Miriam Angeloni, Ingo Thievessen, Felix B. Engel, Paolo Magni and Fulvia Ferrazzi*

Functional genomics meta-analysis to identify gene set enrichment networks in cardiac hypertrophy

https://doi.org/10.1515/hsz-2020-0378
Received November 30, 2020; accepted April 19, 2021;
published online May 5, 2021

Abstract: In order to take advantage of the continuously increasing number of transcriptome studies, it is important to develop strategies that integrate multiple expression datasets addressing the same biological question to allow a robust analysis. Here, we propose a meta-analysis framework that integrates enriched pathways identified through the Gene Set Enrichment Analysis (GSEA) approach and calculates for each meta-pathway an empirical *p*-value. Validation of our approach on benchmark datasets showed comparable or even better performance than existing methods and an increase in robustness with increasing number of integrated datasets. We then applied the meta-analysis framework to 15 functional genomics datasets of physiological and pathological cardiac hypertrophy. Within these datasets we grouped expression sets measured at time points that represent the same hallmarks of heart tissue remodeling ('aggregated time points') and performed meta-analysis on the expression sets assigned to each aggregated time point. To facilitate biological interpretation, results were visualized as gene set enrichment networks. Here, our meta-analysis framework identified well-known biological mechanisms associated with pathological cardiac hypertrophy (e.g., cardiomyocyte apoptosis, cardiac contractile dysfunction, and alteration in energy metabolism). In addition, results highlighted novel, potentially cardioprotective mechanisms in physiological cardiac hypertrophy involving the down-regulation of immune cell response, which are worth further investigation.

Keywords: cardiac hypertrophy; pathological hypertrophy; pathway network; physiological hypertrophy; transcriptome meta-analysis.

***Corresponding author: Fulvia Ferrazzi**, Department of Nephropathology, Institute of Pathology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Krankenhausstr. 8-10, D-91054 Erlangen, Germany; Institute of Pathology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Krankenhausstr. 8-10, D-91054 Erlangen, Germany; and Muscle Research Center Erlangen (MURCE), D-91052 Erlangen, Germany, E-mail: fulvia.ferrazzi@uk-erlangen.de. https://orcid.org/0000-0003-4011-4638
Miriam Angeloni, Department of Nephropathology, Institute of Pathology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Krankenhausstr. 8-10, D-91054 Erlangen, Germany; and Institute of Pathology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Krankenhausstr. 8-10, D-91054 Erlangen, Germany
Ingo Thievessen, Biophysics Group, Department of Physics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Henkestraße 91, D-91052 Erlangen, Germany; and Muscle Research Center Erlangen (MURCE), D-91052 Erlangen, Germany
Felix B. Engel, Experimental Renal and Cardiovascular Research, Department of Nephropathology, Institute of Pathology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Schwabachanlage 12, D-91054 Erlangen, Germany; and Muscle Research Center Erlangen (MURCE), D-91052 Erlangen, Germany
Paolo Magni, Department of Electrical, Computer and Biomedical Engineering, University of Pavia, via Ferrata 5, I-27100 Pavia, Italy

Introduction

In functional genomics, expression microarrays and RNA sequencing (RNA-Seq) allow the measurement and comparison of genome-wide expression data for different biological conditions (Ferrazzi et al. 2015). With the progressive drop of costs of these high-throughput technologies, large-scale expression data have been increasingly made available in public data repositories such as Gene Expression Omnibus (NCBI GEO) (Barrett et al. 2013; Edgar et al. 2002) and ArrayExpress (Athar et al. 2019). This opened the way to the so-called meta-analysis approaches, which identify differentially expressed genes (DEGs) or enriched pathways by integrating the information from multiple high dimensional expression datasets addressing the same biological question.

While single studies might lead to variable results because of technical and biological variability, a meta-analysis has instead the potential to provide more robust and generalizable biological insights (Ramasamy et al. 2008; Walsh et al. 2015). The choice of the meta-analysis approach depends on the biological questions of the study. Naive methods include Venn diagrams and vote counting. More principled methods, instead, identify the feature to integrate (e.g., the fold change or the *p*-value) and establish a statistical procedure to assess the significance of the integrated feature (Ramasamy et al. 2008; Tseng et al.

2012). The integration task can be performed either at gene level or at pathway level (Shen and Tseng 2010). In the meta-analysis at gene level, individual results are integrated right after the differential expression analysis and, once the integrated list of DEGs has been identified, pathway enrichment analysis can be performed on this list. Examples are the Integrative Microarray Analysis of Pathways (IMAP) approach by Setlur et al. (2007), where differential expression *p*-values are converted to z-scores and combined via a weighted average, and the MAPE_G algorithm (Shen and Tseng 2010), where the maximum *p*-value of each gene across all data sets is used as statistic. In the meta-analysis at pathway level, results are integrated after the pathway enrichment analysis. The assumption underlying this analysis is that it has more power in detecting dysregulated pathways irrespectively of potential differences in the expression of the pathway genes across studies, due to technical or biological variability of the studies. An example of pathway meta-analysis approach is the MAPE_P algorithm (Shen and Tseng 2010), which uses as statistic the maximum *p*-value associated with each pathway across datasets. Additionally, the hybrid MAPE_I framework, which integrates the results of MAPE_G and MAPE_P, was proposed by the same authors. Another proposed pathway meta-analysis approach is the 'global approach' by Thomassen et al. (2008). Here, authors used Gene Set Enrichment Analysis (GSEA) (Subramanian et al. 2005) to identify enriched gene sets from individual datasets. GSEA relies on the assumption that a pathway might be enriched even if not all genes associated to it are significantly dysregulated. Moreover, GSEA does not require setting an arbitrary cut-off on the significance of the differential expression of single genes. The enrichment degree of each pathway is quantified by the Normalized Enrichment Score (NES). In the approach by Thomassen et al., a meta-pathway list is built as the intersection of gene sets assessed in the individual datasets. Then, the mean value of the pathway rank across datasets and its associated empirical *p*-value are used to identify statistically significant meta-pathways. Here, we present a meta-analysis framework that extends the approach by Thomassen et al. In addition to the integration of pathways irrespectively of their up- or down-regulation in individual datasets, in our approach the potential presence of missing pathways in the individual studies results is handled. Our goal was to avoid *a priori* excluding pathways that, because of study variability, might not have been identified in a study, or might show in an individual study dysregulation in a different direction from the majority of other studies. Our meta-analysis framework uses the median ranking as non-parametric statistic for the integration and presents the results in the form of gene set enrichment networks, which offer a synthetic graphical overview and allow the identification of clusters of biologically related enriched meta-pathways. We relied on a recently published benchmarking framework to assess the performance of our method. The proposed meta-analysis framework was then applied to functional genomics data of cardiac hypertrophy.

Cardiac hypertrophy is the growth response of the heart to increased hemodynamic stimuli (e.g., pressure or volume overload), characterized by an increase in cardiomyocyte size and/or ventricular wall thickness. There are two forms of cardiac hypertrophy, physiological and pathological. Physiological cardiac hypertrophy is the adaptation to an increased demand for cardiac output (i.e., blood flow), for example in athletes or pregnant women. Physiological hypertrophy is fully reversible when the physiological requirements return to regular levels. In contrast, pathological cardiac hypertrophy can occur as adaptation of the heart to chronically increased pressure overload due to e.g., hypertension or myocardial infarction. It maintains cardiac function but it can lead to a progressive, maladaptive remodeling of the heart over time, often resulting in an irreversible loss of function (Bernardo et al. 2010; Nakamura and Sadoshima 2018; Shimizu and Minamino 2016).

Although cardiac hypertrophy has been a widely studied topic, there are still open questions, including what makes the adaptive morphological stage of pathological hypertrophy different from physiological hypertrophy, and at which time point the irreversible switch from compensated to decompensated hypertrophy occurs. Moreover, to our knowledge there has been no study that investigated the temporal dynamics of dysregulated pathways in both physiological and pathological cardiac hypertrophy. The goal of our meta-analysis was to provide a temporal signature of heart tissue remodeling and identify differences between the two types of hypertrophic conditions.

Results

Assessment of the performance of the proposed meta-analysis framework on benchmark datasets

Our proposed meta-analysis framework integrates enriched pathways. Hereafter, we will use the term 'expression set' to refer to the measured expression values in both case and control samples. As first step of our

framework, GSEA-based pathway enrichment analysis is applied to each of the individual expression sets to be integrated. GSEA analysis requires as input a gene list ranked according to a decreasing measure of differential expression and identifies 'gene sets', hereafter also referred to as 'pathways', which are over-represented either at the top (up-regulated) or bottom (down-regulated) of the ranked list. For each gene set, GSEA gives as output the NES, which measures the enrichment degree of the set also accounting for its size, as well as the set of 'leading edge genes', i.e., those genes that mostly contributed to make the pathway enriched. Individual GSEA results are ranked on the basis of decreasing NES, then a meta-pathway list is obtained as the union of gene set lists from the associated expression sets. The median rank of each meta-pathway across expression sets is calculated and its associated empirical *p*-value is estimated and corrected by the false discovery rate (FDR) procedure.

To evaluate the performance of the proposed meta-analysis method we relied on a very recently published benchmarking framework for gene set enrichment analysis methods (Geistlinger et al. 2020). This benchmarking framework provides a curated collection of transcriptomic datasets investigating different diseases as well as a 'relevance score ranking' for each disease, i.e., a list of disease-relevant gene sets ('gold standard list') with associated 'relevance scores'. In order to benchmark the output of an enrichment method, a ranking score (RS) is calculated by comparing the obtained gene set rankings ('ranking list') with the relevance score ranking. From the RS, a relative ranking score (RRS) can be calculated, which is the ratio (given as percentage) between RS and the theoretical optimum score reached in the case in which the ranking list is identical to the relevance score ranking. We applied our proposed median ranking method to the analysis of five expression sets investigating Alzheimer's disease (Table 1), which had an associated list of 57 gold standard KEGG gene sets. Using as input a list of 341 KEGG gene sets made available within the benchmarking framework, our method achieved RRS = 63.1%. As a comparison, we performed the meta-analysis also with the global approach by Thomassen et al. (2008) and with all MAPE approaches (MAPE_G, MAPE_P, and MAPE_I) proposed by Shen and Tseng (2010). Thomassen et al. obtained similar performance (RRS = 62.5%), MAPE_P slightly worse (RRS = 61.8%), while the other two MAPE strategies had clearly worse performance (RRS = 56.3% and 56.1% respectively for MAPE_G and MAPE_I).

Next, we evaluated how our method's performance varied with increasing number of integrated expression sets (Supplementary Figure 1). As baseline, we assessed the performance of the GSEA approach applied to the individual expression sets. A median RRS of 62.4% was obtained across the five sets, with a minimum and maximum value of 56.8% and 71.6% for GSE5281_HIP and GSE1297 respectively. We then considered all possible combinations of two, three, and four expression sets. While the median RRS remained approximately stable, the variability across the RRS values decreased as the number of integrated sets increased. Thus, the meta-analysis approach became increasingly robust with increasing number of integrated expression sets.

The variability in the performance obtained for each explored number of integrated expression sets suggested that the different 'quality' of the sets was a major factor influencing the performance of the meta-analysis results. To evaluate the influence of study quality on meta-analysis results, we utilized the RRS associated with GSEA results on the single expression sets as a measure of its quality. When applying the median ranking method to the two expression sets with the highest RRS (GSE1297: RRS = 71.6%, GSE5281_EC: RRS = 69.1%), RRS = 72.1% was achieved, which exceeded the RRSs associated with the individual sets. The RRS decreased to 68% when including the third-best expression set (GSE5281_VCX: RRS = 62.4%), but it was higher than the RRS of this set alone. With the fourth dataset (GSE16759: RRS = 59.8%) RRS further decreased to 67.8%, but remained higher than the RRS of both the third- and fourth-best expression set. Additionally, the RRS obtained with all five expression sets (RRS = 63.1%) exceeded the RRSs of three of the five individual studies. Taken together, these results suggested that the power of the meta-analysis improves when good quality expression sets are integrated. In addition, although the integration of low-quality expression sets decreases performance, the meta-analysis is still able to improve the performance of individual analyses on lower quality expression sets.

Retrieval of cardiac hypertrophy genome-wide expression datasets

In order to investigate heart tissue remodeling during cardiac hypertrophy, we searched for large-scale gene expression datasets on physiological and pathological cardiac hypertrophy in the public repositories NCBI GEO (Barrett et al. 2013; Edgar et al. 2002) and ArrayExpress (Athar et al. 2019). We focused on expression datasets measured via microarrays or RNA-Seq in *Mus musculus* or *Rattus norvegicus*. For inclusion in our analysis, physiological cardiac hypertrophy must have been induced by exercise training (e.g., swimming, treadmill running) and

Table 1: Benchmark expression sets.

