), 'Cell division' (Cluster 2), and 'Protein transport' (Cluster 6) (Figure 3B). At the gene level, the Conserved network was representative of GO-BP terms such as 'Regulation of transcription' ( $n = 377$  genes), 'Transport' ( $n = 320$  genes), and 'Signal

transduction' ( $n = 285$  genes), as well as KEGG pathways such as 'Focal adhesion' ( $n = 41$  genes), 'MAPK signaling pathway' ( $n = 39$  genes), and 'Neuroactive ligand-receptor interaction' ( $n = 24$  genes) (Figure 3C, D).

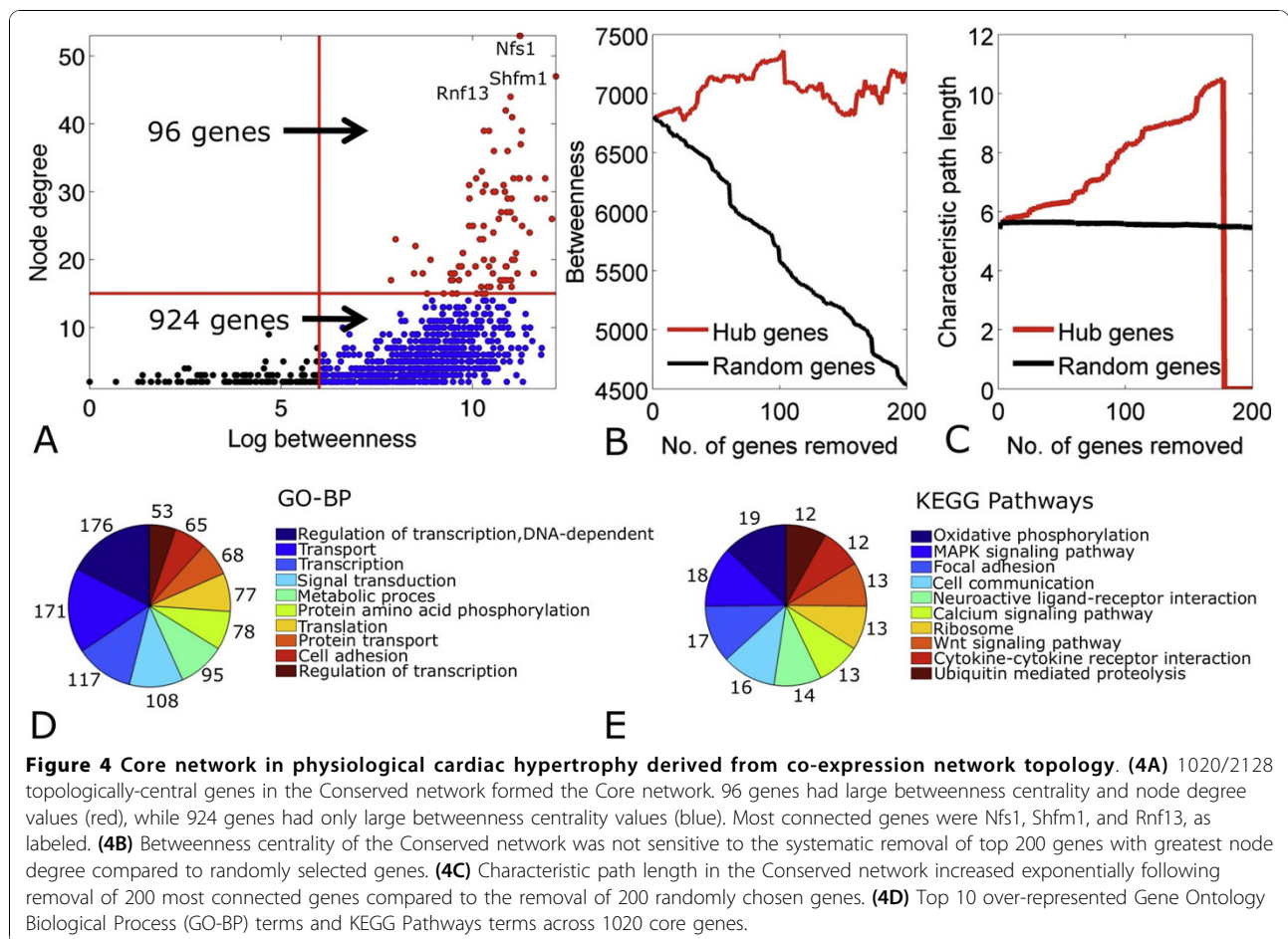
The generation of a Conserved network for physiological cardiac hypertrophy consisting of 2128 genes (Additional file 1) and 4144 interactions (Additional file 2), based on a series of relevant microarray experiments and computational processing of gene expression similarities, is thus a first step towards the discovery of the molecular underpinnings of this phenotype, its basic components and their structural and functional features.

### Identification of Critical Hubs in the Conserved co-expression Network

The topology of the Conserved network was explored further to identify hub genes. Betweenness centrality (relative importance of a gene within a graph) and node degree (number of co-expression links for a gene) were measured for 2128 genes (Figure 4A). There were 1020 genes (48% of 2128) with high betweenness centrality



**Figure 3 Functional interpretation of the Conserved network.** (3A) Network visualization of top 6 MCL clusters in the Conserved network. Each node is a gene, color-coded according to MCL cluster membership. (3B) Functional enrichment of MCL clusters in the Conserved network. Clusters were enriched for Gene Ontology Biological Process (GO-BP) terms and strength of enrichment was measured by the log-odds ratio (see **Methods**). The top over-represented term for each cluster is listed. (3C) Frequency distribution of top 10 Gene Ontology Biological Process (GO-BP) terms across all genes in the Conserved network. (3D) Frequency distribution of top 10 KEGG Pathways terms across all genes in the Conserved network.



