**BMC
Genomics**

Systems analysis of gene ontology and biological pathways involved in post-myocardial infarction responses

Nguyen T Nguyen[1,2], Merry L Lindsey[2,3,4], Yu-Fang Jin[1,2]*

Abstract

Background: Pathway analysis has been widely used to gain insight into essential mechanisms of the response to myocardial infarction (MI). Currently, there exist multiple pathway databases that organize molecular datasets and manually curate pathway maps for biological interpretation at varying forms of organization. However, inconsistencies among different databases in pathway descriptions, frequently due to conflicting results in the literature, can generate incorrect interpretations. Furthermore, although pathway analysis software provides detailed images of interactions among molecules, it does not exhibit how pathways interact with one another or with other biological processes under specific conditions.

Methods: We propose a novel method to standardize descriptions of enriched pathways for a set of genes/proteins using Gene Ontology terms. We used this method to examine the relationships among pathways and biological processes for a set of condition-specific genes/proteins, represented as a functional biological pathway-process network. We applied this algorithm to a set of 613 MI-specific proteins we previously identified.

Results: A total of 96 pathways from Biocarta, KEGG, and Reactome, and 448 Gene Ontology Biological Processes were enriched with these 613 proteins. The pathways were represented as Boolean functions of biological processes, delivering an interactive scheme to organize enriched information with an emphasis on involvement of biological processes in pathways. We extracted a network focusing on MI to demonstrate that tyrosine phosphorylation of Signal Transducer and Activator of Transcription (STAT) protein, positive regulation of collagen metabolic process, coagulation, and positive/negative regulation of blood coagulation have immediate impacts on the MI response.

Conclusions: Our method organized biological processes and pathways in an unbiased approach to provide an intuitive way to identify biological properties of pathways under specific conditions. Pathways from different databases have similar descriptions yet diverse biological processes, indicating variation in their ability to share similar functional characteristics. The coverages of pathways can be expanded with the incorporation of more biological processes, predicting involvement of protein members in pathways. Further, detailed analyses of the functional biological pathway-process network will allow researchers and scientists to explore critical routes in biological systems in the progression of disease.

* Correspondence: yufang.jin@utsa.edu
[1]Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, TX 78249, USA
Full list of author information is available at the end of the article

BioMed Central

Background

The emergence of publicly available pathway databases has provided biologists excellent resources to attain a deeper understanding of biological mechanisms by providing organization to a large list of differentially expressed genes and proteins. Knowledge of molecular-level interactions and reactions has been curated in many knowledge databases, forming biological pathways. These knowledge databases include BioCarta (http://biocarta.com/), Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome, Protein Analysis Through Evolutionary Relationships (PANTHER), and MetaCyc [1-5]. Most often, pathways are organized as directed graphs of interacting molecules and often are accompanied by visualizations that demonstrate relationships among gene products, gene function types (e.g., regulation, activation, and inhibition) and translated protein locations (e.g., extracellular matrix, cell membrane, or nucleus). Recently, the integration of various omics data such as proteomics, genomics, transcriptomics, and metabolomics for knowledge discovery has drawn much attention [6-9]. In addition to the aforementioned pathway knowledge databases, the Gene Ontology (GO) Consortium pursues approaches to standardize the representation of gene products across different species and databases [10]. GO consists of a controlled vocabulary of terms, covering three domains: cellular components, molecular functions and biological processes. A GO Biological Process (GOBP) is a series of molecular events, with a defined beginning and end. However, a biological process is not equivalent to a pathway; GOBPs are assumed to be independent and do not represent the interactions among molecules.