Expression set identifier	GEO dataset accession number	Reference	Platform	Number of samples (cases/controls) in the benchmark framework	Relative Ranking Score (RRS)
GSE1297	GSE1297	Blalock et al. (2004)	Affymetrix Human Genome U133A	16 (7/9)	71.6%
GSE16759	GSE16759	Nunez-Iglesias et al. (2010)	Affymetrix Human Genome U133 Plus 2.0	8 (4/4)	59.8%
GSE5281_EC	GSE5281	Liang et al. (2007)	Affymetrix Human Genome U133 Plus 2.0	21 (9/12)	69.1%
GSE5281 _HIP				23 (10/13)	56.8%
GSE5281 _VCX				31 (19/12)	62.4%

Five Alzheimer's disease expression sets from the benchmark framework by Geistlinger et al. (2020) were used for assessment. For each expression set, its identifier, GEO accession number, reference to the associated publication, array platform, number of samples, and RRS calculated for the results of GSEA applied to the individual set are reported. Note, three expression sets are associated with the same GEO dataset and are differentiated by the brain region in which expression was measured (EC: entorhinal cortex, HIP: hippocampus, VCX: primary visual cortex); for dataset GSE1297 only cases relative to severe disease were considered, as in the benchmark framework.

pathological cardiac hypertrophy by transaortic constriction (TAC) or aortic banding. In total, 15 expression datasets were identified (Table 2). Among these datasets, some were time series (i.e., containing gene expression levels measured at different time points after the induction of hypertrophy) and some others were static. Here, an expression set contains gene expression measurements on samples belonging to a given time point, so that a static dataset coincided with an expression set, whereas a time series dataset contained as many expression sets as time points. Each expression set included both cardiac hypertrophy samples (cases) and controls.

Individual datasets were pre-processed starting from raw data, whenever available, in order to remove as much as possible any non-biological variability across studies attributable to pre-processing methods. Afterwards, differential expression analysis (comparing cases versus controls) of each expression set was performed, followed by GSEA-based pathway enrichment analysis based on the Gene Ontology (GO) Biological Process (BP) gene sets retrieved from the Molecular Signature Database (MSigDB https://www.gsea-msigdb.org/gsea/msigdb/collections.jsp) (Liberzon et al. 2015; Subramanian et al. 2005).

Table 2: Cardiac hypertrophy studies included in the meta-analysis.

Hypertrophy type	ArrayExpress accession	Reference	Platform	Sex
Pathological cardiac hypertrophy	E-GEOD-36074	Skrbic et al. (2013)	Affymetrix Mouse430_2	Male
	E-GEOD-6970	Witt et al. (2008)	Affymetrix Mouse430a_2	Mixed
	E-GEOD-24242	Park et al. (2011)	Affymetrix MoEx-1_0-st-v1	Male
	E-GEOD-76	//	Affymetrix MG_U74Av2	Not specified
	E-GEOD-1621	Zhao et al. (2004)	Affymetrix MG_U74Av2	Male
	E-GEOD-738	Strøm et al. (2004)	Affymetrix RG_U34A	Male
	E-GEOD-29446	Lee et al. (2011)	Illumina Genome Analyzer II	Male
	E-GEOD-35350	Yang et al. (2012)	Illumina HiSeq 2000	Male
	E-MTAB-1507	Papait et al. (2013)	AB 5500 Genetic Analyzer	Male
	E-MTAB-727	Song et al. (2012)	Illumina Genome Analyzer II	Male
	GSE95140	Nomura et al. (2018)	Illumina HiSeq 2500	Male
Physiological cardiac hypertrophy	E-GEOD-36330	Chung et al. (2012)	Affymetrix Mouse430_2	Female
	E-GEOD-77	//	Affymetrix MG_U74Av2	Not specified
	E-GEOD-739	Strøm et al. (2005)	Affymetrix RG_U34A	Male
	E-GEOD-776	Kang et al. (2004)	Affymetrix RG_U34A	Not specified
	E-MTAB-727	Song et al. (2012)	Illumina Genome Analyzer II	Male

ArrayExpress accession number, reference to literature (when available), technological platform, and sex of the animals for the studies included in the meta-analysis.

The number of dysregulated meta-pathways peaks 1–2 weeks after hypertrophy induction

One issue in performing the meta-analysis of cardiac hypertrophy is that the studies were differently designed. Static studies were performed at different time points after hypertrophy induction and temporal studies were performed with varying temporal grids. In order to increase the robustness of our study, we grouped expression sets performed at time points that represent the same hallmarks of heart tissue remodeling ('aggregated time points') (Table 3). For pathological cardiac hypertrophy, four aggregated time points were inferred on the basis of literature knowledge (Brenes-Castro et al. 2018; Chang et al. 2017; Fard et al. 2000; Mirotsou et al. 2006; Park et al. 2011; Van den Bosch et al. 2006): within three days (immediate pressure overload response), 7–10 days (hypertrophic cardiac growth phase), 2–4 weeks (sustained cardiac hypertrophy), and 6–8 weeks (late stage pathological cardiac hypertrophy). For physiological cardiac hypertrophy, instead, no consensus could be found in the literature regarding temporal hallmarks. Given that the time points of the collected datasets ranged from 1 to 7 weeks, the following three aggregated time points were defined: 1–2, 3–4, and 6–7 weeks. Meta-analysis was then performed on the expression sets assigned to each aggregated time point. The leading edge genes associated to each meta-pathway were taken as the union of the leading edge genes identified in the individual expression sets results. A meta-pathway was considered up-regulated if its median rank was less than the empirical null density distribution mode, otherwise a meta-pathway was considered down-regulated.

Next, we examined the number of statistically significant dysregulated meta-pathways (FDR < 0.1). During pathological cardiac hypertrophy, dysregulated meta-pathways could be detected from the very first time point (within three days) (Figure 1A). The number of meta-pathways reached a maximum at 7–10 days after hypertrophy induction, which thus appeared to be a critical stage in pathological hypertrophy. After this time point, the number of significantly dysregulated meta-pathways slowly decreased. In physiological cardiac hypertrophy our analysis identified over 40 significantly dysregulated meta-pathways at the first aggregated time point (1–2 weeks) (Figure 1B). In contrast to pathological cardiac hypertrophy, these showed a considerable prevalence of down-regulated meta-pathways. After this time point a decrease in the number of significantly dysregulated meta-pathways was observed. At 3–4 weeks, there were nine dysregulated meta-pathways, seven of which were up-regulated. At the last time point (6–7 weeks), no significantly dysregulated meta-pathways were found. Together, these data demonstrate that the number of dysregulated pathways peaks early after the induction of both pathological and physiological cardiac hypertrophy, and that the peak in physiological, but not pathological, cardiac hypertrophy is characterized by a higher proportion of down-regulated meta-pathways.

Table 3: Aggregated time points with associated datasets and number of expression sets.

Hypertrophy type	Aggregated time point	Associated datasets (number of expression sets)
Pathological cardiac hypertrophy	Within three days	E-GEOD-76 (4)
		E-GEOD-1621 (1)
		GSE95140 (1)
	7–10 days	E-GEOD-76 (1)
		E-GEOD-1621 (1)
		E-MTAB-1507 (1)
		E-GEOD-35350 (1)
		E-GEOD-29446 (1)
		E-MTAB-727 (1)
		GSE95140 (1)
		E-GEOD-24242 (1)
	2–4 weeks	E-GEOD-6970 (1)
		GSE95140 (2)
		E-GEOD-1621 (1)
		E-GEOD-36074 (1)
		E-GEOD-24242 (1)
	6–8 weeks	E-GEOD-738 (1)
		E-GEOD-76 (1)
		E-GEOD-29446 (1)
		GSE95140 (1)
Physiological cardiac hypertrophy	1–2 weeks	E-GEOD-77 (2)
		E-GEOD-36330 (1)
	3–4 weeks	E-GEOD-77 (3)
		E-GEOD-36330 (1)
		E-GEOD-776 (1)
		E-MTAB-727 (1)
	6–7 weeks	E-GEOD-776 (1)
		E-GEOD-739 (1)

Aggregated time points were identified in order to reconstruct cardiac remodeling over time. For each aggregated time point the associated datasets and the number of analyzed expression sets are shown.

Meta-analysis identifies well-known biological processes associated with pathological cardiac hypertrophy

In order to better highlight changes across time points and facilitate the identification of the biological mechanisms involved in heart tissue remodeling, we visualized

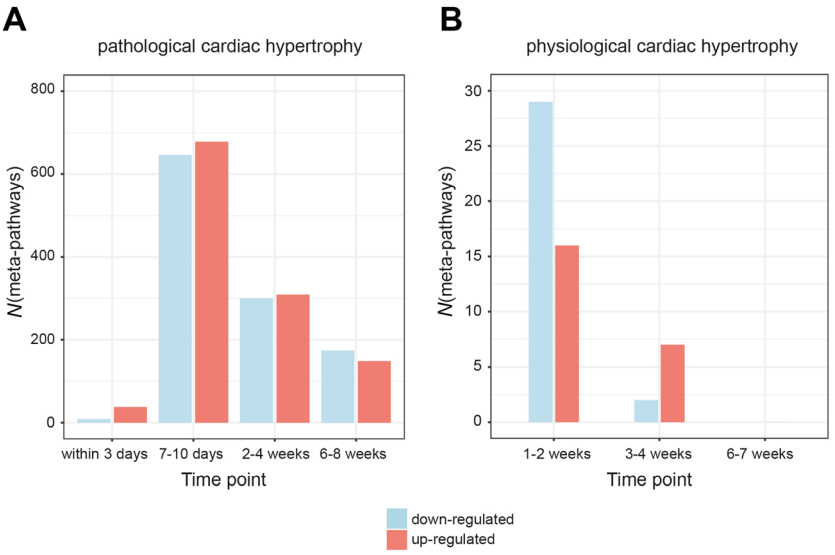


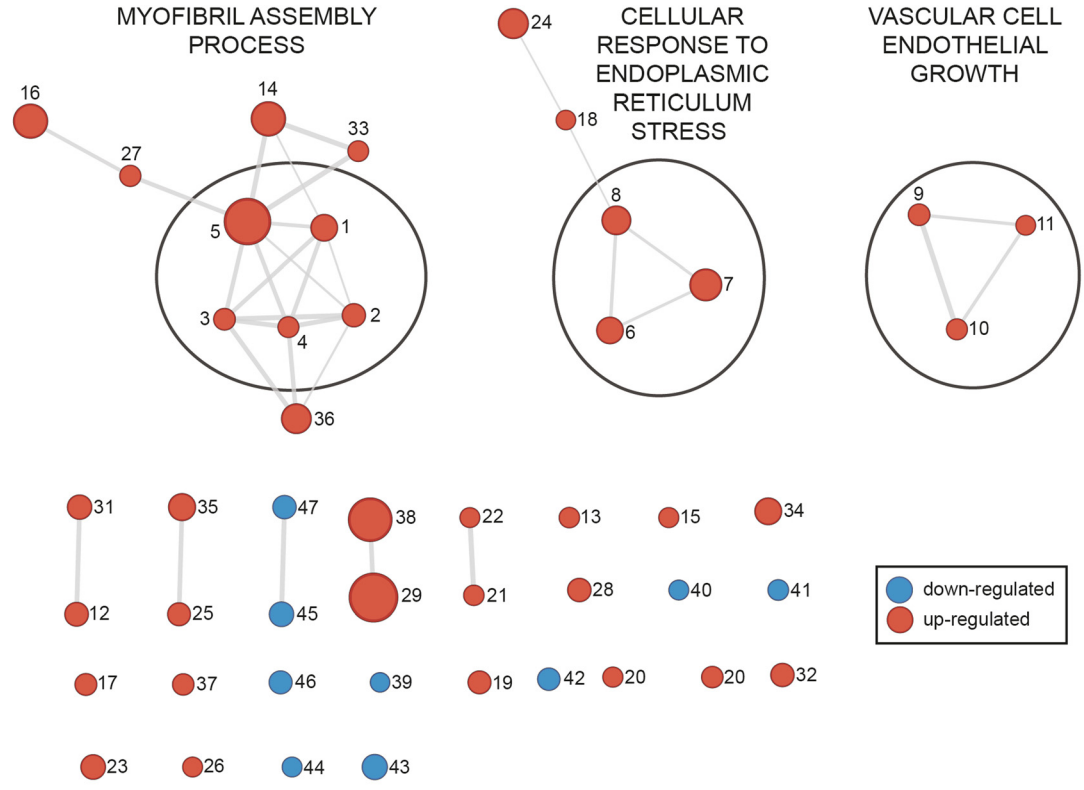
Figure 1: Number of dysregulated meta-pathways over time.
Number of significantly dysregulated (FDR < 0.1) meta-pathways identified for (A) pathological and (B) physiological cardiac hypertrophy for each aggregated time point.

significantly dysregulated meta-pathways as gene set enrichment networks. In these networks, nodes represent meta-pathways and edges the similarity between two connected meta-pathways, based on the proportion of shared genes. Such a network offers an overview of all significantly dysregulated meta-pathways and allows the identification of clusters of densely connected and thus biologically related meta-pathways. A cluster was considered up-regulated when the majority of the associated meta-pathways were up-regulated, otherwise it was considered down-regulated.