(log betweenness centrality > 6), connected by 3047 interactions (74% of 4144). These 1020 genes formed the “core” of the Conserved network (defined as the Core network) mainly because changes in their expression and/or structure are likely to alter behavior and topology of the overall network. Remarkably, 96 out of 1020 genes had both high betweenness centrality and node degrees (log betweenness centrality > 6 and node degree > 15). These genes tended to localize at the center of the network, while the other 924 genes aligned along the periphery (Figure 4B). The three genes with the greatest values for both topological parameters were Nfs1 (nitrogen fixation 1 homolog), Shfm1 (deleted in split hand/split foot protein 1), and Rnf13 (ring finger protein 13). It is interesting to note that Nfs1 is an aminotransferase [23] with a cysteine desulfurase function [24] implicated in Freidrich’s ataxia [25], a complex disease often associated with a hypertrophic cardiomyopathy phenotype [26]. Furthermore, Shfm1 is the gene most likely associated with Split hand/split foot malformation (SHFM) in region 7q21.3-q22.1 [27], a disease exhibiting congenital heart defect phenotypes [28]. Finally, Rnf13 is a transmembrane RING-type E3 ubiquitin ligase highly

expressed in pancreatic ductal adenocarcinoma [29], but also expressed in chicken embryo brain and heart [30]. It follows that most of the other 96 genes uncovered by using the above mentioned topological parameters might also be implicated in expression patterns with a phenotype associated with heart tissue.

To test the hypothesis that hub genes may be crucial to the overall structure of the discovered network, the 200 most connected genes were systematically removed from the network. To assess network integrity, average betweenness and characteristic path length (see **Methods**) were measured. Betweenness did not change drastically following systematic removal of the top connected nodes compared to random node removal (Figure 4C). However, systematic removal of hubs increased characteristic path length to a threshold beyond which it rapidly collapsed due to splintering of the core network into small subnetworks. Characteristic path length was unaffected by removal of random genes and the network remained intact (Figure 4D).

It was then of interest to identify biological processes represented by the core network. Of 1020 core genes, 176 participated in ‘DNA-dependent regulation of

transcription', 171 in 'Transport', and 117 in 'Transcription'. Additionally, the 1020 genes were mapped to KEGG pathways such as 'Oxidative phosphorylation' ( $n = 19$  genes), 'MAPK signaling pathway' ( $n = 18$  genes), and 'Focal adhesion' ( $n = 17$  genes) (Figure 4E). Evidently, not all genes can be associated with GO or KEGG classes.

The topology of the core network was further interrogated by MCL clustering (inflation parameter = 1.3). MCL partitioned the core network into 48 clusters. The largest cluster contained 252 genes. Overall, there were 7 clusters with 20 or more genes, representing 70% of the core network (Figure 5A). Because each cluster may contain genes involved in a common molecular pathway, over-represented KEGG pathways for each cluster were identified using the log-odds ratio (see **Methods**). Only the largest three clusters could be detected as enriched by KEGG pathways, due to low counts. For example, Clusters 1 was mostly representative of 'Apoptosis' and 'Valine, leucine, and isoleucine degradation' ( $n = 8$  genes and  $n = 4$  genes respectively) while cluster 2 was representative of 'Proteasome' ( $n = 6$  genes). (Figure 5B).

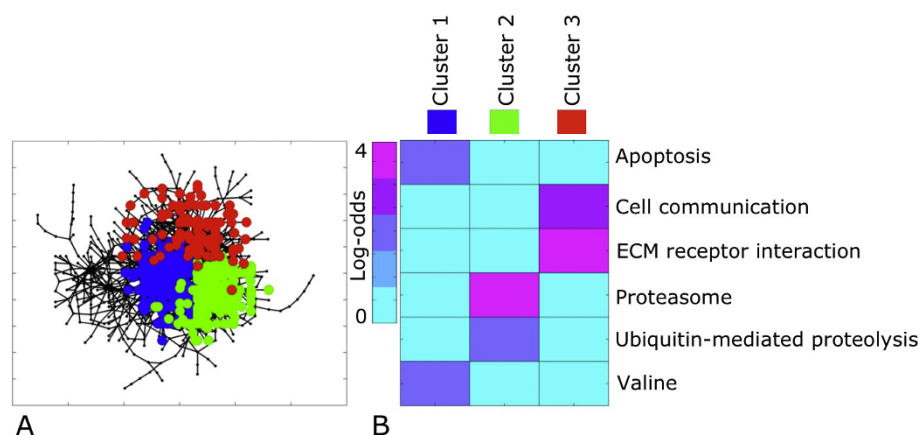
The question arises whether the 1020 genes of the Core network are also evolutionarily conserved. These genes were compared against the complete genomes of 287 species stored in the COGENT database (see **Methods**), resulting in a network of 100532 pairwise sequence similarities covering 64550 unique homologues (Figure 6A, Additional file 3). There are 271 genes that match 200 species or more, while the frequency distribution of core gene homologues is a typical distribution for sequence similarities (Figure 6B) [31]. Only 7 genes do not have a homologue apart from human or mouse (Ensembl Mouse Ids: 00000027596, 00000031329, 00000039153, 00000045107, 00000061555, 00000078135,

00000079523), most of them encoding proteins of unknown function, except 00000078135 which encodes the 'EP300 interacting inhibitor of differentiation 1' gene. The following numbers of core genes have detected a number of unique homologues, respectively as follows: 993 detect 8928 in human, 999 detect 6235 in mouse, 794 detect 3697 in *Drosophila melanogaster*, 728 detect 2424 in *Caenorhabditis elegans*, 413 detect 697 in *Saccharomyces cerevisiae* S28, just 72 detect 79 in *Escherichia coli* and 29 detect 26 in *Helicobacter pylori* J99 strain (Figure 6C), strongly indicating that the majority of the detected genes are confined to the mammalian taxonomic range. The high numbers for the animal model species as well as mouse and human are derived from extensive paralogous families within this set. Further investigation is necessary to understand the evolutionary history of the detected core network.