Despite manual curation and careful revision, different knowledge databases could have different descriptions, participating molecules, interacting diagrams, and supporting literature for similar pathways. For example, considering the Transforming Growth Factors Beta (TGF-beta) signaling pathway in human, KEGG reported as *hsa04350: TGF-beta signaling pathway*, Reactome reported as *REACT_111102.4: Signaling by TGF-beta Receptor Complex*, and Biocarta reported as *h_tgfbpathway*. In detail, KEGG annotated 80 genes/proteins, Reactome annotated 120 genes/proteins, and Biocarta annotated 17 genes/proteins with TGF-beta signaling pathway. Descriptions of TGF-beta signaling pathway in the nucleus were excerpted to show related yet distinctive contents among KEGG, Reactome and Biocarta databases (Material in quote marks and italic type represents verbatim quotation from the knowledge databases):

KEGG - "*Once phosphorylated, R-Smads associate with the co-mediator Smad, Smad4, and the heteromeric complex then translocates into the nucleus. In the nucleus, Smad complexes activate specific genes through cooperative interactions with other DNA-binding and coactivator (or co-repressor) proteins*".

(http://www.genome.jp/kegg-bin/show_pathway?hsa04350)

Reactome - "*The general signaling scheme is rather simple: upon binding of a ligand, an activated plasma membrane receptor complex is formed, which passes on the signal towards the nucleus through a phosphorylated receptor SMAD (R-SMAD). In the nucleus, the activated R-SMAD promotes transcription in complex with a closely related helper molecule termed Co-SMAD (SMAD4)*".

(http://www.reactome.org/PathwayBrowser/#DIAGRAM = 170834&PATH = 162582)

Biocarta - "*The activated TGF-beta R1 phosphorylates SMAD2 and SMAD3, which bind to the SMAD4 mediator to move into the nucleus and form complexes that regulate transcription. SMADs regulate transcription in several ways, including binding to DNA, interacting with other transcription factors, and interacting with transcription corepressors and coactivators like p300 and CBP*".

(http://www.biocarta.com/pathfiles/h_tgfbpathway.asp).

These variations in knowledge representation among different databases prompt an urgent need for standard pathway representations. For a set of proteins or genes with enriched pathways and GOBPs, we propose a method that integrates molecular interaction, biological pathways and GOBP to standardize descriptions of pathways using GOBPs through the establishment of the functional biological pathway-process network. We demonstrated with the set of 613 proteins related to myocardial infarction (MI) from the MI-specific protein-protein interaction network [11].

Methods

In this study, we started with 613 MI-specific proteins to find enriched pathways and GOBPs [11]. We performed analyses to statistically examine the similarities between pathways and biological processes and identify the hierarchical structures for the GOBPs. Based on the similarity score matrix and the structure of GOBPs, we established the logical circuitry between GOBPs and pathways, and visualize the circuitry with networks.

Selection of condition-specific genes/proteins

We previously identified 613 proteins specific to MI in an MI-specific protein-protein interaction network (MIPIN); the network and its protein members were used here to demonstrate the developed method [11].

Functional annotation analysis

Many tools are available to provide gene-annotation enrichment analysis and pathway mapping. We performed functional annotation analysis using DAVID

Functional Annotation Tool, with the parameters *Count* to be 2 and *EASE* to be 0.05, to obtain enriched GOBP terms, KEGG and Reactome pathways [12].

Statistical measure of inter-annotator agreement

We evaluated the pairwise similarity between different annotation terms, including GO terms and pathways using Kappa statistics because annotation terms sharing common members might be related to one another [13]. Considering a set of all annotated genes/proteins G, two annotation terms T_i and T_j annotated by two set of genes G_i and G_j ($i \neq j$; $i,j = 1, 2, ..., N$), we denoted the number of proteins annotated by both terms as a_{ij}, the number of proteins annotated by T_i but not T_j as b_{ij}, the number of proteins annotated by T_j but not T_i as c_{ij}, and the number of proteins not annotated by neither terms among the union of proteins annotated by N annotation terms as d_{ij}.