The enrichment network within three days from the induction of pathological cardiac hypertrophy contained three up-regulated clusters: 'MYOFIBRIL ASSEMBLY PROCESS', 'CELLULAR RESPONSE TO ENDOPLASMIC RETICULUM STRESS', and 'VASCULAR CELL ENDOTHELIAL GROWTH' (Figure 2; Supplementary Table 1). Hypertrophy, an increase in cardiomyocytes size and thus muscle mass, during its acute phase is characterized by the biosynthesis of new contractile proteins. These proteins assemble into myofibrils, which in turn organize into new sarcomeres to form an enhanced contractile apparatus able to sustain the increasing pressure overload (Bernardo et al. 2010). Accordingly, the enrichment network showed the up-regulation of meta-pathways related to actomyosin structure organization clustered together in 'MYOFIBRIL ASSEMBLY PROCESS'. At the molecular level, the leading edge genes of two out of five up-regulated meta-pathways belonging to the cluster contained the fetal gene *Acta1*, which encodes the actin alpha skeletal muscle. This is in agreement with the literature, which describes a reactivation of the fetal gene program in the maladaptive form of cardiac hypertrophy (Bernardo et al. 2010, 2012; Nakamura and Sadoshima 2018). The increased myofibril assembly in hypertrophic cardiomyocytes is in line with the up-regulation of the un-clustered meta-pathway 'GO_CYTOPLASMIC_TRANSLATION', which represents mRNA-translation by cytoplasmic ribosomes, as well as the clustered meta-pathways 'CELLULAR RESPONSE TO ENDOPLASMIC RETICULUM STRESS', as a result of persistent endoplasmic reticulum (ER) stress. The ER plays an essential role during protein-synthesis, by facilitating the translation and folding of newly formed proteins, as well as the degradation of misfolded proteins. A substantial increase in the protein-synthesis rate in growing cardiomyocytes must be balanced by increased protein-folding to avoid the accumulation of toxic misfolded proteins, which is mediated by the unfolded protein response (UPR), an adaptive cellular process to accommodate protein-folding stress (Blackwood et al. 2019). Upon activation, three signaling arms cooperate to restore cellular homeostasis, including protein kinase RNA-like ER kinase (PERK), inositol-requiring protein 1 (IRE1), and activating transcription factor 6 (ATF6). However, during severe and persistent protein-folding stress, such as during TAC-induced hypertrophy, the adaptation of UPR starts to fail (Wang et al. 2018b). This continued ER stress ultimately leads to heart failure by inducing inflammatory and apoptotic pathways to eliminate terminally affected cells (Blackwood et al. 2019; Okada et al. 2004; Wang et al. 2018a). In addition to enhanced protein translation, it has been shown that oxidative stress, energy deprivation, disordered calcium content, and inflammation lead to ER stress (Binder et al. 2019). It is worth noting that 'GO_CYTOPLASMIC_-TRANSLATION' is not integrated into the 'CELLULAR RESPONSE TO ENDOPLASMIC RETICULUM STRESS' cluster, as the similarity between the un-clustered meta-pathway and the clustered meta-pathways did not exceed the pre-set

overlap coefficient (OC) threshold. Indeed, only the gene *Nck1* was shared by the meta-pathways belonging to the cluster 'CELLULAR RESPONSE TO ER STRESS' and the meta-pathway 'GO_CYTOPLASMIC_TRANSLATION'.

The increase in cardiac muscle mass during hypertrophy requires a matching increase in the number of blood vessels to supply nutrients and oxygen. Consequently, we found an up-regulation of meta-pathways related to



1: GO_ACTOMYOSIN_STRUCTURE_ORGANIZATION
2: GO_CELLULAR_COMPONENT_ASSEMBLY_INVOLVED_IN_MORPHOGENESIS
3: GO_MYOFIBRIL_ASSEMBLY
4: GO_SARCOMERE_ORGANIZATION
5: GO_SUPRAMOLECULAR_FIBER_ORGANIZATION
6: GO_CELLULAR_RESPONSE_TO_TOPOLOGICALLY_INCORRECT_PROTEIN
7: GO_RESPONSE_TO_ENDOPLASMIC_RETICULUM_STRESS
8: GO_RESPONSE_TO_TOPOLOGICALLY_INCORRECT_PROTEIN
9: GO_ENDOTHELIAL_CELL_DEVELOPMENT
10: GO_ESTABLISHMENT_OF_ENDOTHELIAL_BARRIER
11: GO_REGULATION_OF_ENDOTHELIAL_CELL_DEVELOPMENT
12: GO_CELLULAR_TRANSITION_METAL_ION_HOMEOSTASIS
13: GO_EPIBOLY
14: GO_ACTIN_FILAMENT_ORGANIZATION
15: GO_NEGATIVE_REGULATION_OF_PROTEIN_LOCALIZATION_TO_CELL_PERIPHERY
16: GO_EXTRACELLULAR_STRUCTURE_ORGANIZATION
17: GO_PROTEIN_TRANSMEMBRANE_TRANSPORT
18: GO_CHAPERONE_MEDIATED_PROTEIN_COMPLEX_ASSEMBLY
19: GO_HOMOTYPIC_CELL_CELL_ADHESION
20: GO_PROTEASOMAL_UBIQUITIN_INDEPENDENT_PROTEIN_CATABOLIC_PROCESS
21: GO_PROTEIN_LOCALIZATION_TO_LYSOSOME
22: GO_PROTEIN_TARGETING_TO_LYSOSOME
23: GO_PLATELET_DEGRANULATION
24: GO_PROTEIN_FOLDING
25: GO_NEGATIVE_REGULATION_OF_RESPONSE_TO_WOUNDING
26: GO_PEPTIDE_CROSS_LINKING
27: GO_COLLAGEN_FIBRIL_ORGANIZATION
28: GO_CYTOPLASMIC_TRANSLATION
29: GO_CELL_ACTIVATION_INVOLVED_IN_IMMUNE_RESPONSE
30: GO_REGULATION_OF_EXTRINSIC_APOPTOTIC_SIGNALING_PATHWAY_VIA_DE ATH_DOMAIN_RECEPTORS
31: GO_TRANSITION_METAL_ION_HOMEOSTASIS
32: GO_INTEGRIN_MEDIATED_SIGNALING_PATHWAY
33: GO_ACTIN_FILAMENT_DEPOLYMERIZATION
34: GO_TRANSFORMING_GROWTH_FACTOR_BETA_RECEPTOR_SIGNALING_PATHWAY
35: GO_REGULATION_OF_RESPONSE_TO_WOUNDING
36: GO_STRIATED_MUSCLE_CELL_DIFFERENTIATION
37: GO_MONOCYTE_CHEMOTAXIS
38: GO_MYELOID_LEUKOCYTE_MEDIATED_IMMUNITY
39: GO_NEGATIVE_REGULATION_OF_SODIUM_ION_TRANSPORT
40: GO_WALKING_BEHAVIOR
41: GO_NEGATIVE_REGULATION_OF_DNA_RECOMBINATION
42: GO_POSITIVE_REGULATION_OF_NUCLEAR_DIVISION
43: GO_REGULATION_OF_POSTSYNAPTIC_MEMBRANE_POTENTIAL
44: GO_AXONEME_ASSEMBLY
45: GO_EAR_MORPHOGENESIS
46: GO_PROTEIN_LOCALIZATION_TO_CHROMOSOME
47: GO_INNER_EAR_MORPHOGENESIS

Figure 2: Gene set enrichment network within three days after pathological hypertrophy induction.
Nodes represent meta-pathways and edges the similarity between two connected meta-pathways, based on the proportion of shared genes. The three identified clusters are highlighted with circles. Details on the identified meta-pathways, together with the cluster assignment, are in Supplementary Table 1. Node size is proportional to the number of leading edge genes and edge thickness to the similarity (overlap coefficient [OC]).

vascular endothelial growth (cluster 'VASCULAR CELL ENDOTHELIAL GROWTH'). These meta-pathways included in the leading edge genes *Vegfa*, encoding the vascular endothelial growth factor A (VEGF-A), which has critical functions in myocardial angiogenesis (Nakamura and Sadoshima 2018; Oka et al. 2014).

At 7–10 days post-induction of pathological hypertrophy, over one thousand meta-pathways were found significantly dysregulated, with a similar number of up- and down-regulated clusters (Figure 3; Supplementary Table 2). As response to the acute insult caused by the TAC hypertrophic stimulus, hormones such as angiotensin II play a crucial role in triggering the cascade of events that can ultimately lead to heart failure. These events include pro-inflammatory cytokine secretion and inflammatory pathway activation (Brenes-Castro et al. 2018; Nakamura and Sadoshima 2018). Accordingly, our analysis revealed the up-regulation of various meta-pathways related to immune cell responses. The cluster 'CYTOKINE SECRETION' was found up-regulated. Meta-pathways of the cluster included leading edge genes encoding pro-inflammatory cytokines such as interleukin-6 (IL-6), interleukin-1 beta (IL-1 beta), and tumor necrosis factor (TNF-a). Cytokines secretion can induce monocyte chemotaxis and infiltration, responsible for promoting the inflammatory cascade (Brenes-Castro et al. 2018). Accordingly, the enrichment map showed also the up-regulation of the cluster 'MONOCYTE AND LEUKOCYTE CHEMOTAXIS' in addition to other immune-system related clusters ('T-CELL AND LEUKOCYTE ACTIVATION', 'T-CELL DIFFERENTIATION'). Transcription factor (TF) prediction analysis applied to the cluster 'T-CELL AND LEUKOCYTE ACTIVATION' identified the nuclear factor NF-kappa-B p105 subunit, encoded by *Nfkb1*, as the most significant TF (FDR = 1.27e-26). Its transcriptional activity is known to be involved in the hypertrophic growth of heart tissue through the activation of pro-inflammatory cytokines, fibrosis, and cell apoptosis (Bernardo et al. 2012; Gordon et al. 2011; Nakamura and Sadoshima 2018). Accordingly, the activation of the pro-fibrotic cascade releasing fibrogenic mediators associated with the initial adaptation of heart tissue to the persistent hemodynamic load (Kong et al. 2014) was reflected in the up-regulation of the cluster 'MAST CELL ACTIVATION INVOLVED IN IMMUNE RESPONSE'. Here, genes encoding the transforming growth factors TGF-beta-1, TGF-beta-2, and TGF-beta-3 were found among the leading edge genes. Together, these data show that 7–10 days post-induction pathological cardiac hypertrophy is associated with the up-regulation of different immune-modulatory pathways.

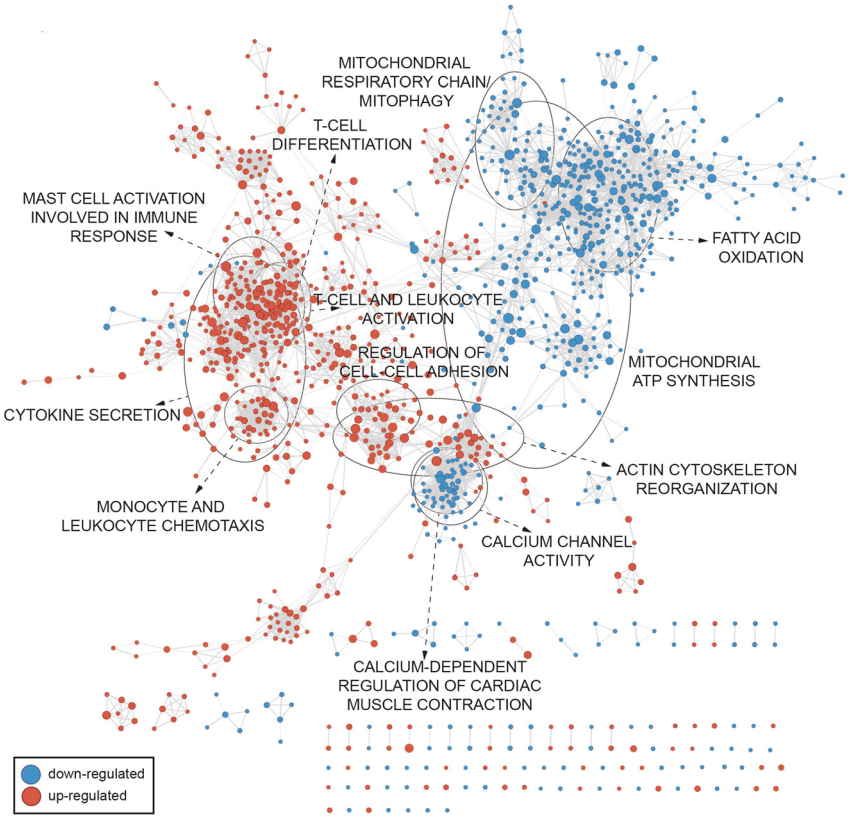


Figure 3: Gene set enrichment network at 7–10 days after pathological hypertrophy induction.
Nodes represent meta-pathways and edges the similarity between two connected meta-pathways, based on the proportion of shared genes. A subset of the identified clusters is highlighted with circles. Supplementary Table 2 lists all identified meta-pathways, together with the assignment to clusters. Node size is proportional to the number of leading edge genes and edge thickness to the similarity (OC).

Furthermore, it is known that extracellular matrix (ECM) remodeling occurs within the first weeks of pathological cardiac hypertrophy to counteract the increasing hemodynamic load. This process involves cytoskeleton components as well as transmembrane receptors such as integrins (Johnatty et al. 2000; Van den Bosch et al. 2006; Wakatsuki et al. 2004; Zhao et al. 2004). In line with this, the clusters 'ACTIN CYTOSKELETON REORGANIZATION' and 'REGULATION OF CELL-CELL ADHESION' together with the top (according to FDR) up-regulated un-clustered meta-pathways ('GO COLLAGEN FIBRIL ORGANIZATION', 'GO INTEGRIN MEDIATED SIGNALING PATHWAYS', 'GO POSITIVE REGULATION OF CELL SUBSTRATE ADHESION') suggest an up-regulation of processes that regulate ECM-remodeling.