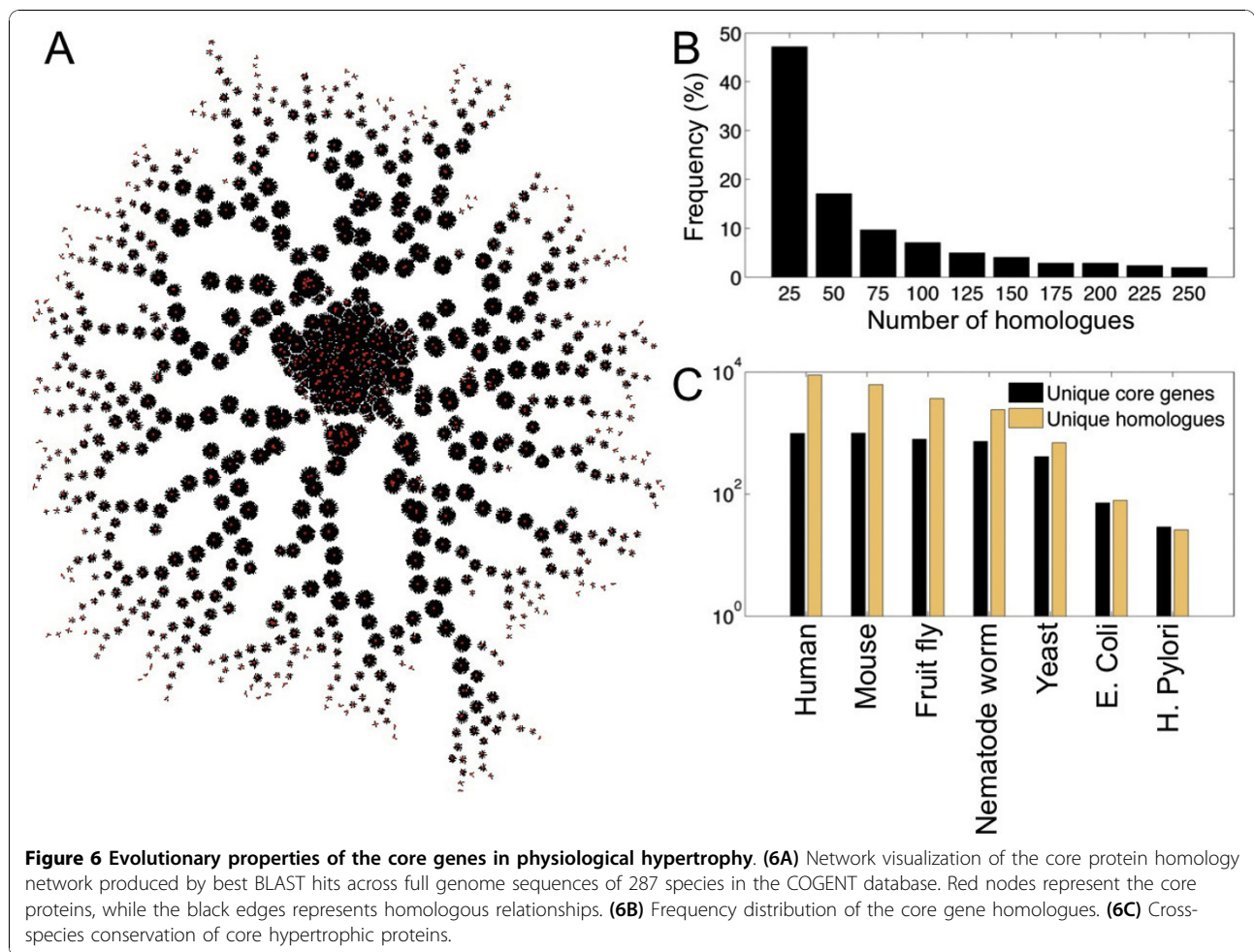
Thus, the generation of a core network for physiological cardiac hypertrophy reduces the initial number of genes to just over a thousand and consequently allows the further study of a more compact dataset, based on topological feature detection. The discovery of both known and newly detected cases in terms of genes and gene sets, along with their functional and evolutionary properties represents a consolidation of information that can be obtained from multiple microarray experiments for this key phenotype.

## Discussion

Physiological stimuli such as chronic exercise lead to compensatory growth and remodeling of the heart associated with preserved or improved cardiac function. Recently, class IA phosphoinositide 3-kinase (PI3K) and Akt1 have emerged as important regulators of physiological adaptation [32,33] but the broader signaling



**Figure 5 Clustering and KEGG pathways enrichment of the "core" conserved network. (5A)** MCL clustering of the core gene network identified 3 major clusters, each represented by a unique color **(5B)** Heatmap of gene counts in the core network clusters 1-3, enriched for KEGG Pathway terms (log-odds ratio enrichment, see **Methods**). The top over-represented term for each cluster is listed.



cascades associated with physiological LVH remain poorly understood. In this study we show that network analysis has the potential to infer genome-wide biological mechanisms related to physiological LVH phenotype. Importantly, we report on the network topology and functional properties of the physiological LVH networks, the first such analysis in a mammalian cardiovascular system.

Gene expression profiles were used to identify conserved gene co-expression patterns in PI3K, Akt1, and Swimming models of physiological LVH and to obtain a global overview of biological functions involved in physiological cardiac remodeling. Previous reports have explored gene co-expression networks derived from heterogeneous microarray platforms [14,34] and confirm that observing a conserved gene co-expression suggests a biological relevance [9,35]. The consensus gene co-expression model, referred to as the Conserved network, consisted of 2128 genes and 4144 links (Additional files 1 and 2). It was confirmed to be scale-free, highly structured, and non-random, suggesting the presence of a

small number of critical hub genes that may be biologically relevant. Additionally, the Conserved network had only a trivial intersection with the Normal interactome (88 genes, 57 links), suggesting that our consensus model may present a reliable physiological LVH signature. Topological features were consistent with the general behavior of biological networks [36] and topologies detected in protein-protein interaction collections such as STRING [37]. At  $PCC \geq 0.70$ , 31% of all genes in the Conserved network were identified in the KEGG pathways database. This coverage increased exponentially with PCC threshold, approaching 80% at  $PCC = 0.88$  (Figure 2C). These results are comparable to previous studies of co-expression networks [20] and suggest that an increase in PCC stringency produces a marked positive effect on network precision.