Thus, we have,

$$G = \bigcup_N G_i, a_{ij} = G_i \cap G_j, b_{ij} = G_i \backslash G_j, c_{ij} = G_j \backslash G_i, d_{ij} = G \backslash (G_i \cup G_j)$$

The Kappa score κ_{ij} was defined as,

$$\kappa_{ij} = \frac{\Pr(agree_{ij}) - \Pr(random_{ij})}{1 - \Pr(random_{ij})},$$

where $\Pr(agree_{ij})$ was the observed percentage agreement and $\Pr(random_{ij})$ was the overall probability of random agreement for annotation terms T_i and T_j. The observed percentage agreement $\Pr(agree_{ij})$ could be calculated as follows,

$$P(agree_{ij}) = \frac{a_{ij} + d_{ij}}{a_{ij} + b_{ij} + c_{ij} + d_{ij}}.$$

Out of total number of associated proteins, T_i annotates $(a_{ij} + b_{ij})/(a_{ij} + b_{ij} + c_{ij} + d_{ij})$ and T_j annotates $(a_{ij} + c_{ij})/(a_{ij} + b_{ij} + c_{ij} + d_{ij})$. Thus, the probability that both annotation terms randomly annotate the same proteins was $(a_{ij} + b_{ij})(a_{ij} + c_{ij})/(a_{ij} + b_{ij} + c_{ij} + d_{ij})^2$. Similarly, the probability that neither pathway randomly annotate the same protein was $(b_{ij} + d_{ij})(c_{ij} + d_{ij})/(a_{ij} + b_{ij} + c_{ij} + d_{ij})^2$. As a result, the overall probability of random agreement $\Pr(random_{ij})$ could be calculated as,

$$P(random_{ij}) = \frac{(a_{ij} + b_{ij})(a_{ij} + c_{ij}) + (b_{ij} + d_{ij})(c_{ij} + d_{ij})}{(a_{ij} + b_{ij} + c_{ij} + d_{ij})^2}.$$

A high Kappa score indicated that two annotation terms share many common proteins.
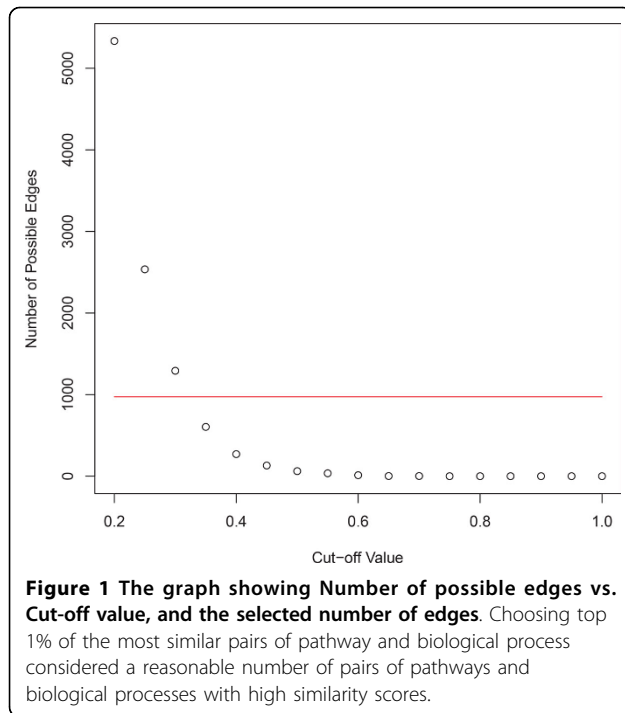
Construction of undirected GOBP graph

An undirected GOBP graph *GraphGOBPenriched* was constructed to describe the relationships among $N_{enrichedGOBP}$ enriched GOBP terms, i.e., *GraphGOBPenriched* = (V_{GOBP},

E_{GOBP}), $|V_{GOBP}| = N_{enrichedGOBP}$, and E_{GOBP} defines the set of edges in the graph. The relationships between GOBP terms, represented by edges connecting them, were evaluated based on the ancestor/offspring relationships in the complete directed acyclic graph of all GOBP terms from the Gene Ontology Consortium. We mapped $N_{enrichedGOBP}$ enriched GOBP terms to the corresponding vertices of the complete directed acyclic graph of all GOBP terms from the Gene Ontology Consortium using the package "GO.db" from Bioconductor [14]. Let *GraphGOBPComplete* = ($V_{completeGOBP}$, $E_{completeGOBP}$) be the complete directed acyclic graph of all GOBP terms. Then, V_{GOBP} is mapped to $V_{completeGOBP}$ ($V_{GOBP} \subset V'$ and $V' \subset V_{completeGOBP}$). Two GOBP terms would be connected if there existed a link between this pair of vertices in the complete graph of GOBP. All networks and graphs in this study were constructed and analyzed with the assistance of the package 'igraph' in R [15].