Further distinctive features of pathological cardiac hypertrophy are the impairment of heart tissue contractility function and energy metabolism (Bernardo et al. 2010, 2012; Nakamura and Sadoshima 2018). In the 7–10 days network a dysregulation in heart tissue contractile function could be already detected. More specifically, an impairment in calcium handling due to a dysfunction of the sarcoplasmic/ER calcium ATPase 2 (SERCA2) and of the ryanodine receptor 2 (RYR-2) was associated to failing hearts (Tham et al. 2015). In our network the genes *Atp2a2* and *Ryr2*, which code for these two proteins, were found among the leading edge genes of the meta-pathways of both clusters 'CALCIUM-DEPENDENT REGULATION OF CARDIAC MUSCLE CONTRACTION' and 'CALCIUM CHANNEL ACTIVITY'. TF prediction analysis of the cluster 'CALCIUM-DEPENDENT REGULATION OF CARDIAC MUSCLE CONTRACTION' identified the nuclear respiratory factor 1 (NRF-1) as the top key TF (FDR = 0.003). NRF-1 regulates the expression of mitochondrial electron transport chain-related genes (Aubert et al. 2013). In addition, *Tbx5* and *Gata4* genes, which encode respectively the T-box TF TBX5 and the GATA-binding factor 4, were found within the top 10 TFs, with an FDR of 0.007. These two TFs are known to regulate genes whose expression is induced in response to pathological stimuli (e.g., natriuretic peptides *Anp*, *Bnp*) (Akazawa and Komuro 2003; Frey and Olson 2003; Nomura et al. 2018).

Changes in metabolic pathways that characterize pathological hypertrophy (Bernardo et al. 2010, 2012; Nakamura and Sadoshima 2018) were reflected by the down-regulation of the clusters 'FATTY ACID OXIDATION', 'MITOCHONDRIAL RESPIRATORY CHAIN/MITOPHAGY,' and 'MITOCHONDRIAL ATP SYNTHESIS'. In accordance with the known metabolic remodeling from fatty acid oxidation to glucose utilization, the peroxisome proliferator-activated receptor delta (PPAR-delta) and the peroxisome proliferator-activated receptor alpha (PPAR-alpha), were found among the top 10 predicted TFs of the down-regulated cluster 'FATTY ACID OXIDATION', with FDRs respectively of 7.76e-06 and 1.04e-05. Furthermore, TF prediction analysis performed on the cluster 'MITOCHONDRIAL ATP SYNTHESIS' identified the estrogen-related receptor alpha (ERR-alpha), encoded by *Essrα*, as one of the most significant key TFs (FDR = 0.01). These findings are in line with literature knowledge describing the nuclear receptors families PPARs and ERRs as playing a key role in mediating fatty acid oxidation and mitochondrial energy metabolism, respectively (Abel and Doenst 2011; Fan and Evans 2015; Huss et al. 2004; Nakamura and Sadoshima 2018). Of note was also the presence of the peroxisome proliferator-activated receptor gamma (PPAR-gamma) and the peroxisome proliferator-activated receptor gamma coactivator 1-alpha (PCG-1α), encoded by the gene *Ppargc1α*, within the top 10 predicted TFs regulating the 'MITOCHONDRIAL ATP SYNTHESIS' cluster with an FDR respectively of 0.03 and 0.047. PCG-1α is known to be one of the main transcriptional regulators governing mitochondrial activity and whose down-regulation significantly contributes to heart failure (Arany et al. 2006). Moreover, its transcriptional activity is known to promote ERRs expression (Huss et al. 2004). TF prediction analysis of the clusters 'FATTY ACID OXIDATION' and 'MITOCHONDRIAL ATP SYNTHESIS' also identified sterol regulatory element-binding protein 1 (SREBP-1), encoded by the gene *Srebf1*, as the most significant TF in both clusters, with an FDR of 2.25e-07 and 0.007 respectively. SREBP-1 regulates cholesterol and fatty acid synthesis (Engelking et al. 2018), is known to modulate PPAR-gamma activity (Kim et al. 1998), and is involved in fibrotic pathways in the kidney (Dorotea et al. 2020). While SREPB-1 has not been described so far in pathological hypertrophy, SREBP-1 expression correlates with cardiac PPAR-gamma expression, intramyocyte lipid accumulation, and reduced cardiac function in metabolic syndrome patients (Marfella et al. 2009). This suggests that interactions between PPARs and SREBP-1 might be involved in cardiac remodeling in pathological hypertrophy.

At 2–4 weeks after the induction of pathological cardiac hypertrophy the enrichment network retained prominent clusters related to actin cytoskeleton reorganization, cardiac contractility and energy metabolism (Figure 4; Supplementary Table 3). The network showed cardiac dysfunction (down-regulated cluster 'CALCIUM-DEPENDENT REGULATION OF CARDIAC MUSCLE CONTRACTION', up-regulated cluster 'APOPTOSIS'), as well as deficiency in mitochondrial ATP synthesis and fatty acid metabolism (e.g., down-regulated clusters 'MITOCHONDRIAL ATP SYNTHESIS' and 'FATTY ACID OXIDATION'). TF prediction analysis

identified ERR-alpha as the most significant key TF (FDR = 5.65e-06) regulating 'FATTY ACID OXIDATION' cluster genes, as well as the three peroxisome proliferator activator receptors PPAR-delta (FDR = 0.005), PPAR-alpha (FDR = 0.008), and PPAR-gamma (FDR = 0.01). PPAR-gamma and ERR-alpha were also found among the key TFs regulating the cluster 'MITOCHONDRIAL ATP SYNTHESIS', with an FDR of 1.88e-05 and 0.005 respectively. Similarly to the time point 7–10 days after TAC, we found SREBP-1 as the top TF regulating the cluster 'MITOCHONDRIAL ATP SYNTHESIS' (FDR = 1.88e-05) but also within the top five TFs (FDR = 3.91e-05) identified for the cluster 'FATTY ACID OXIDATION'. In addition, previous studies observed a slight increase in glucose utilization at early stages of pathological cardiac hypertrophy, followed by a decrease in glucose metabolism during maladaptive cardiac remodeling (Bernardo et al. 2010; Nakamura and Sadoshima 2018). Consistent with this, we observed the down-regulation of over thirty meta-pathways related to glucose oxidation clustered together in 'CARBOHYDRATE METABOLISM'.

The up-regulation of the cluster 'ACTIN CYTOSKELETON REORGANIZATION' suggests that myocardial tissue continues to grow. This growth is accompanied by an increase in blood vessels and capillarity density, as shown by the up-regulation of the cluster 'ENDOTHELIAL CELL MIGRATION INDUCED ANGIOGENESIS'. However, the concomitant up-regulation of the cluster 'APOPTOSIS' suggests the possibility that the supply of oxygen and nutrients may not be able to sustain the increase in cardiac muscle mass over time (Nakamura and Sadoshima 2018).

Surprisingly, as opposed to the 7–10 days aggregated time point, the enrichment network at 2–4 weeks did not show any of the large clusters related to cytokine secretion or monocyte/T-cell/mast cell activation/migration that were up-regulated at the earlier time point. Instead, the large cluster 'ANTIGEN PROCESSING' was the only up-regulated immune cell-related cluster. Prominent meta-pathways in this cluster were related to antigen processing and presentation as well as proteasome-dependent and -independent protein catabolism, indicating that different, i.e., adaptive rather than innate, immune-modulatory processes are active at this later stage of pathological cardiac hypertrophy. Predicted TFs for this cluster included the nuclear factor NF-kappa-B p105 subunit encoded by the gene *Nfkb1* (FDR = 0.008).

Dysregulated clusters similar to the 2–4 weeks time point were identified also 6–8 weeks after hypertrophy induction (Figure 5; Supplementary Table 4). We still observed the down-regulation of the 'CALCIUM-DEPENDENT REGULATION OF CARDIAC MUSCLE CONTRACTION' cluster. In addition, the network confirmed the up-regulation of cytoskeleton remodeling and the synthesis of new myofibrils to enhance the contractile apparatus ('ACTIN CYTOSKELETON REMODELING', 'STRESS FIBER AND FOCAL ADHESION FORMATION'), which seemed to accompany heart tissue pathogenesis up to these late stages. Mitochondrial dysfunction and deficiency in fatty acid metabolism persist

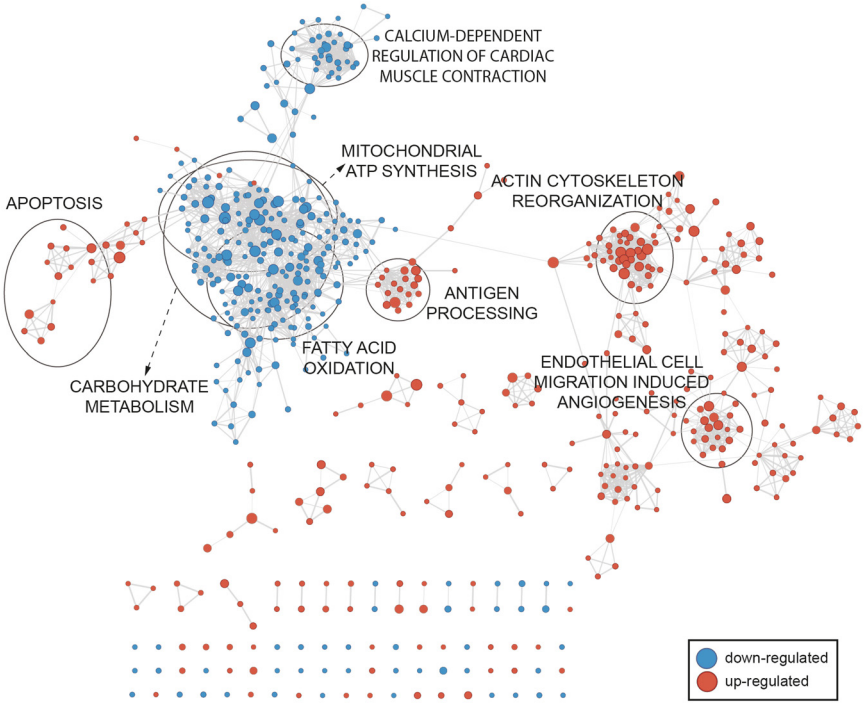


Figure 4: Gene set enrichment network at 2–4 weeks after pathological hypertrophy induction.
Nodes represent meta-pathways and edges the similarity between two connected meta-pathways, based on the proportion of shared genes. A subset of the identified clusters is highlighted with circles. Supplementary Table 3 lists all identified meta-pathways, together with the assignment to clusters. Node size is proportional to the number of leading edge genes and edge thickness to the similarity (OC).

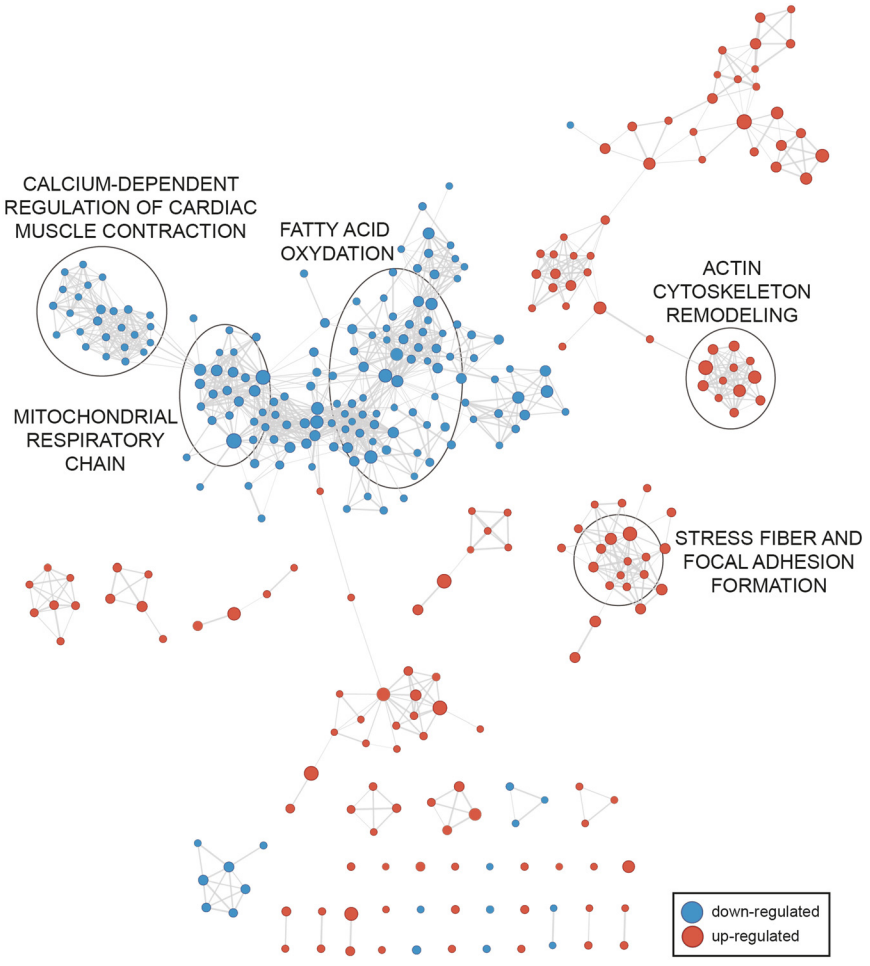


Figure 5: Gene set enrichment network at 6–8 weeks after pathological hypertrophy induction.
Nodes represent meta-pathways and edges the similarity between two connected meta-pathways, based on the proportion of shared genes. A subset of the identified clusters is highlighted with circles. Supplementary Table 4 lists all identified meta-pathways, together with the assignment to clusters. Node size is proportional to the number of leading edge genes and edge thickness to the similarity (OC).

in response to the maladaptive remodeling of heart tissue (Nakamura and Sadoshima 2018), as shown by the down-regulation of the 'MITOCHONDRIAL RESPIRATORY CHAIN' and 'FATTY ACID OXIDATION' clusters. Accordingly, PPAR-gamma (FDR = 4.47e-06), PPAR-alpha (FDR = 1.67e-05), ERR-alpha (FDR = 4.25e-05) and SREBP-1 (FDR = 3.98e-07) were found within the top 10 TFs regulating 'FATTY ACID OXIDATION' related genes, with SREBP-1 being the most significant TF during this late stage phase of cardiac remodeling.