Due to a large number of co-expression links (3 million), it is possible that some of these links are artifacts or byproducts of systematic error. Thus, evaluation of conserved co-expression links across three physiological LVH networks has a number of strengths compared to



conventional statistical approaches. First, reproducible co-expressions are less likely to be false-positives and may reflect biologically relevant links, thus presenting a reliable interactome for further experimental validation [9,38]. For example, in a recent meta-analysis of >300 tissue samples of gastric cancer, this hypothesis helped to identify a functional link between prognostic marker PLA2G2A and the EphB2 receptor [34]. Second, network intersections account for putative platform- or experiment-dependent variability (e.g. number of transcripts) between multiple microarray datasets [14]. Third, due to the heterogeneous (molecular and physiological) nature of physiological LVH models, conserved co-expressions provide an overview of common regulatory mechanisms.

These assumptions were confirmed using automated PubMed queries, whereby each gene in the Conserved network was searched in the context of 'hypertrophy', 'heart', or 'heart failure'. Indeed, 933 out of 2128 (44%) genes in the Conserved network had at least one abstract per search term while 50 of those have at least one hundred abstracts for all terms, suggesting that the Conserved network provides an acceptable coverage of current molecular knowledge of cardiac biology (Additional file 4).

The Conserved network may be used to describe the regulatory mechanisms underpinning the cardiac remodeling response to physiological stress. 'Oxidative phosphorylation' was noted as one of the most abundant KEGG pathways ( $n = 27$  genes). The most over-represented members of this pathway were genes encoding subunits of mitochondrial cytochrome c oxidase (COX) ( $n = 6$  genes). COX is localized to the inner membrane of mitochondria and is the last component of respiratory chain. To sustain respiration, this enzyme catalyzes the transfer of electrons from cytochrome c to molecular oxygen and facilitates the aerobic production of ATP by ATPsynthase ( $n = 2$  genes in the Conserved network) [39]. To maintain efficient cardiac contractility under increased energetic demand, the regulation of COX function must be preserved. In post-myocardial infarction this mechanism is disrupted by the generation of reactive oxygen species (ROS) such as superoxide, leading to a marked loss of COX activity [40]. These results are consistent with the well-established concept that suppression of mitochondrial energy metabolism can lead to depression of cardiac contractile function [41].

The Conserved network was useful in the delineation of the cardiac response to increased protein synthesis and energy deprivation through activation of autophagy. This is a highly conserved cellular pro-survival mechanism for bulk lysosomal degradation of cytoplasmic components that mobilizes energy resources in response to starvation or hypoxia [42]. Autophagy also has a protein quality-control housekeeping function. The Conserved network

identified two key genes related to autophagic processes, Atg5 (Autophagy-related protein 5) and Becn1 (Beclin-1). Both of these genes were topologically central to the Conserved network (betweenness centrality of 53356.0 and 12262.3 respectively), implicating them in critical mediation of network information flow. Recent studies in mice with temporally controlled cardiac-specific deficiency of Atg5 demonstrated that Atg5 was essential for normal physiological growth and function of the heart. However, Atg5-deficient animals developed contractile dysfunction and heart failure accompanied by increased levels of ubiquitinated proteins. Furthermore, Atg5-deficient hearts showed disorganized sarcomere structure and mitochondrial misalignment and aggregation [43]. These abnormalities were suggested, at least in part, to be due to loss of the protein quality control function of autophagy. Becn1 is part of a PI3K complex that plays an important role during the initiation of autophagosome formation [42,44]. Interestingly, mice with heterozygous disruption of Becn1 (Becn1<sup>+/-</sup>) exhibited reduced levels of autophagy during reperfusion but had decreased apoptosis and reduced infarct size compared to wild type mice [45], suggesting that in this case autophagy was detrimental. However, Becn1 is an important point of crosstalk with apoptotic pathways through its interaction with anti-apoptotic proteins such as Bcl-2 [46]. Disruption of Becn1 could therefore have pro- or anti-survival effects [42]. Of note, in the Conserved network, Becn1 localized to the same MCL cluster as Bcl-2, which is known to inhibit Becn1-dependent autophagy [46]. Thus, in physiological LVH, autophagy compatible with cell survival, rather than cell death, may be regulated by coordinated changes in Atg5, Becn1 and Bcl-2. Indeed, autophagy- and proteolysis-related genes localized to the same cluster as genes involved in cell cycle regulation, providing further support for this hypothesis.

To explore if key regulatory mechanisms may be encoded by topologically significant nodes, the Conserved network was studied using concepts of betweenness centrality and node degree. These approaches are known to detect essential hubs in interaction networks [47] and previous studies have demonstrated that betweenness is a good indicator of biological essentiality [48]. Interestingly, when the top 200 hub genes were systematically removed from the Conserved network, average network betweenness remained mostly constant and high, while characteristic path length increased dramatically, to a threshold beyond which the network collapsed. This may suggest a presence of a large number of well-connected genes that preserve network information flow, possibly an indicator of maintained functional cardiac integrity during physiological remodeling. Additionally, topologically-central genes (core genes) localized to

KEGG pathways including 'Oxidative phosphorylation' ( $n = 19$  genes), 'MAPK signaling pathway' ( $n = 18$  genes), and 'Focal adhesion' ( $n = 17$  genes) (Figure 4E).

Several genes associated with the mammalian target of rapamycin (mTOR) pathway (Cab39, Hif1a, Tsc2) were also identified. The mTOR pathway controls changes in cell size following activation of the PI3K/Akt system. Akt phosphorylates the Tsc2 gene product tuberin, and thereby reduces its ability to stimulate GTP hydrolysis on the Ras-like G protein Rheb, leading to increased protein synthesis via ribosome biogenesis - a key feature of cardiac hypertrophy - and cell growth [49]. Recently, inhibition of the mTOR pathway by rapamycin was demonstrated to alleviate load-induced cardiac hypertrophy in mice, making it a potential therapeutic target [50]. Indeed, Tsc2 had a very large betweenness centrality value (174802.9, top 1%), confirming that it is one of the key constituents of the Conserved network. Core genes present in the 'MAPK signaling pathway' included Map4k3, Map3k7, Rap1a, Mapkapk2, Cacng2, and Ppm1b. Of these, Ppm1b (protein phosphatase 1B) had the greatest node degree (32) and betweenness centrality (73822.0) values, supporting its biological importance. These findings are reinforced by demonstration of direct inhibition of Map3k7 by Ppm1b [51], thus providing further evidence that Map3k7 activity is reduced in physiological hypertrophy protecting the heart from interstitial fibrosis, severe myocardial dysfunction, and apoptosis [52].