Construction of undirected Boolean bipartite pathway and GOBP graph

The relationships between pathways and GOBP terms were represented as an undirected graph where edges between pathways and GOBP terms were evaluated based on Kappa statistics. We computed the Kappa similarity matrix of size $N_{totalGOBP}$ x $N_{totalPathway}$, where $N_{totalPathway}$ is the total number of pathways including Biocarta, KEGG and Reactome pathways. Each row of the similarity matrix represents a GOBP term, and each column represents a pathway. Top 1% of the most similar pairs of pathway and biological process were selected and connected based on the Kappa similarity scores. Figure 1 showed that choosing the top 1% of the most similar pairs allowed the selection of a reasonable number of edges with high similarity scores (the average of Kappa scores was 0.025, and the chose cut-off value was 0.27). The set of pairs of pathway and GOBP terms satisfying such condition as was denoted as $E_{PathwayGOBP}$. We then established the pathway and GOBP graph as an undirected bipartite graph *BipartiteGraphPathwayGOBP* = {$V_{Pathway}$, V_{GOBP}, $E_{PathwayGOBP}$} where $V_{Pathway}$ is the set of pathways and V_{GOBP} is the set of GOBP terms included in $E_{PathwayGOBP}$ ($|V_{Pathway}| \leq N_{totalPathway}$ and $|V_B| \leq N_{totalGOBP}$). Thus, the graph *BipartiteGraphPathwayGOBP* would consist of pathways that could be well represented by GOBP terms.

We further introduced Boolean rules to *BipartiteGraphPathwayGOBP* to represent pathways as Boolean functions of biological processes, assuming that connected biological processes have direct impacts on the pathways. Since a pathway contains dynamics and dependencies among participating molecules, which are annotated by biological processes, we assume that different combinations of biological process states can affect the

Figure 1 The graph showing Number of possible edges vs. Cut-off value, and the selected number of edges. Choosing top 1% of the most similar pairs of pathway and biological process considered a reasonable number of pairs of pathways and biological processes with high similarity scores.

state of the pathway, which is either 'active' (binary state 1) or 'inactive' (binary state 0). For every pathway $V_{Pathwayi}$ in the graph *BipartiteGraphPathwayGOBP*, let $V_{PathwayGOBPi}$ be the set of GOBP terms connected to that pathway and $V_{GOBP} = \cup V_{PathwayGOBPi}$, we performed Boolean mapping such that the pathway $V_{Pathwayi}$ could be described as a Boolean algebra functions of its connected GOBP terms, $V_{Pathwayi} = f(V_{PathwayGOBPi})$.

The Boolean rules were derived from the relationships between GOBP terms connected to the pathway. If two GOBP terms were connected, then the Boolean relationship between these GOBP terms would be "OR." Such assumption arose from the fact two connected GOBP terms would share a significant amount of protein; thus, if a biological process was active, then its connected process must be simultaneously active as well. The relationship between two unconnected GOBP terms would be "AND." For example, considering a small network with 3 GOBP terms, $GOBP_1$, $GOBP_2$ and $GOBP_3$, and a pathway P, where $GOBP_1$ and $GOBP_2$ were connected, $GOPB_3$ was not connected with $GOBP_1$ and $GOBP_2$, and all GOBP terms were connected to pathway P. Then, the Boolean function for P could be written as, $V_P = (V_{GOBP1} \cup V_{GOBP2}) \cap V_{GOBP3}$.