Collectively, the results validated our meta-analysis framework demonstrating that it is well suited to identify biological mechanisms underlying cardiac hypertrophy.

Comparison with other meta-analysis methods on the aggregated time point 7–10 days after TAC

As further assessment of our framework, meta-analysis at 7–10 days after TAC was performed also with the global approach by Thomassen et al. and the MAPE framework. Thomassen et al. identified 1589 significantly dysregulated (FDR < 0.1) meta-pathways, with 1162 in common with our method. Differences with the results by MAPE were more pronounced. MAPE_G did not detect any significantly dysregulated meta-pathways, MAPE_P identified 898 significantly dysregulated meta-pathways and MAPE_I 434, 421 of which in common with MAPE_P. Four hundred and fourteen of the MAPE_P meta-pathways were in common with our results. Note, 135 significant meta-pathways identified by our method were not present in the results of MAPE_P and Thomassen et al. Indeed, these pathways were missing in the GSEA results relative to at least one expression set (because of the thresholds on minimum and maximum gene set size) and the other meta-analysis methods take into account only pathways in common to all expression sets. For the top 20 meta-pathways out of the 135, the proportion of expression sets in which a given meta-pathway was significantly dysregulated (FDR < 0.1) out of the total number of expression sets in which the pathway was assessed was calculated (Supplementary

Table 5). On average, the top 20 meta-pathways were significantly dysregulated in 70.8% of the expression sets. This shows that the fact that our method handles the presence of missing pathways in the individual studies results is beneficial in identifying additional potentially interesting meta-pathways.

Impact of time point aggregation on meta-analysis results

In order to investigate the impact of aggregating time points on meta-analysis results, we performed meta-analyses on expression sets associated with subgroups of the time points belonging to the aggregated time point 2–4 weeks after TAC. This aggregated time point included six expression sets: two measured at two weeks, one measured at three weeks, and three measured at four weeks. We evaluated the number of significant meta-pathways (FDR < 0.1) identified when progressively grouping expression sets associated with the time points 2, 3, and 4 weeks (Figure 6). Results showed that the number of significantly dysregulated meta-pathways increased with increasing number of grouped time points. Indeed, the meta-analysis of subgroup 2&4 weeks identified a higher number of significantly dysregulated meta-pathways compared to those recovered separately in the two subgroups 2 and 4 weeks. Meta-analysis on all expression sets (2–4 weeks) detected 609 meta-pathways. Here, most of the meta-pathways first identified in the subgroups 2, 4, and 2&4 weeks were conserved and additional 196 meta-pathways were identified. Taken together, a higher number of integrated time points appeared to improve the robustness of the obtained results.

Meta-analysis reveals novel, potentially cardioprotective mechanisms in physiological cardiac hypertrophy

Physiological hypertrophy is the adaptation of heart tissue to physiological needs, for example during exercise or pregnancy. It is a reversible condition characterized by enhanced cardiac function without the adverse effects associated with pathological cardiac hypertrophy such as cardiac fibrosis or cardiomyocyte apoptosis (Nakamura and Sadoshima 2018). At 1–2 weeks after induction of physiological cardiac hypertrophy, three down-regulated clusters were identified in the enrichment network (Figure 7; Supplementary Table 6). Surprisingly, one down-regulated cluster was 'ANTIGEN PROCESSING', as opposed to our previous finding of an up-regulation of various immune-related meta-pathways/clusters at the time points 7–10 days and 2–4 weeks in pathological hypertrophy. Notably, right at 2–4 weeks after pathological hypertrophy induction, the enrichment network showed the up-regulation of the 'ANTIGEN PROCESSING' cluster (Figure 4, Supplementary Table 3). This cluster contained 16 meta-pathways, eight of which were in common with the corresponding cluster for physiological hypertrophy, demonstrating that identical meta-pathways related to antigen processing were up-regulated in pathological hypertrophy but down-regulated in physiological hypertrophy at comparable time points. Similarly, the cluster

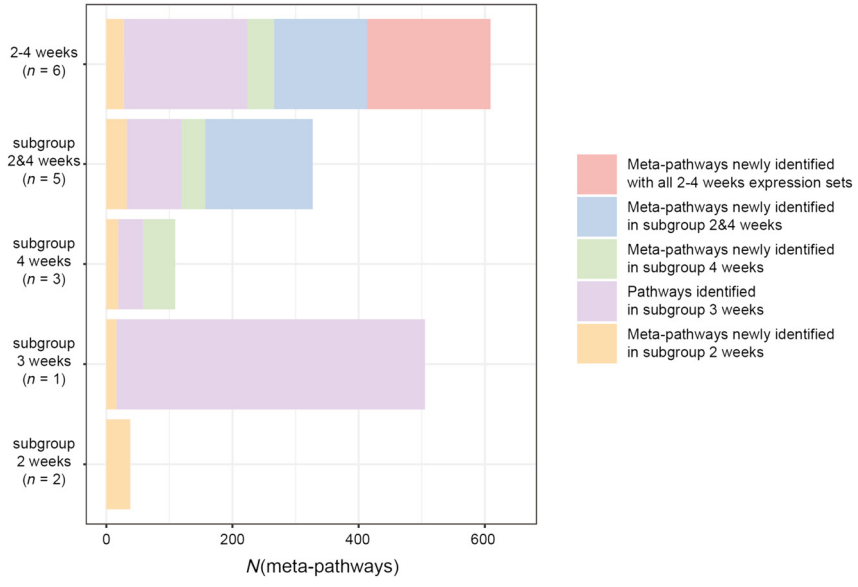


Figure 6: Impact of time point aggregation on identified meta-pathways.
Number of significantly dysregulated (FDR < 0.1) meta-pathways identified when performing meta-analysis on expression sets associated with subgroups of time points belonging to the aggregated time point 2–4 weeks after TAC. The barplots of the number of identified meta-pathways are colored according to the subgroups in which meta-pathways were first identified. For each subgroup the number *n* of jointly analyzed expression sets is shown.

'CELLULAR RESPONSE TO ENDOPLASMIC RETICULUM STRESS' (containing four meta-pathways) was down-regulated in physiological hypertrophy, while the cluster with the corresponding label in the enrichment network within three days from pathological hypertrophy, consisting of three out of these four meta-pathways, was up-regulated (Figure 2; Supplementary Table 1). TF prediction analysis of the down-regulated cluster 'ANTIGEN PROCESSING' identified the cellular tumor antigen p53 and the nuclear factor NF-kappa-B p105 subunit among the top predicted regulators (FDR = 2.13e-05). In addition, also the CCAAT/enhancer-binding protein beta (C/EBP beta), known as a key factor in physiological cardiac hypertrophy to protect against pathological remodeling (Bostrom et al. 2010; Nakamura and Sadoshima 2018), was found in the predicted TFs (FDR = 0.03) and was significantly down-regulated in all the three expression sets associated with this time point.

Taken together, these data indicate that meta-pathways with critical functions in antigen processing by immune cells and the ER stress response are regulated in an opposing manner in physiological and pathological cardiac hypertrophy, suggesting potential novel mechanisms that may differentiate between physiological and pathological cardiac hypertrophy on the molecular level.

Discussion

We have here proposed a novel framework for meta-analysis of enriched pathways from transcriptomics datasets. The assessment of our method on benchmark data showed that it had comparable or even better performance than the previously published MAPE approaches (Shen and Tseng 2010) and the method by Thomassen et al. (2008). It addition, our method appeared to be increasingly

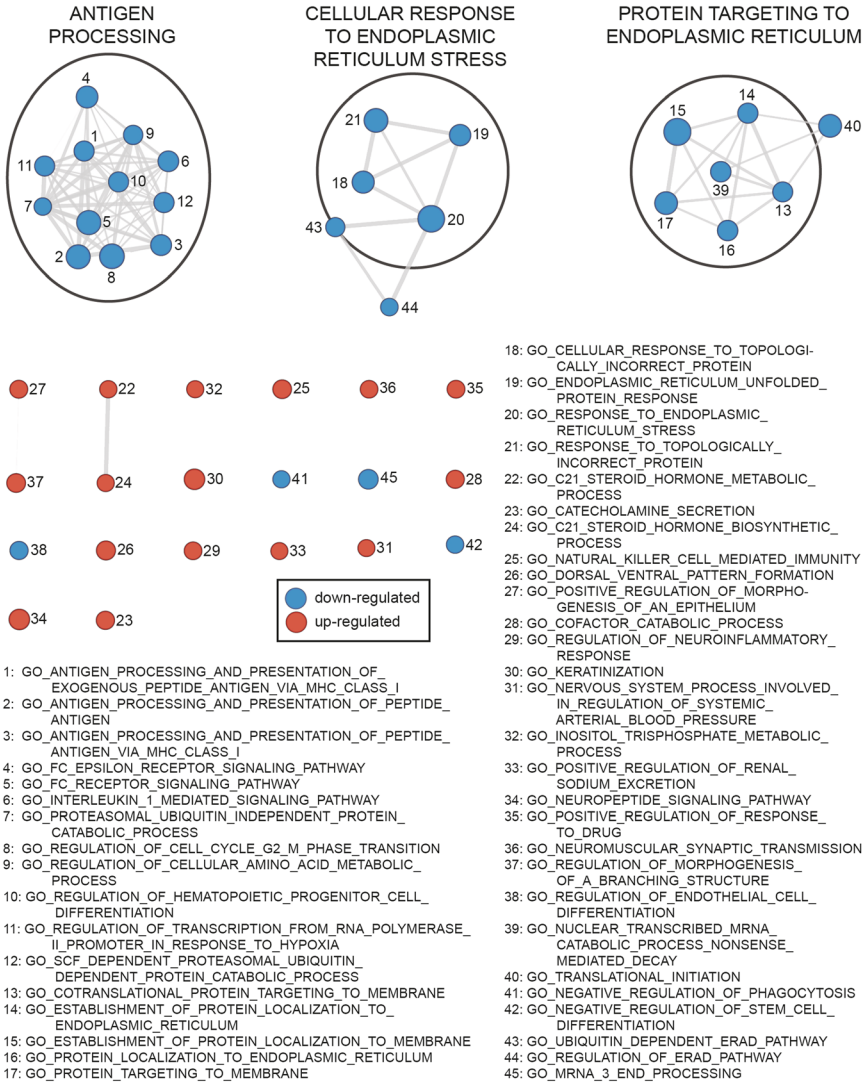


Figure 7: Gene set enrichment network at 1–2 weeks after physiological hypertrophy induction.
Nodes represent meta-pathways and edges the similarity between two connected meta-pathways, based on the proportion of shared genes. The three identified clusters are highlighted with circles. Details on the identified meta-pathways, together with the cluster assignment, are in Supplementary Table 6. Node size is proportional to the number of leading edge genes and edge thickness to the similarity (OC).

robust with increasing number of integrated studies. We then applied our proposed framework to expression datasets of physiological and pathological cardiac hypertrophy with the aim of dissecting mechanisms occurring over time during heart tissue response to hemodynamic overload and of identifying differences between the two types of hypertrophy. In pathological cardiac hypertrophy the decrease in the number of significantly dysregulated pathways over time may suggest the onset of a steady-state phase of cardiac hypertrophy, which might correspond to a late stage of pressure overload adaptation (Brenes-Castro et al. 2018; Van den Bosch et al. 2006). Results from pathological enrichment networks showed that the developed meta-analysis framework was able to detect mechanisms well-established in the literature, such as cardiomyocyte apoptosis, cardiac contractile dysfunction, and alteration in energy metabolism. The down-regulation of mitochondrial energy-related genes was detected also in previous transcriptomics studies, including microarray analyses (Witt et al. 2008; Park et al. 2011) and bulk RNA-seq analyses (Song et al. 2012; Yang et al. 2012). Note, in a single-cell RNA-seq analysis of mouse cardiomyocytes, an initial transient increase of mitochondrial biogenesis was observed (Nomura et al. 2018). The results of TF prediction analyses on network cluster genes correlated with literature knowledge on key TFs governing cardiac hypertrophy. In addition, comparison of the results of our meta-analysis framework at 7–10 days after TAC with Thomassen et al. and MAPE_P showed that handling the presence of missing pathways in the individual studies allows identifying additional potentially interesting meta-pathways.

Similarly to pathological cardiac hypertrophy, the number of dysregulated meta-pathways in physiological cardiac hypertrophy was highest at the first aggregated time point (1–2 weeks) and then decreased to zero at 6–7 weeks, suggesting that a steady-state condition might have been reached, in which the heart had fully adapted to the increased workload.