Similarly, the core Conserved network suggests that the genes involved in KEGG 'Calcium signaling pathway' may be involved in physiological LVH. There were 13 genes (e.g. Ppp3ca, Egfr, Vdac3, Slc25a4, Tnnc1) allocated to 'Calcium signaling pathway', of which Ppp3ca (calcineurin A alpha) had the largest betweenness centrality value (71043.2). Ppp3ca has been shown to be a key regulator of cardiac hypertrophy through activation of the transcription factor NFAT (nuclear factor of activated T-cells) which promotes the expression of pro-hypertrophic genes in concert with other transcription factors such as GATA4 and MEF2 [53]. It can also inhibit Map3k7 signaling [54]. The Conserved network also provides further evidence that calcineurin activity is highly regulated under physiological conditions by elucidation of the Rcn2 gene, which is known to inhibit calcineurin signaling [55].

The use of MCL in the core network (Figure 5) identified enriched clusters of genes participating in similar biological pathways. For example, cluster 1 was enriched for KEGG pathway 'Apoptosis' ( $n = 5$  genes: Birc2, Irak1, Pik3ca, Prkaca, Ppp3ca). Birc2 (baculoviral IAP repeat-containing 2, betweenness = 3316.0) encodes a protein that inhibits apoptosis by binding to tumor necrosis factor receptor-associated factors TRAF1 and TRAF2.

Although previously not reported in the mammalian heart, Birc2 was confirmed as a critical regulator of vascular integrity and endothelial cell survival in zebrafish [56]. Null mutants for Birc2 showed severe hemorrhage and vascular regression due to endothelial cell integrity defects and activation of Caspase-8-dependent apoptosis program. Coordinated regulation of angiogenesis is essential for preserved cardiac contractile function [57] and our results provide further molecular evidence for angiogenic gene programs in physiological LVH that merits further exploration.

## Conclusions

This report presents the first integrative analysis of genome-wide expression data and computational network inference in the context of physiological LVH. The identification of several mechanisms already known to be involved in physiological cardiac remodeling based on prior experimental studies provides confirmation to the validity of the approaches used in this study. In addition to supporting current molecular understanding of the cardiac physiological response to stress, this work characterizes topological and functional properties of 2128 potential molecular targets involved in the systematic regulation of physiological LVH. Additionally, we demonstrate for the first time the evolutionary complexity of the hypertrophic response. Our study suggests that evaluation of higher order relationships between genes and their neighbors, rather than mere individual over- or under-expression, may facilitate a better understanding of function in physiological and pathological phenotypes. Overall, the results offer new support for the utility of co-expression network modeling and the quality of public microarray data in the context of cardiac hypertrophy, facilitating further analysis of complex physiological and pathological phenotypes.

## Methods

### Data Preparation

Three publicly available mouse microarray datasets were included in this study, corresponding to 51 arrays. Individual mouse phenotypes under experimental conditions were reviewed carefully to ensure that each met physiological inclusion criteria (LVH with preserved or improved heart function and corresponding normal controls). Raw expression values were obtained from ArrayExpress database [58] and normalized using Robust Multi-array Average (RMA) [59]. Probesets with very low expression across experiments were removed and, in cases where multiple probesets mapped to a single gene, only those genes with the highest intensities were retained. To standardize annotation across multiple microarray platforms, Affymetrix probe identifiers were mapped to their corresponding Ensembl (August 12, 2009) gene identifiers (IDs) [60].

Pairwise similarity in gene expression vectors was expressed by the Pearson correlation coefficient (PCC). Gene pairs that correlated above a predefined PCC threshold value were represented in the form of an undirected unweighted network, where nodes (vertices) correspond to genes and links (edges) correspond to co-expression between genes. Randomized networks were generated by rewiring edges in the original networks while preserving the degrees of the respective nodes [21]. The number of rewiring steps taken for each model was  $4 \times$  (number of edges). This method ensures that topological structure of the network is retained during randomization.

### Network consensus and topological analysis

A co-expression link between two genes was considered as a 'consensus' link, if it was observed in all three datasets. Topological properties examined were node degree, network diameter, betweenness centrality, connected components, clustering coefficient, and characteristic path length [61]. Node degree is defined as the total number of edges that connect to a given node. Network diameter is defined as the average shortest path between any pair of nodes in the network. Betweenness centrality is the measure of node importance within a graph, where nodes that occur on many shortest paths between nodes have higher betweenness. Connected components are maximal connected subgraphs of an undirected graph in which any two vertices are connected to each other by edges. Clustering coefficient is the degree to which nodes tend to cluster together. Characteristic path length is the average distance between pairs of vertices.

### Cluster Analysis and Functional Enrichment

Significant clusters of genes in a co-expression network were identified using Markov Cluster Algorithm (MCL) [62]. This is an efficient, unsupervised, and accurate graph clustering approach based on simulation of stochastic flow in graphs. To ensure significance of enrichment, only resulting clusters with 10 or more genes were further retained. A distinct advantage of MCL is its ability to avoid incorrect clustering assignments in the presence of false negative edges [6,62]. This is due to the fact that MCL discovers clusters by virtue of genes sharing higher-order connectivity in their local neighborhoods and not merely pairwise linkages. Genes identified to be present in the same cluster were analyzed for over-represented (enriched) Gene Ontology Biological Process (GO-BP) terms and KEGG pathways [22] using the log-odds ratio. Higher ratio indicates a higher relative abundance of a GO-BP term or KEGG pathway in a cluster compared to the entire network. While all KEGG pathways were considered for enrichment, to avoid broad annotation terms, only GO-BP categories with fewer than 1,500 genes (mouse annotations) were retained [63].