The functional biological pathway-process network and the extracted MI network

We combined the GOBP graph *GraphGOBPenriched* from section 2.4 and the bipartite graph *BipartiteGraphPathwayGOBP* from section 2.5 to have a complete

functional biological pathway-process network, where there were connections among GOBPs, and pathways communicated with each other through biological processes. As the complete network had many vertices and edges, we presented the MI pathway, *h_amiPathway*, from Biocarta, to illustrate the result. We retained important GOBP terms which were crossed by the shortest paths among other pathways to the MI pathway. Shortest paths were calculated using the unweighted breadth-first search method. The extracted network allowed us to identify how the MI pathway could lead to other pathways and vice versa, initiating cardiac remodeling post-MI.

Results
Undirected GOBP graph
Using DAVID Functional Annotation Tool, we obtained 993 enriched GOBP terms from the list of 613 MI-specific proteins. From the ancestor/offspring relationships, the graph *GraphGOBPenriched* was constructed, resulting in a network of 993 vertices and 4284 edges. *GraphGOBPenriched* had 16 connected sub-graphs having more than 1 vertex and 46 isolated vertices. The largest connected sub-graph consisted of 885 vertices and 4199 edges.

It is interesting to note that GOBP terms with the highest degree, measuring the number of direct links incident on a vertex in a graph, were related to phosphorylation, phosphate, phosphorus, and kinase activity (Table 1). Since phosphorus and phosphate metabolic processes have the highest connections, this could mean that the chemical reactions and pathways involving intracellular signaling might initiate the cascade of events post-MI. In fact, serum phosphorus has been shown to serve as a sensitive indicator of MI and is linked to all-cause mortality and heart failure in patients after MI [16,17]. Hypophosphatemia in MI is associated with a greater degree of dysfunction of the left ventricle (LV), resulting in increased 30 days mortality [18]. In patients with MI, plasma sphingosine-1-phosphate concentration is reduced, leading to decreases protective action on cardiomyocyte viability [19].

In addition, biological processes involved with phosphorylation accounted for 4 GOBP terms while there were 5 kinase-activity-related GOBPs in Table 1. Phosphorylation is a major post-translational modification to regulate protein function. In a phosphorylation process, a protein kinase modifies target proteins, or substrates, by chemically adding phosphate groups to them. This result corresponded well with our previous work which identified Kinase Pathways as one of the major groups of pathways significantly enriched following MI [11].

Table 1. Top 20 GO Biological Processes ranked by degree measurements

GOBP ID	Name	Degree
GO:0006793	phosphorus metabolic process*	55
GO:0006796	phosphate metabolic process*	52
GO:0006955	immune response	47
GO:0010033	response to organic substance	47
GO:0016310	phosphorylation*	45
GO:0048584	positive regulation of response to stimulus	45
GO:0051174	regulation of phosphorus metabolic process*	44
GO:0019220	regulation of phosphate metabolic process*	43
GO:0043507	positive regulation of JUN kinase activity*	43
GO:0042325	regulation of phosphorylation*	39
GO:0043406	positive regulation of MAP kinase activity*	38
GO:0000187	activation of MAPK activity*	38
GO:0032268	regulation of cellular protein metabolic process	37
GO:0031659	positive regulation of cyclin-dependent protein kinase activity during G1/S*	37
GO:0006468	protein amino acid phosphorylation*	36
GO:0010604	positive regulation of macromolecule metabolic process	34
GO:0001932	regulation of protein amino acid phosphorylation*	34
GO:0001775	cell activation	34
GO:0045860	positive regulation of protein kinase activity*	33
GO:0006952	defense response	33