Interestingly, we observed an opposing regulation of meta-pathways related to immune-system activation in physiological and pathological hypertrophy in our analysis. The up-regulation of meta-pathways related to pro-inflammatory cytokine secretion, monocyte chemotaxis and T-cell recruitment, is consistent with literature describing immune cell infiltration of myocardial tissue during pathological cardiac hypertrophy. In addition, gene expression changes in immune-system related pathways, particularly the innate immune-system (i.e., leukocyte chemotaxis, migration, and cytokine signaling), have been recently detected in the analysis of a human aortic stenosis expression dataset (Yu et al. 2019). Instead, the same authors did not detect immune-system changes in a meta-analysis of expression datasets from early stage TAC models, while they detected them in the later stages by relaxing their meta-analysis criteria. Possible reasons for the differences to our study might be the different included datasets and time points as well as differences in the meta-analysis strategy. Notably, using our meta-analysis approach we found that following immune cell infiltration and differentiation at the 7–10 days time point of pathological hypertrophy, processes involving antigen processing and presentation were up-regulated. This suggests that immune cell infiltration of the hypertrophic myocardium is followed by the activation particularly of adaptive immune processes, which likely has significant impact on the progression towards maladaptive cardiac remodeling. Importantly, we found a down-regulation of the same meta-pathways at comparable time points during physiological cardiac hypertrophy, which, to our knowledge, has not yet been previously described. These data suggest the possibility that an active suppression of adaptive immune cell responses might serve to protect myocardial tissue from inflammatory and fibrotic processes during physiological cardiac hypertrophy and might represent a potential novel cardioprotective mechanism to explore in future studies.

Differently from pathological hypertrophy, at 1–2 weeks after induction of physiological hypertrophy the response to ER stress was down-regulated. This might indicate that, as opposed to pathological hypertrophy, metabolic or mechanical stresses under increased physiological load are not high enough to require efforts from ER to maintain protein homeostasis.

Datasets retrieval was a key phase in our meta-analysis, during which stringent inclusion criteria needed to be established in order to reduce inter-studies variability. Yet, while for pathological cardiac hypertrophy it was possible to focus only on experiments where hypertrophy was induced by TAC or aortic banding, for the physiological condition different experimental procedures used to induce hypertrophy (e.g., swimming, treadmill, running, wheeling), as well as different training protocols, had to be considered due to the lower number of datasets available in public repositories. Moreover, the hypertrophic stimulus induced by exercise is not as strong as the stimulus induced by TAC or aortic banding. This, together with datasets heterogeneity, might have impaired the identification of enriched meta-pathways in physiological hypertrophy. This impairment might be the reason why some pathways characteristic for physiological hypertrophy, such as those related to the enhanced fatty acid oxidation and glucose utilization (Bernardo et al. 2010;

Nakamura and Sadoshima 2018), were not identified. Future availability of a larger number of studies should mitigate these issues.

It is known that sex affects several aspects of heart disease (Gerdts and Regitz-Zagrosek 2019; Perrino et al. 2021). However, a comparative analysis of over 400 publications on TAC in rodent models did not find a significant gender effect (Bosch et al. 2020). Furthermore, Witt et al. observed no sex differences at two weeks utilizing a TAC model (Witt et al. 2008). However, they observed significant sex differences in left ventricular mass at 6 and 9 weeks. Gene expression analysis at two weeks showed similar up-regulation of classical hypertrophy marker genes for males and females. Yet, genes related to mitochondrial function, ECM-remodeling, and genes encoding ribosomal proteins were differentially regulated in males and females. The authors thus concluded that such changes might explain the sex differences at later stages. Here, we have not considered gender effects in order to maximize the number of studies to include in the meta-analysis. In the future it will be important to apply our meta-analysis approach considering the gender aspect, which would require the availability of a larger cohort of female datasets.

Taken together, our analysis in murine models revealed novel, potentially cardioprotective mechanisms in physiological cardiac hypertrophy. In addition to the experimental validation of our novel findings, a comparison of our results with those from a recently published temporal single-cell RNA-Seq analysis of cardiomyocyte remodeling (Nomura et al. 2018) would be of interest. Furthermore, it will be important in the future to validate our findings in humans.

Materials and methods

Benchmark datasets and benchmarking procedure

In order to assess the performance of our meta-analysis framework we used five Alzheimer's disease expression sets (Table 1) made available by the *GSEABenchmarkeR* package v. 1.8.0. Alzheimer's disease was associated with a gold standard list of 57 KEGG gene sets, each associated with a relevance score. Pre-processing and differential expression analysis of the five expression sets were conducted relying on the functions *maPreproc* and *runDE* provided by the *GSEABenchmarkeR* package choosing limma as *runDE* argument. GSEA, as implemented in the *fgseaSimple* function of the Bioconductor *fgsea* package v.1.14.0 (Korotkevich et al. 2021), was performed for each of the five expression sets utilizing the 341 gene sets belonging to the KEGG pathway database (Kanehisa et al. 2014) retrieved via the function *getGenesets* of the *EnrichmentBrowser* package v.2.18.2. Pathways resulting from meta-analysis on all, or a subset of, the five benchmark expression sets were ranked according to their *p*-value. Then the score of a pathway ranking (RS) was calculated using the function *evalRelevance* with the method weighted sum ('wsum') as argument. If some of the 341 pathways were missing in the obtained ranking, in order to calculate RS, *p*-value = 1 was assigned to the missing pathways. The RRS corresponding to RS was calculated using the theoretical optimum score obtained through the function *compOpt* applied to the entire list of 341 KEGG pathways.

Physiological and pathological cardiac hypertrophy datasets search and identification of aggregated time points for meta-analysis

Genome-wide expression datasets associated with both physiological and pathological cardiac hypertrophy were selected relying on the public data repositories Gene Expression Omnibus (NCBI GEO) (Barrett et al. 2013; Edgar et al. 2002) and ArrayExpress (Athar et al. 2019). The search was performed in April 2019 using as search keywords 'hypertrophy[Description]' for NCBI GEO and 'hypertrophy' for ArrayExpress. Datasets measured after TAC or aortic banding were chosen as models for pathological cardiac hypertrophy whereas datasets containing measurements in exercise training (e.g., running, swimming, treadmill) were chosen as models for physiological cardiac hypertrophy. Only studies containing oligonucleotide expression microarray or RNA-Seq data measured in *M. musculus* and *R. norvegicus* were kept. In addition, all studies involving transgenic test subjects or where hypertrophy was induced by knock-out experiments were excluded. In the included studies, staging was performed based on time after induction of hypertrophy. Successful induction of hypertrophy was validated in most cases by changes in function and wall thickness (echocardiography), morphology of isolated cardiomyocytes, heart weight to body weight, left ventricular weight to tibia length, and/or expression of typical hypertrophic markers (e.g., *Nppa*, *Nppb*, *Myh7*, *Acta1*, and/or *Pln*). Among the collected datasets, some were static and some others were time series experiments (i.e., containing samples referring to different time points). Thus, if a temporal dataset is characterized by expression values at *N* different time points, *N* is also the number of expression sets associated with that dataset. For static experiments dataset and expression set coincide. Each expression set contained both cases and controls: in pathological cardiac hypertrophy cases were test subjects that underwent TAC/aortic banding and controls were sham operated test subjects, in physiological cardiac hypertrophy cases were test subjects that underwent exercise training and controls were sedentary test subjects. Only expression sets with at least two biological replicates for both cases and controls were analyzed.

For both hypertrophy conditions we identified time points that represent 'hallmarks' of heart tissue remodeling. Expression sets performed at time points corresponding to the same hallmark were then grouped in 'aggregated time points' and jointly analyzed according to the developed meta-analysis framework.

Pre-processing and differential expression analysis of cardiac hypertrophy expression sets

For microarray experiments, whenever raw data were available, CEL files were retrieved from either NCBI GEO or ArrayExpress and pre-processed relying on the *justRMA* function of the Bioconductor *affy*

package v.1.66.0 (Gautier et al. 2004). Pre-processing and differential expression analysis were performed at probe set level. Probe set IDs were then annotated to gene symbols and Entrez IDs. Probe sets not mapping to any gene or mapping to more than one gene were removed (in the latter case they never constituted more than 9% of total probe sets). When different probe set IDs mapped to the same gene, the one with the highest absolute moderated t-statistic was chosen to represent the gene.

All RNA-Seq datasets were measured in *M. musculus* and raw data were downloadable for all from the European Nucleotide Archive (ENA, https://www.ebi.ac.uk/ena). After quality control performed using FastQC v.0.11.8 (Andrews 2010), raw sequence reads were quality-trimmed with Cutadapt v.2.3 (Martin 2011). Reads were then mapped against the *M. musculus* reference genome (Ensembl GRCm38 [Yates et al. 2020]) using the STAR aligner v.2.7.0f (Dobin et al. 2013), and a STAR genome directory created by supplying the Ensembl gtf annotation file (release 97) for GRCm38. STAR also provided raw read counts per gene.

For both microarray and RNA-Seq experiments differential expression analysis was performed within the R/Bioconductor environment v.4.0.2 (Gentleman et al. 2004; R Core Team 2020) relying on the *limma* package v.3.44.3 (Ritchie et al. 2015). The *limma* function *lmFit* was used to fit a linear model for each gene and the function *eBayes* to perform differential expression analysis by comparing, within each expression set, cases to controls. For each gene a moderated *t*-statistic, *p*-value and adjusted *p*-value are given as output.

Pathway enrichment analysis

Pathway enrichment analysis was performed relying on the GSEA approach (Subramanian et al. 2005), as implemented in the function *fgseaSimple* of the Bioconductor *fgsea* package v.1.14.0 (Korotkevich et al. 2021). The input to GSEA should be a list L of genes, ranked according to a decreasing measure of differential expression estimated in a genome-wide experiment. Statistically significant dysregulated gene sets, i.e., group of genes belonging a pre-defined biological category, are expected to be those with members primarily found either at the top or at the bottom of L, depending on whether they are up- or down-regulated. The degree by which a gene set S is over-represented at the top or at the bottom of the list is quantified by the Enrichment Score (ES): high positive ES characterizes an up-regulated gene set whereas a high negative ES is characteristic of a down-regulated gene set. A *p*-value is estimated via permutation. To account for differences in gene sets size, a normalization is performed yielding a NES. The genes that most contribute to make a given pathway enriched are identified and referred to as 'leading edge genes'.

GSEA analysis for each expression set was performed sorting genes according to decreasing moderated t-statistic values and with the following parameter settings: *nperm* = 10^5, *minSize* = 10, *maxSize* = 500, where *nperm* is the number of permutations performed to estimate the nominal *p*-value for a given pathway, *minSize* and *maxSize* are respectively the minimum and maximum size of a gene set to be tested. As query gene sets the GO (Ashburner et al. 2000; The Gene Ontology Consortium 2019) BP collection retrieved from the Molecular Signature Database (MSigDB, https://www.gsea-msigdb.org/gsea/msigdb/collections.jsp) (Liberzon et al. 2015; Subramanian et al. 2005) was utilized. Gene annotation files for *M. musculus* and *R. norvegicus* were prepared using the R packages *msigdbr* v.7.1.1 (Dolgalev 2020) and *msigdb* v.0.2.0 (Wang 2020).

Pathway meta-analysis approach: the median ranking method

Our proposed pathway meta-analysis approach, hereafter also referred to as median ranking method, relies on the following steps:

- **Ranking of individual lists.** The list of enriched pathways from each individual expression set are ranked according to decreasing NES, so that up-regulated pathways are at the top of the lists (NES > 0), down-regulated ones at the bottom (NES < 0), and those not significantly dysregulated are found in the middle positions with NES close to or equal to zero.
- **Building the meta-pathway list.** The meta-pathway list is taken as the union of all the individual lists of the associated expression sets. Taking the union avoids excluding gene sets that might be missing in an individual list as they have not satisfied the *maxSize* or *minSize* requirements.
- **Missing pathway imputation.** Pathways missing in individual lists are added by assigning to them NES = 0.
- **Median ranking calculation.** For each meta-pathway the median rank across the expression sets is computed.
- **Leading edge genes and assignment of the direction of regulation.** The 'leading edge genes' of a meta-pathway p_i are taken as the union of the leading edge genes identified in the individual expression sets. In the cardiac hypertrophy case study p_i was considered up-regulated if its median rank $m_i < M$, where M is the mode of the empirical null density distribution, otherwise it was considered down-regulated.
- **Significance assessment**: The significance of the enrichment of each meta-pathway is assessed by calculating an empirical *p*-value associated with the median ranking statistic.

To assess the significance, we assume that the null hypothesis H_0 is that the meta-pathway is not dysregulated and the alternative hypothesis H_1 is that the meta-pathway is dysregulated either up or down. Under H_0, the pathway rank in each individual list is assumed to be a random value between 1 and K, where K is the total number of analyzed gene sets (Thomassen et al. 2008). Thus, if N is the number of expression sets to aggregate, in order to build the null hypothesis distribution of the median ranking values across the N expression sets, a random sampling of N ranking values, uniformly distributed between one and K, was performed 10^7 times and the median value of the N values was calculated each time. Starting from this vector of median ranking values, both the empirical Probability Density Function (ePDF) and the empirical Cumulative Distribution Function (eCDF) were estimated through the R package *EnvStats* v.2.3.1 (Millard 2014). In order to estimate the *p*-value, the following procedure, valid also in the case of an asymmetric density distribution, was utilized:

- If the observed median ranking value x_1 is lower than the mode x_{mode} of the ePDF:
  - compute the area under the ePDF defined by the $(-\infty; x_1]$ interval;
  - find the value $x_2 > x_{mode}$ such that $\text{ePDF}(x_2) = \text{ePDF}(x_1)$;
  - compute the area defined by $[x_2; +\infty)$;
  - calculate the *p*-value as the sum of the two areas.