### Evolutionary gene analysis

Evolutionary conservation was computed by comparing selected protein sequences from the core network (corresponding to 1020 genes) against the complete genomes of 287 species available in the Complete Genome Tracking (COGENT) database [64] database using BLAST [65] with default parameters. Significant hits from this run have been retained with a p-value cut-off  $< 10^{-06}$ , corresponding to 100532 pairwise similarity relationships. Homology networks were visualized using the Large Graph Layout (LGL) software [66].

### Additional material

**Additional file 1: Table of Conserved network genes in physiological cardiac hypertrophy (sorted by node degree).**

**Additional file 2: Gene edge list for the Conserved physiological cardiac hypertrophy network.**

**Additional file 3: An edge list for 100532 pairwise sequence similarities covering 64550 unique homologues.**

**Additional file 4: Specificity of the Conserved network genes to the current knowledge of cardiac biology. (A)** 933/2128 (44%) genes in the Conserved network had at least one abstract per PubMed search terms 'Hypertrophy', 'Heart', or 'Heart Failure'. Number of abstracts per gene is reflected both by a color scale blue-red and node size.

**(B)** PubMed abstract frequency histogram of 933 genes confirmed to be related to cardiac biology.

### Acknowledgements

This work was supported by the British Heart Foundation (BHF) through awards RE/08/003 and CH/99001 (AMS) and a PhD studentship for ID. Parts of this work have been supported by the Network of Excellence ENFIN (contract number LSHG-CT-2005-518254) funded by the European Commission, and the Leducq Fondation.

### Author details

<sup>1</sup>King's College London (KCL) BHF Centre of Research Excellence - Cardiovascular Division - School of Medicine - James Black Centre - 125 Coldharbour Lane, London SE5 9NU - UK. <sup>2</sup>Centre for Bioinformatics - School of Physical Sciences & Engineering - King's College London (KCL) - Strand, London WC2R 2LS - UK. <sup>3</sup>Computational Genomics Unit, Institute of Agrobiotechnology - Centre for Research & Technology Hellas - PO Box 361, GR-57001 Thessaloniki - Greece.

### Authors' contributions

ID performed data integration, statistical and network analysis and drafted the initial manuscript. ST has contributed to the analysis and the manuscript. CAO carried out computational analysis, edited the manuscript and coordinated the research. AMS coordinated the project design and edited the manuscript. All authors read and approved the manuscript.

Received: 6 June 2010 Accepted: 11 October 2010

Published: 11 October 2010

### References

1. Catalucci D, Latronico MV, Ellingsen O, Condorelli G: **Physiological myocardial hypertrophy: how and why?** *Front Biosci* 2008, **13**:312-324.
2. Raskoff WJ, Goldman S, Cohn K: **The "athletic heart". Prevalence and physiological significance of left ventricular enlargement in distance runners.** *JAMA* 1976, **236**(2):158-162.
3. Mann DL, Bristow MR: **Mechanisms and models in heart failure: the biomechanical model and beyond.** *Circulation* 2005, **111**(21):2837-2849.