*GOBPs related with phosphorus, phosphate, phosphorylation, and kinase activity.
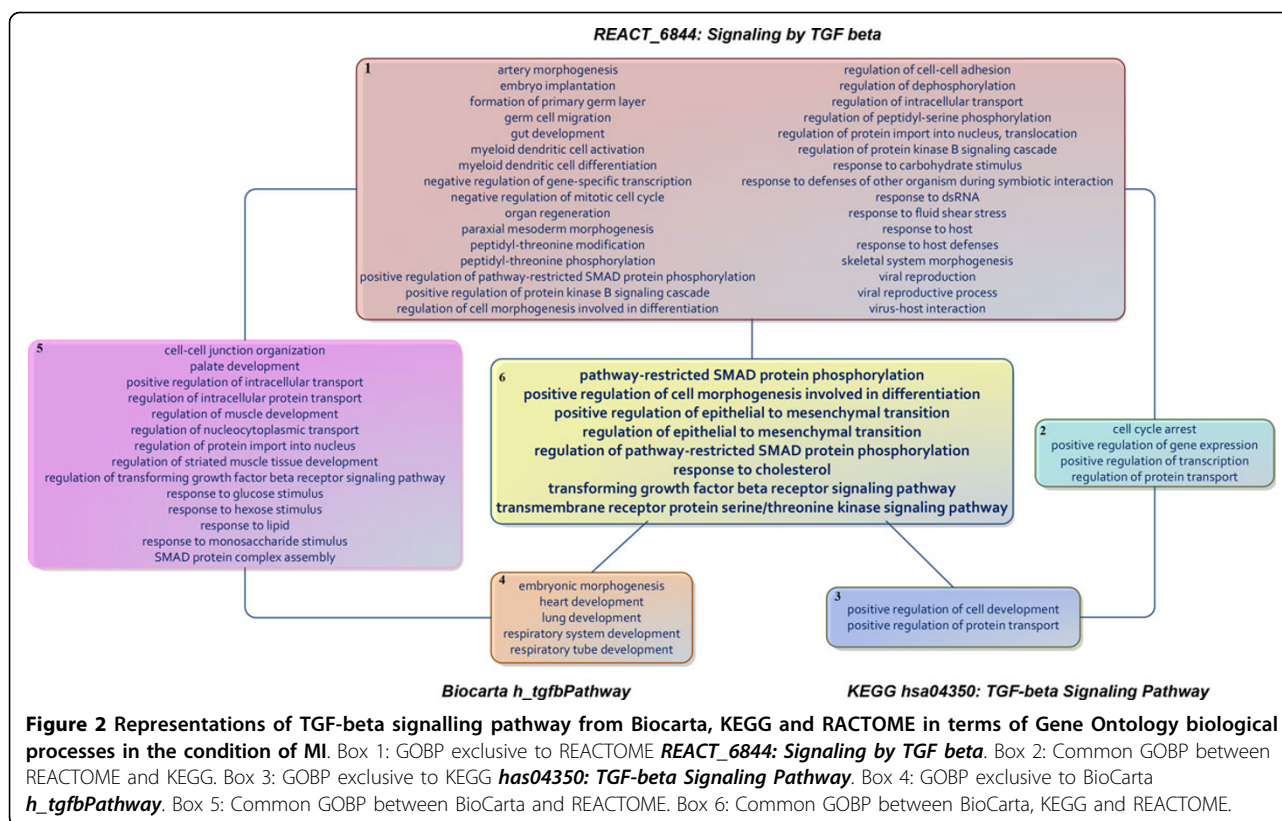
Network of biological pathways and GOBP showed similarities and differences among pathways in regard to GOBP annotation

At selected parameters, we retrieved 98 pathways, including 37 KEGG, 13 Reactome, and 48 Biocarta pathways using DAVID Functional Annotation Tool. Analysing statistical measures of inter-annotator agreement between 98 pathways and 993 GOBP terms, we established a graph *BipartiteGraphPathwayGOBP* with 544 vertices, containing 96 pathways 448 associated GOBPs, and 973 edges. These edges represented the most significantly enriched pairs of pathways and GOBP in the context of MI. This graph consisted of 8 sub-graphs, with the largest connected component having 76 pathways and 396 GOBP terms.