- If, instead, the observed median ranking value x_1 is greater than the mode x_{mode} of the ePDF:
  - compute the area under the ePDF defined by the $[x_1; \infty)$ interval;
  - find the value $x_2 < x_{mode}$ such as ePDF(x_2) = ePDF(x_1);
  - compute the area defined by $(-\infty; x_2]$;
  - calculate the *p*-value as the sum of the two areas.

The function *epdfPlot* from the *EnvStats* package was used to obtain x_2 and the *ecdfPlot* function was applied to estimate the *p*-value as sum of the two areas defined above. Empirical *p*-values were then corrected according to the Benjamini-Hochberg procedure (Benjamini and Hochberg 1995), as implemented in the *p.adjust* R function. A meta-pathway was considered statistically significant if its FDR < 0.1.

Gene set enrichment networks

Meta-analyses results for each aggregated time point of the cardiac hypertrophy study were visualized as gene set enrichment networks relying on the Enrichment Map plug-in v.3.2.1 (Merico et al. 2010) within Cytoscape v.3.7.1 (Shannon et al. 2003). In the network, nodes represent gene sets and edges the similarity between nodes, calculated by means of the OC (Merico et al. 2010). Given gene sets *A* and *B*, and the operator | |, where |*X*| is the number of genes belonging to gene set *X*, the OC is defined as:

$$OC = |A \cap B| / \min(|A|, |B|) \quad (1)$$

OC takes values between 0 and 1 and it is suitable for hierarchical gene sets collections as the GO. Indeed, as Eq. (1) shows, when all the genes belonging to a 'child' gene set are found in the 'parent' set, OC = 1. The following parameters were used in Enrichment Map: node FDR threshold = 0.1; OC threshold = 0.5.

To facilitate examination of the networks, node clusters were identified relying on the Molecular Complex Detection (MCODE) (Bader and Hogue 2003) plug-in, with the following parameters: haircut = false; fluff = false; node score threshold = 0.2; k-core threshold = 2; max depth from seed = 100. A cluster score is provided by MCODE, which is higher for more densely connected clusters. A cluster of meta-pathways was considered up-regulated when the majority of the belonging meta-pathways were up-regulated, otherwise it was considered down-regulated. A label for each cluster was identified by means of the AutoAnnotate plug-in (Kucera et al. 2016); this label was afterwards manually revised taking as reference the hierarchy of the cluster meta-pathways provided by the GO annotation browser QuickGO (https://www.ebi.ac.uk/QuickGO/) (Binns et al. 2009).

Transcription factor prediction analysis

The Transcriptional Regulatory Relationships Unraveled by Sentence-based Text mining (TRRUST) version 2 database (https://www.grnpedia.org/trrust/) (Han et al. 2018) was used to predict the key TFs associated with meta-pathway clusters, utilizing mouse as species. The genes obtained from the union of the leading edge genes belonging to the top 10 significantly dysregulated meta-pathways within a given cluster were provided in input to the tool as query genes. Significant TFs were identified considering FDR threshold = 0.05.

Implementation of the meta-analysis methods used for comparison

The 'global approach' by Thomassen et al. (2008) was implemented within the R/Bioconductor environment v.4.0.2, with the following differences with respect to our approach:

- The meta-pathway list was taken as the intersection (and not the union) of all the individual lists of the integrated expression sets;
- For each meta-pathway the mean (and not the median) rank across expression sets was computed.

To assess significance, a null hypothesis distribution of mean ranking values was created. Since not otherwise specified by the authors, the estimation of the empirical *p*-value was performed according to the same procedure established for our meta-analysis method. Empirical *p*-values were then corrected according to the FDR procedure relying on the *p.adjust* R function.

The Meta-Analysis for Pathway Enrichment (MAPE) framework by Shen and Tseng (2010) includes three different meta-analysis strategies, namely MAPE_G, in which meta-analysis is performed at gene level, MAPE_P, in which meta-analysis is performed at pathway level, and MAPE_I, which integrates results from MAPE_G and MAPE_P. For the implementation of the three MAPE strategies we relied on the R package *MetaPath* v.1.0 (Shen and Tseng 2015). Its main function *MAPE* takes as input the expression sets to be integrated, first it performs differential expression analysis and then performs meta-analysis according to all three strategies. In order to focus the comparison between MAPE and our method on the meta-analysis algorithms themselves, all compared meta-analysis methods should be based on the same differential expression analysis results. Thus, the *MAPE* code was modified to directly take as input previously calculated differential expression *p*-values.

For MAPE_G only genes in common across all datasets were considered and the maximum *p*-value (max*P*) statistic was used as integrative statistic, as proposed in the original paper (Shen and Tseng 2010). In order to assess the significance of the max*P* statistic, we relied on the fact that under the null hypothesis of a gene not differentially expressed, the *p*-value *p* is uniformly distributed over the range [0, 1] and max*P* calculated over *N* expression sets max$P = \text{argmax}_{1 \le n \le N}(p_n)$ follows a beta distribution with degrees of freedom $\alpha = N$, $\beta = 1$ (Qin and Lu 2018; Song and Tseng 2014). After performing meta-analysis at the gene level, pathway enrichment analysis was performed relying on the Kolmogorov–Smirnov (KS) statistics and statistical significance was assessed by performing a gene-wise permutation test, as implemented in the *Enrichment_KS_gene* function of the *MetaPath* package.

In the case of MAPE_P, first pathway enrichment analysis was performed separately on each expression set using the KS statistic and *p*-values computed through gene permutation tests. Then, the max*P* of the pathways in common across all datasets was taken as statistic for the meta-analysis and its *p*-value determined through a beta null distribution of parameters $\alpha = N$, $\beta = 1$. These *p*-values were then corrected according to the FDR method.

For MAPE_I, the pathways in common between the results of MAPE_G and MAPE_P were considered and the minimum of the two *p*-values (min*P*) was taken as statistic, according to (Shen and Tseng 2010). The *p*-value of min*P* was calculated relying on the fact that min$P = \text{argmin}_{1 \le n \le N}(p_n)$ follows a beta distribution with $\alpha = 1$, $\beta = N$ degrees of freedom (Song and Tseng 2014; Qin and Lu 2018). The *p*-values were then corrected according to the FDR procedure.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.
Research funding: None declared.
Conflict of interest statement: The authors declare that they have no conflicts of interests.

References

Abel, E.D. and Doenst, T. (2011). Mitochondrial adaptations to physiological vs. pathological cardiac hypertrophy. Cardiovasc. Res. 90: 234–242.

Akazawa, H. and Komuro, I. (2003). Roles of cardiac transcription factors in cardiac hypertrophy. Circ. Res. 92: 1079–1088.

Andrews, S. (2010). FASTQC. A quality control tool for high throughput sequence data, Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc.

Arany, Z., Novikov, M., Chin, S., Ma, Y., Rosenzweig, A., and Spiegelman, B.M. (2006). Transverse aortic constriction leads to accelerated heart failure in mice lacking PPAR-γ coactivator 1α. Proc. Natl. Acad. Sci. U.S.A. 103: 10086–10091.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., et al. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25: 25–29.

Athar, A., Füllgrabe, A., George, N., Iqbal, H., Huerta, L., Ali, A., Snow, C., Fonseca, N.A., Petryszak, R., et al. (2019). ArrayExpress update – from bulk to single-cell expression data. Nucleic Acids Res. 47: D711–D715.

Aubert, G., Vega, R.B., and Kelly, D.P. (2013). Perturbations in the gene regulatory pathways controlling mitochondrial energy production in the failing heart. Biochim. Biophys. Acta Mol. Cell Res. 1833: 840–847.

Bader, G.D. and Hogue, C.W. (2003). An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinf. 4: 2.

Barrett, T., Wilhite, S.E., Ledoux, P., Evangelista, C., Kim, I.F., Tomashevsky, M., Marshall, K.A., Phillippy, K.H., Sherman, P.M., Holko, M., et al. (2013). NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res. 41: D991–D995.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B 57: 289–300.

Bernardo, B.C., Weeks, K.L., Pretorius, L., and McMullen, J.R. (2010). Molecular distinction between physiological and pathological cardiac hypertrophy: experimental findings and therapeutic strategies. Pharmacol. Ther. 128: 191–227.

Bernardo, B.C., Ooi, J.Y.Y., and McMullen, J.R. (2012). The yin and yang of adaptive and maladaptive processes in heart failure. Drug Discov. Today Ther. Strat. 9: e163–e172.

Binder, P., Wang, S., Radu, M., Zin, M., Collins, L., Khan, S., Li, Y., Sekeres, K., Humphreys, N., Swanton, E., et al. (2019). Pak2 as a novel therapeutic target for cardioprotective endoplasmic reticulum stress response. Circ. Res. 124: 696–711.

Binns, D., Dimmer, E., Huntley, R., Barrell, D., O'donovan, C., and Apweiler, R. (2009). QuickGO: a web-based tool for Gene Ontology searching. Bioinformatics 25: 3045–3046.

Blackwood, E.A., Hofmann, C., Santo Domingo, M., Bilal, A.S., Sarakki, A., Stauffer, W., Arrieta, A., Thuerauf, D.J., Kolkhorst, F.W., Muller, O.J., et al. (2019). ATF6 regulates cardiac hypertrophy by transcriptional induction of the mTORC1 activator, Rheb. Circ. Res. 124: 79–93.

Blalock, E.M., Geddes, J.W., Chen, K.C., Porter, N.M., Markesbery, W.R., and Landfield, P.W., (2004). Incipient Alzheimer's disease: microarray correlation analyses reveal major transcriptional and tumor suppressor responses. Proc. Natl. Acad. Sci. U.S.A 101: 2173–2178.

Bosch, L., de Haan, J.J., Bastemeijer, M., van der Burg, J., van der Worp, E., Wesseling, M., Viola, M., Odille, C., El Azzouzi, H., Pasterkamp, G., et al. (2020). The transverse aortic constriction heart failure animal model: a systematic review and meta-analysis. Heart Fail. Rev., https://doi.org/10.1007/s10741-020-09960-w.

Bostrom, P., Mann, N., Wu, J., Quintero, P.A., Plovie, E.R., Panakova, D., Gupta, R.K., Xiao, C., MacRae, C.A., Rosenzweig, A., et al. (2010). C/EBPbeta controls exercise-induced cardiac growth and protects against pathological cardiac remodeling. Cell 143: 1072–1083.

Brenes-Castro, D., Castillo, E.C., Vázquez-Garza, E., Torre-Amione, G., and García-Rivas, G. (2018). Temporal frame of immune cell infiltration during heart failure establishment: lessons from animal models. Int. J. Mol. Sci. 19: 3719.

Chang, Y.-M., Ling, L., Chang, Y.-T., Chang, Y.-W., Li, W.-H., Shih, A.C.-C., and Chen, C.-C. (2017). Three TF co-expression modules regulate pressure-overload cardiac hypertrophy in male mice. Sci. Rep. 7: 1–13.

Chung, E., Heimiller, J., and Leinwand, L.A. (2012). Distinct cardiac transcriptional profiles defining pregnancy and exercise. PloS One 7: e42297.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29: 15–21.

Dolgalev, I. (2020). Msigdbr: MSigDB gene sets for multiple organisms in a tidy data format, Available at: https://CRAN.R-project.org/package=msigdbr.

Dorotea, D., Koya, D., and Ha, H. (2020). Recent insights into SREBP as a direct mediator of kidney fibrosis via lipid-independent pathways. Front. Pharmacol. 11: 265.

Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 30: 207–210.

Engelking, L.J., Cantoria, M.J., Xu, Y., and Liang, G. (2018). Developmental and extrahepatic physiological functions of SREBP pathway genes in mice. Semin. Cell Dev. Biol. 81: 98–109.

Fan, W. and Evans, R. (2015). PPARs and ERRs: molecular mediators of mitochondrial metabolism. Curr. Opin. Cell Biol. 33: 49–54.

Fard, A., Wang, C.Y., Takuma, S., Skopicki, H.A., Pinsky, D.J., Di Tullio, M.R., and Homma, S. (2000). Noninvasive assessment and necropsy validation of changes in left ventricular mass in ascending aortic banded mice. J. Am. Soc. Echocardiogr. 13: 582–587.

Ferrazzi, F., Bellazzi, R., and Engel, F.B. (2015). Gene network analysis: from heart development to cardiac therapy. Thromb. Haemostasis 113: 522–531.

Frey, N. and Olson, E. (2003). Cardiac hypertrophy: the good, the bad, and the ugly. Annu. Rev. Physiol. 65: 45–79.

Gautier, L., Cope, L., Bolstad, B.M., and Irizarry, R.A. (2004). affy—analysis of Affymetrix GeneChip data at the probe level. Bioinformatics 20: 307–315.

Geistlinger, L., Csaba, G., Santarelli, M., Ramos, M., Schiffer, L., Turaga, N., Law, C., Davis, S., Carey, V., Morgan, M., et al. (2020). Toward a gold standard for benchmarking gene set enrichment analysis. Briefings Bioinf. 22: 545–556.