4. Iemitsu M, Maeda S, Miyauchi T, Matsuda M, Tanaka H: **Gene expression profiling of exercise-induced cardiac hypertrophy in rats.** *Acta Physiol Scand* 2005, **185**(4):259-270.
5. Kong SW, Bodyak N, Yue P, Liu Z, Brown J, Izumo S, Kang PM: **Genetic expression profiles during physiological and pathological cardiac hypertrophy and heart failure in rats.** *Physiol Genomics* 2005, **21**(1):34-42.
6. Enright AJ, Van Dongen S, Ouzounis CA: **An efficient algorithm for large-scale detection of protein families.** *Nucleic Acids Res* 2002, **30**(7):1575-1584.
7. Zampetaki A, Kiechl S, Drozdzov I, Willeit P, Mayr U, Prokopi M, Mayr A, Weger S, Oberhollenzer F, Bonora E: **Plasma MicroRNA Profiling Reveals Loss of Endothelial MiR-126 and Other MicroRNAs in Type 2 Diabetes.** *Circ Res* 2010, **107**(6):810-7.
8. Pujana MA, Han JD, Starita LM, Stevens KN, Tewari M, Ahn JS, Rennert G, Moreno V, Kirchhoff T, Gold B: **Network modeling links breast cancer susceptibility and centrosome dysfunction.** *Nat Genet* 2007, **39**(11):1338-1349.
9. Stuart JM, Segal E, Koller D, Kim SK: **A gene-coexpression network for global discovery of conserved genetic modules.** *Science* 2003, **302**(5643):249-255.
10. Shreenivasiah PK, Rho SH, Kim T, Kim do H: **An overview of cardiac systems biology.** *J Mol Cell Cardiol* 2008, **44**(3):460-469.
11. McWilliam H, Valentin F, Goujon M, Li W, Narayanasamy M, Martin J, Miyar T, Lopez R: **Web services at the European Bioinformatics Institute-2009.** *Nucleic Acids Res* 2009, **37** Web Server: W6-10.
12. Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Edgar R: **NCBI GEO: mining tens of millions of expression profiles—database and tools update.** *Nucleic Acids Res* 2007, **35** Database: D760-765.
13. Oldham MC, Horvath S, Geschwind DH: **Conservation and evolution of gene coexpression networks in human and chimpanzee brains.** *Proc Natl Acad Sci USA* 2006, **103**(47):17973-17978.
14. Lee HK, Hsu AK, Sajdak J, Qin J, Pavlidis P: **Coexpression analysis of human genes across many microarray data sets.** *Genome Res* 2004, **14**(6):1085-1094.
15. Schiekofe S, Shiojima I, Sato K, Galasso G, Oshima Y, Walsh K: **Microarray analysis of Akt1 activation in transgenic mouse hearts reveals transcript expression profiles associated with compensatory hypertrophy and failure.** *Physiol Genomics* 2006, **27**(2):156-170.
16. McMullen JR, Amirahmadi F, Woodcock EA, Schinke-Braun M, Bouwman RD, Hewitt KA, Mollica JP, Zhang L, Zhang Y, Shioi T: **Protective effects of exercise and phosphoinositide 3-kinase(p110alpha) signaling in dilated and hypertrophic cardiomyopathy.** *Proc Natl Acad Sci USA* 2007, **104**(2):612-617.
17. Kaplan ML, Cheslow Y, Vikstrom K, Malhotra A, Geenen DL, Nakouzi A, Leinwand LA, Buttrick PM: **Cardiac adaptations to chronic exercise in mice.** *Am J Physiol* 1994, **267**(3 Pt 2):H1167-1173.
18. Su AI, Cooke MP, Ching KA, Hakak Y, Walker JR, Wiltshire T, Orth AP, Vega RG, Sapinoso LM, Moqrich A: **Large-scale analysis of the human and mouse transcriptomes.** *Proc Natl Acad Sci USA* 2002, **99**(7):4465-4470.
19. Elo LL, Jarvenpaa H, Oresic M, Lahesmaa R, Aittokallio T: **Systematic construction of gene coexpression networks with applications to human T helper cell differentiation process.** *Bioinformatics* 2007, **23**(16):2096-2103.
20. Prieto C, Risueno A, Fontanillo C, De las Rivas J: **Human gene coexpression landscape: confident network derived from tissue transcriptomic profiles.** *PLoS One* 2008, **3**(12):e3911.
21. Maslov S, Sneppen K: **Specificity and stability in topology of protein networks.** *Science* 2002, **296**(5569):910-913.
22. Kanehisa M, Goto S: **KEGG: kyoto encyclopedia of genes and genomes.** *Nucleic Acids Res* 2000, **28**(1):27-30.
23. Ouzounis C, Sander C: **Homology of the NifS family of proteins to a new class of pyridoxal phosphate-dependent enzymes.** *FEBS Lett* 1993, **322**(2):159-164.
24. Zheng L, White RH, Cash VL, Jack RF, Dean DR: **Cysteine desulfurase activity indicates a role for NIFS in metallocluster biosynthesis.** *Proc Natl Acad Sci USA* 1993, **90**(7):2754-2758.
25. Huang ML, Becker EM, Whitnall M, Rahmanto YS, Ponka P, Richardson DR: **Elucidation of the mechanism of mitochondrial iron loading in Friedreich's ataxia by analysis of a mouse mutant.** *Proc Natl Acad Sci USA* 2009, **106**(38):16381-16386.
26. Pandolfo M: **Friedreich ataxia.** *Arch Neurol* 2008, **65**(10):1296-1303.
27. Crackower MA, Scherer SW, Rommens JM, Hui CC, Poorkaj P, Soder S, Cobben JM, Hudgins L, Evans JP, Tsui LC: **Characterization of the split hand/split foot malformation locus SHFM1 at 7q21.3-q22.1 and analysis of a candidate gene for its expression during limb development.** *Hum Mol Genet* 1996, **5**(5):571-579.
28. Elliott AM, Evans JA: **The association of split hand foot malformation (SHFM) and congenital heart defects.** *Birth Defects Res A Clin Mol Teratol* 2008, **82**(6):425-434.
29. Zhang Q, Meng Y, Zhang L, Chen J, Zhu D: **RNF13: a novel RING-type ubiquitin ligase over-expressed in pancreatic cancer.** *Cell Res* 2009, **19**(3):348-357.
30. Tranque P, Crossin KL, Cirelli C, Edelman GM, Mauro VP: **Identification and characterization of a RING zinc finger gene (C-RZF) expressed in chicken embryo cells.** *Proc Natl Acad Sci USA* 1996, **93**(7):3105-3109.
31. Enright AJ, Kunin V, Ouzounis CA: **Protein families and TRIBES in genome sequence space.** *Nucleic Acids Res* 2003, **31**(15):4632-4638.
32. Luo J, McMullen JR, Sobkiw CL, Zhang L, Dorfman AL, Sherwood MC, Logsdon MN, Horner JW, DePinho RA, Izumo S: **Class IA phosphoinositide 3-kinase regulates heart size and physiological cardiac hypertrophy.** *Mol Cell Biol* 2005, **25**(21):9491-9502.
33. DeBosch B, Treskov I, Lupu TS, Weinheimer C, Kovacs A, Courtois M, Muslin AJ: **Akt1 is required for physiological cardiac growth.** *Circulation* 2006, **113**(17):2097-2104.
34. Aggarwal A, Guo DL, Hoshida Y, Yuen ST, Chu KM, So S, Boussiouas A, Chen X, Bowtell D, Aburatani H: **Topological and functional discovery in a gene coexpression meta-network of gastric cancer.** *Cancer Res* 2006, **66**(1):232-241.
35. Graeber TG, Eisenberg D: **Bioinformatic identification of potential autocrine signaling loops in cancers from gene expression profiles.** *Nat Genet* 2001, **29**(3):295-300.
36. Jeong H, Mason SP, Barabasi AL, Oltvai ZN: **Lethality and centrality in protein networks.** *Nature* 2001, **411**(6833):41-42.
37. von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B: **STRING: a database of predicted functional associations between proteins.** *Nucleic Acids Res* 2003, **31**(1):258-261.
38. Aid-Pavlidis T, Pavlidis P, Timmusk T: **Meta-coexpression conservation analysis of microarray data: a "subset" approach provides insight into brain-derived neurotrophic factor regulation.** *BMC Genomics* 2009, **10**: 420.
39. Proshlyakov DA, Pressler MA, DeMaso C, Leykam JF, DeWitt DL, Babcock GT: **Oxygen activation and reduction in respiration: involvement of redox-active tyrosine 244.** *Science* 2000, **290**(5496):1588-1591.
40. Paradies G, Petrosillo G, Pistolesse M, Ruggiero FM: **The effect of reactive oxygen species generated from the mitochondrial electron transport chain on the cytochrome c oxidase activity and on the cardiolipin content in bovine heart submitochondrial particles.** *FEBS Lett* 2000, **466**(2-3):323-326.
41. Ingwall JS, Weiss RG: **Is the failing heart energy starved? On using chemical energy to support cardiac function.** *Circ Res* 2004, **95**(2):135-145.
42. Levine B, Kroemer G: **Autophagy in the pathogenesis of disease.** *Cell* 2008, **132**(1):27-42.
43. Nakai A, Yamaguchi O, Takeda T, Higuchi Y, Hikoso S, Taniike M, Omiya S, Mizote I, Matsumura Y, Asahi M: **The role of autophagy in cardiomyocytes in the basal state and in response to hemodynamic stress.** *Nat Med* 2007, **13**(5):619-624.
44. Kihara A, Noda T, Ishihara N, Ohsumi Y: **Two distinct Vps34 phosphatidylinositol 3-kinase complexes function in autophagy and carboxypeptidase Y sorting in *Saccharomyces cerevisiae*.** *J Cell Biol* 2001, **152**(3):519-530.
45. Matsui Y, Takagi H, Qu X, Abdellatif M, Sakoda H, Asano T, Levine B, Sadoshima J: **Distinct roles of autophagy in the heart during ischemia and reperfusion: roles of AMP-activated protein kinase and Beclin 1 in mediating autophagy.** *Circ Res* 2007, **100**(6):914-922.
46. Pattengire S, Tassa A, Qu X, Garuti R, Liang XH, Mizushima N, Packer M, Schneider MD, Levine B: **Bcl-2 antiapoptotic proteins inhibit Beclin 1-dependent autophagy.** *Cell* 2005, **122**(6):927-939.
47. Girvan M, Newman ME: **Community structure in social and biological networks.** *Proc Natl Acad Sci USA* 2002, **99**(12):7821-7826.
48. Yu H, Kim PM, Sprecher E, Trifonov V, Gerstein M: **The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics.** *PLoS Comput Biol* 2007, **3**(4):e59.
49. Dorn GW, Force T: **Protein kinase cascades in the regulation of cardiac hypertrophy.** *J Clin Invest* 2005, **115**(3):527-537.

50. Shioi T, McMullen JR, Tamavski O, Converso K, Sherwood MC, Manning WJ, Izumo S: **Rapamycin attenuates load-induced cardiac hypertrophy in mice.** *Circulation* 2003, **107**(12):1664-1670.
51. Hanada M, Ninomiya-Tsuji J, Komaki K, Ohnishi M, Katsura K, Kanamaru R, Matsumoto K, Tamura S: **Regulation of the TAK1 signaling pathway by protein phosphatase 2C.** *J Biol Chem* 2001, **276**(8):5753-5759.
52. Zhang D, Gaussin V, Taffet GE, Belaguli NS, Yamada M, Schwartz RJ, Michael LH, Overbeek PA, Schneider MD: **TAK1 is activated in the myocardium after pressure overload and is sufficient to provoke heart failure in transgenic mice.** *Nat Med* 2000, **6**(5):556-563.
53. Heineke J, Molkentin JD: **Regulation of cardiac hypertrophy by intracellular signalling pathways.** *Nat Rev Mol Cell Biol* 2006, **7**(8):589-600.
54. Liu Q, Busby JC, Molkentin JD: **Interaction between TAK1-TAB1-TAB2 and RCAN1-calcineurin defines a signalling nodal control point.** *Nat Cell Biol* 2009, **11**(2):154-161.
55. Mehta S, Li H, Hogan PG, Cunningham KW: **Domain architecture of the regulators of calcineurin (RCANs) and identification of a divergent RCAN in yeast.** *Mol Cell Biol* 2009, **29**(10):2777-2793.
56. Santoro MM, Samuel T, Mitchell T, Reed JC, Stainier DY: **Birc2 (clap1) regulates endothelial cell integrity and blood vessel homeostasis.** *Nat Genet* 2007, **39**(11):1397-1402.
57. Shiojima I, Sato K, Izumiya Y, Schiekofe S, Ito M, Liao R, Colucci WS, Walsh K: **Disruption of coordinated cardiac hypertrophy and angiogenesis contributes to the transition to heart failure.** *J Clin Invest* 2005, **115**(8):2108-2118.
58. Parkinson H, Kapushesky M, Kolesnikov N, Rustici G, Shojatalab M, Abeygunawardena N, Berube H, Dylag M, Emam I, Farne A: **ArrayExpress update—from an archive of functional genomics experiments to the atlas of gene expression.** *Nucleic Acids Res* 2009, **37** Database: D868-872.
59. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP: **Exploration, normalization, and summaries of high density oligonucleotide array probe level data.** *Biostatistics* 2003, **4**(2):249-264.
60. Birney E, Andrews TD, Bevan P, Caccamo M, Chen Y, Clarke L, Coates G, Cuff J, Curwen V, Cutts T: **An overview of Ensembl.** *Genome Res* 2004, **14**(5):925-928.
61. Freeman TC, Goldovsky L, Brosch M, van Dongen S, Maziere P, Grocock RJ, Freilich S, Thornton J, Enright AJ: **Construction, visualisation, and clustering of transcription networks from microarray expression data.** *PLoS Comput Biol* 2007, **3**(10):2032-2042.
62. van Dongen S: **Graph clustering by flow simulation.** PhD University of Utrecht 2000.
63. Wang K, Narayanan M, Zhong H, Tompa M, Schadt EE, Zhu J: **Meta-analysis of inter-species liver co-expression networks elucidates traits associated with common human diseases.** *PLoS Comput Biol* 2009, **5**(12):e1000616.
64. Janssen P, Enright AJ, Audit B, Cases I, Goldovsky L, Harte N, Kunin V, Ouzounis CA: **CComplete GENome Tracking (COGENT): a flexible data environment for computational genomics.** *Bioinformatics* 2003, **19**(11):1451-1452.
65. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**(3):403-410.
66. Adai AT, Date SV, Wieland S, Marcotte EM: **LGL: creating a map of protein function with an algorithm for visualizing very large biological networks.** *J Mol Biol* 2004, **340**(1):179-190.

doi:10.1186/1471-2164-11-557

Cite this article as: Drozdov *et al.*: Genome-wide expression patterns in physiological cardiac hypertrophy. *BMC Genomics* 2010 **11**:557.

Submit your next manuscript to BioMed Central  
and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