Gentleman, R.C., Carey, V.J., Bates, D.M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., et al. (2004). Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 5: R80.

Gerdts, E. and Regitz-Zagrosek, V. (2019). Sex differences in cardiometabolic disorders. Nat. Med. 25: 1657–1666.

Gordon, J.W., Shaw, J.A., and Kirshenbaum, L.A. (2011). Multiple facets of NF-κB in the heart: to be or not to NF-κB. Circ. Res. 108: 1122–1132.

Han, H., Cho, J.W., Lee, S., Yun, A., Kim, H., Bae, D., Yang, S., Kim, C.Y., Lee, M., Kim, E., et al. (2018). TRRUST v2: an expanded reference database of human and mouse transcriptional regulatory interactions. Nucleic Acids Res. 46: D380–D386.

Huss, J.M., Torra, I.P., Staels, B., Giguere, V., and Kelly, D.P. (2004). Estrogen-related receptor α directs peroxisome proliferator-activated receptor α signaling in the transcriptional control of energy metabolism in cardiac and skeletal muscle. Mol. Cell Biol. 24: 9079–9091.

Johnatty, S.E., Dyck, J.R., Michael, L.H., Olson, E.N., and Abdellatif, M. (2000). Identification of genes regulated during mechanical load-induced cardiac hypertrophy. J. Mol. Cell. Cardiol. 32: 805–815.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014). Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 42: D199–D205.

Kang, P.M., Yue, P., Liu, Z., Tarnavski, O., Bodyak, N., and Izumo, S. (2004). Alterations in apoptosis regulatory factors during hypertrophy and heart failure. Am. J. Physiol. Heart Circ. Physiol. 287: H72–H80.

Kim, J.B., Wright, H.M., Wright, M., and Spiegelman, B.M. (1998). ADD1/SREBP1 activates PPARgamma through the production of endogenous ligand. Proc. Natl. Acad. Sci. U.S.A. 95: 4333–4337.

Kong, P., Christia, P., and Frangogiannis, N.G. (2014). The pathogenesis of cardiac fibrosis. Cell. Mol. Life Sci. 71: 549–574.

Korotkevich, G., Sukhov, V., Budin, N., Shpak, B., Artyomov, M.N., and Sergushichev, A. (2021). Fast gene set enrichment analysis. bioRxiv 060012, https://doi.org/10.1101/060012, Available at: https://www.biorxiv.org/content/10.1101/060012v3.

Kucera, M., Isserlin, R., Arkhangorodsky, A., and Bader, G.D. (2016). AutoAnnotate: a Cytoscape app for summarizing networks with semantic annotations. F1000Research 5: 1717.

Lee, J.-H., Gao, C., Peng, G., Greer, C., Ren, S., Wang, Y., and Xiao, X. (2011). Analysis of transcriptome complexity through RNA sequencing in normal and failing murine hearts. Circ. Res. 109: 1332–1341.

Liang, W.S., Dunckley, T., Beach, T.G., Grover, A., Mastroeni, D., Walker, D.G., Caselli, R.J., Kukull, W.A., McKeel, D., and Morris, J.C. (2007). Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. Physiol. Genom 28: 311–322.

Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J.P., and Tamayo, P. (2015). The molecular signatures database hallmark gene set collection. Cell Syst. 1: 417–425.

Marfella, R., Di Filippo, C., Portoghese, M., Barbieri, M., Ferraraccio, F., Siniscalchi, M., Cacciapuoti, F., Rossi, F., D'Amico, M., and Paolisso, G. (2009). Myocardial lipid accumulation in patients with pressure-overloaded heart and metabolic syndrome. J. Lipid Res. 50: 2314–2323.

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J 17: 10.

Merico, D., Isserlin, R., Stueker, O., Emili, A., and Bader, G.D. (2010). Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. PloS One 5: e13984.

Millard, S.P. (2013). EnvStats: an R package for environmental statistics. Springer, New York. ISBN 978-1-4614-8455-4.

Mirotsou, M., Dzau, V.J., Pratt, R.E., and Weinberg, E.O. (2006). Physiological genomics of cardiac disease: quantitative relationships between gene expression and left ventricular hypertrophy. Physiol. Genom. 27: 86–94.

Nakamura, M. and Sadoshima, J. (2018). Mechanisms of physiological and pathological cardiac hypertrophy. Nat. Rev. Cardiol. 15: 387–407.

Nomura, S., Satoh, M., Fujita, T., Higo, T., Sumida, T., Ko, T., Yamaguchi, T., Tobita, T., Naito, A.T., Ito, M., et al. (2018). Cardiomyocyte gene programs encoding morphological and functional signatures in cardiac hypertrophy and failure. Nat. Commun. 9: 4435.

Nunez-Iglesias, J., Liu, C.-C., Morgan, T.E., Finch, C.E., and Zhou, X.J. (2010). Joint genome-wide profiling of miRNA and mRNA expression in Alzheimer's disease cortex reveals altered miRNA regulation. PloS One 5: e8898.

Oka, T., Akazawa, H., Naito, A.T., and Komuro, I. (2014). Angiogenesis and cardiac hypertrophy: maintenance of cardiac function and causative roles in heart failure. Circ. Res. 114: 565–571.

Okada, K.-I., Minamino, T., Tsukamoto, Y., Liao, Y., Tsukamoto, O., Takashima, S., Hirata, A., Fujita, M., Nagamachi, Y., and Nakatani, T. (2004). Prolonged endoplasmic reticulum stress in hypertrophic and failing heart after aortic constriction: possible contribution of endoplasmic reticulum stress to cardiac myocyte apoptosis. Circulation 110: 705–712.

Papait, R., Cattaneo, P., Kunderfranco, P., Greco, C., Carullo, P., Guffanti, A., Viganò, V., Stirparo, G.G., Latronico, M.V.G., Hasenfuss, G., et al. (2013). Genome-wide analysis of histone marks identifying an epigenetic signature of promoters and enhancers underlying cardiac hypertrophy. Proc. Natl. Acad. Sci. U.S.A. 110: 20164–20169.

Park, J.Y., Li, W., Zheng, D., Zhai, P., Zhao, Y., Matsuda, T., Vatner, S.F., Sadoshima, J., and Tian, B. (2011). Comparative analysis of mRNA isoform expression in cardiac hypertrophy and development reveals multiple post-transcriptional regulatory modules. PloS One 6: e22391.

Perrino, C., Ferdinandy, P., Botker, H.E., Brundel, B., Collins, P., Davidson, S.M., den Ruijter, H.M., Engel, F.B., Gerdts, E., Girao, H., et al. (2021). Improving translational research in sex-specific effects of comorbidities and risk factors in ischaemic heart disease and cardioprotection: position paper and recommendations of the ESC Working Group on Cellular Biology of the Heart. Cardiovasc. Res. 117: 367–385.

Qin, W. and Lu, H. (2018). A novel joint analysis framework improves identification of differentially expressed genes in cross disease transcriptomic analysis. BioData Min. 11: 1–17.

R Core Team (2020). R: a language and environment for statistical computing, Available at: http://www.R-project.org/.

Ramasamy, A., Mondry, A., Holmes, C.C., and Altman, D.G. (2008). Key issues in conducting a meta-analysis of gene expression microarray datasets. PLoS Med. 5: e184.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 43: e47.

Setlur, S.R., Royce, T.E., Sboner, A., Mosquera, J.-M., Demichelis, F., Hofer, M.D., Mertz, K.D., Gerstein, M., and Rubin, M.A. (2007). Integrative microarray analysis of pathways dysregulated in metastatic prostate cancer. Canc. Res. 67: 10296–10303.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13: 2498–2504.

Shen, K. and Tseng, G.C. (2010). Meta-analysis for pathway enrichment analysis when combining multiple genomic studies. Bioinformatics 26: 1316–1323.

Shen, K. and Tseng, G.C. (2015). MetaPath: perform the meta-analysis for pathway enrichment analysis (MAPE). R package version 1.0.

Shimizu, I. and Minamino, T. (2016). Physiological and pathological cardiac hypertrophy. J. Mol. Cell. Cardiol. 97: 245–262.

Skrbic, B., Bjørnstad, J.L., Marstein, H.S., Carlson, C.R., Sjaastad, I., Nygård, S., Bjørnstad, S., Christensen, G., and Tønnessen, T. (2013). Differential regulation of extracellular matrix constituents in myocardial remodeling with and without heart failure following pressure overload. Matrix Biol. 32: 133–142.

Song, C. and Tseng, G.C. (2014). Hypothesis setting and order statistic for robust genomic meta-analysis. Ann. Appl. Stat. 8: 777–800.

Song, H.K., Hong, S.-E., Kim, T., and Kim, D.H. (2012). Deep RNA sequencing reveals novel cardiac transcriptomic signatures for physiological and pathological hypertrophy. PloS One 7: e35552.

Strøm, C.C., Kruhøffer, M., Knudsen, S., Stensgaard-Hansen, F., Jonassen, T.E.N., Orntoft, T.F., Haunsø, S., and Sheikh, S.P. (2004). Identification of a core set of genes that signifies pathways underlying cardiac hypertrophy. Comp. Funct. Genom. 5: 459–470.

Strøm, C.C., Aplin, M., Ploug, T., Christoffersen, T.E.H., Langfort, J., Viese, M., Galbo, H., Haunsø, S., and Sheikh, S.P. (2005). Expression profiling reveals differences in metabolic gene expression between exercise-induced cardiac effects and maladaptive cardiac hypertrophy. FEBS J. 272: 2684–2695.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc. Natl. Acad. Sci. U.S.A. 102: 15545–15550.

Tham, Y.K., Bernardo, B.C., Ooi, J.Y., Weeks, K.L., and McMullen, J.R. (2015). Pathophysiology of cardiac hypertrophy and heart failure: signaling pathways and novel therapeutic targets. Arch. Toxicol. 89: 1401–1438.

The Gene Ontology Consortium (2019). The gene ontology resource: 20 years and still GOing strong. Nucleic Acids Res. 47: D330–D338.

Thomassen, M., Tan, Q., and Kruse, T.A. (2008). Gene expression meta-analysis identifies metastatic pathways and transcription factors in breast cancer. BMC Canc. 8: 1–12.

Tseng, G.C., Ghosh, D., and Feingold, E. (2012). Comprehensive literature review and statistical considerations for microarray meta-analysis. Nucleic Acids Res. 40: 3785–3799.

Van den Bosch, B., Lindsey, P., Van den Burg, C., Van der Vlies, S., Lips, D., Van der Vusse, G., Ayoubi, T., Doevendans, P., and Smeets, H. (2006). Early and transient gene expression changes in pressure overload-induced cardiac hypertrophy in mice. Genomics 88: 480–488.

Wakatsuki, T., Schlessinger, J., and Elson, E.L. (2004). The biochemical response of the heart to hypertension and exercise. Trends Biochem. Sci. 29: 609–617.

Walsh, C.J., Hu, P., Batt, J., and Santos, C.C.D. (2015). Microarray meta-analysis and cross-platform normalization: integrative genomics for robust biomarker discovery. Microarrays 4: 389–406.

Wang, M. (2020). msigdb: MSigDB gene set collections, Available at: https://rdrr.io/github/mw201608/msigdb/.

Wang, S., Binder, P., Fang, Q., Wang, Z., Xiao, W., Liu, W., and Wang, X. (2018a). Endoplasmic reticulum stress in the heart: insights into mechanisms and drug targets. Br. J. Pharmacol. 175: 1293–1304.

Wang, X., Xu, L., Gillette, T.G., Jiang, X., and Wang, Z.V. (2018b). The unfolded protein response in ischemic heart disease. J. Mol. Cell. Cardiol. 117: 19–25.

Witt, H., Schubert, C., Jaekel, J., Fliegner, D., Penkalla, A., Tiemann, K., Stypmann, J., Roepcke, S., Brokat, S., Mahmoodzadeh, S., et al. (2008). Sex-specific pathways in early cardiac response to pressure overload in mice. J. Mol. Med. (Berl.) 86: 1013–1024.

Yang, K.-C., Ku, Y.-C., Lovett, M., and Nerbonne, J.M. (2012). Combined deep microRNA and mRNA sequencing identifies protective transcriptomal signature of enhanced PI3Kα signaling in cardiac hypertrophy. J. Mol. Cell. Cardiol. 53: 101–112.

Yates, A.D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M.R., Armean, I.M., Azov, A.G., Bennett, R., et al. (2020). Ensembl 2020. Nucleic Acids Res. 48: D682–D688.

Yu, P., Zhang, B., Liu, M., Yu, Y., Zhao, J., Zhang, C., Li, Y., Zhang, L., Yang, X., and Jiang, H. (2019). Transcriptome analysis of hypertrophic heart tissues from murine transverse aortic constriction and human aortic stenosis reveals key genes and transcription factors involved in cardiac remodeling Induced by mechanical stress. Dis. Markers 2019: 5058313, https://doi.org/10.1155/2019/5058313. eCollection 2019.

Zhao, M., Chow, A., Powers, J., Fajardo, G., and Bernstein, D. (2004). Microarray analysis of gene expression after transverse aortic constriction in mice. Physiol. Genom. 19: 93–105.

Supplementary Material: The online version of this article offers supplementary material (https://doi.org/10.1515/hsz-2020-0378).