Earlier, we mentioned the TGF-beta signaling pathway and how it was defined differently among the KEGG, Reactome, and Biocarta pathway databases. We further examined the associated GOBP terms to compare these 3 pathways (Figure 2). The variations were due to different literature being used to construct the pathways: *REACT_6844: Signaling by TGF beta* were involved with 56 GOBP terms, *hsa04350: TGF-beta Signaling Pathway* was associated with 14 GOBP terms, and the *h_tgfbPathway* was linked to 27 GOBP terms. Nonetheless, the common biological processes among these pathways included phosphorylation of SMAD proteins, serine/threonine kinase signaling pathway, epithelial-mesenchymal transition, and response to cholesterol and cell morphogenesis

involved in differentiation (Figure 2: Box 6). It can be seen that the *REACT_6844* provided a more complete description of TGFβ signaling pathway (Figure 2: Box 1-2&5), *hsa04350* mainly focused on protein transport, transcription, gene expression and cell development (Figure 2: Box 2-3), whereas *h_tgfbPathway* emphasized organ development (Figure 2: Box 4-5). As a result, we can understand the different characteristics assigned for each pathway under the different circumstances. Individually, TGF-beta signaling pathways from KEGG, Biocarta, and Reactome annotated 21, 12, and 7 proteins, respectively, from the initial 613 MI-specific proteins. Thus, by incorporating the signaling pathways from different sources, we updated the knowledge of TGF-beta signaling pathways with more biological processes, and identified additional proteins participating in the pathway. Using this approach, the total number of proteins annotated with TGF-beta signaling pathways, by combining proteins from KEGG, Biocarta and Reactome, was expanded to 25 proteins.

Additionally, we investigated how this system acts using three other cardiovascular disease processes, namely *hsa05412: Arrhythmogenic Right Ventricular Cardiomyopathy* (ARVC), *hsa05410: Hypertrophic Cardiomyopathy* (HCM), and *hsa05414: Dilated Cardiomyopathy* (DCM). These analyses provide additional examples to demonstrate how representing pathways in terms of biological processes helped us to quickly understand the characteristics of such conditions under specific circumstances (Figure 3). ARVC is an inherited disease that results in fat

Figure 2 Representations of TGF-beta signalling pathway from Biocarta, KEGG and RACTOME in terms of Gene Ontology biological processes in the condition of MI. Box 1: GOBP exclusive to REACTOME *REACT_6844: Signaling by TGF beta*. Box 2: Common GOBP between REACTOME and KEGG. Box 3: GOBP exclusive to KEGG *has04350: TGF-beta Signaling Pathway*. Box 4: GOBP exclusive to BioCarta *h_tgfbPathway*. Box 5: Common GOBP between BioCarta and REACTOME. Box 6: Common GOBP between BioCarta, KEGG and REACTOME.

and fibrous tissues replacing the heart muscle of the right ventricle and subepicardial region of the left ventricle. With HCM, a portion of the myocardium is hypertrophied, forcing the heart to work harder to pump blood because of the thickened heart muscle. DCM is a condition in which the heart weakens and becomes dilated, resulting in inefficient blood pumping to other organs. All three aforementioned cardiomyopathy pathways involve integrin-mediated signaling pathway, cell-matrix adhesion, and cell-substrate adhesion. However, HCM and DCM are specifically related to leukocyte adhesion. It has been confirmed that human leukocyte antigens are associated with HCM and DCM [20-23].

We showed a visualization of a sub-graph consisting of 7 pathways and 34 GOBP terms that intersected with the MI response (Figure 4). Two pathways having the largest number of associated GOBP terms were *hsa04610: Complement And Coagulation Cascades* (characterized by 17 GOBP terms) and *h_fibrinolysisPathway* (characterized by 22 GOBP terms). The center of this sub-network is the MI pathway from Biocarta, *h_amiPathway*. Altogether, 3 pathways were represented by 32 out of 34 GOBP terms in this sub-network, and there were 8 common GOBP terms, including *coagulation, regulation of coagulation, negative regulation of coagulation, blood coagulation, regulation of blood coagulation, negative*

regulation of blood coagulation, homeostasis and *regulation of body fluid levels* (Table 2). As a result, we noticed that blood coagulation, coagulation, homeostasis and regulation of body fluid levels were the underlying processes in these pathways. Table 2 and Figure 4 also pointed out the differences among these pathways: *hsa04610* was associated with activation of proteins involved in acute inflammatory response and wound healing, whereas the fibrinolysis pathway was specifically involved with fibrinolysis, platelet activation, protein phosphorylation, collagen process and tissue regeneration.

Associations between the MI response and biological processes have been experimentally and clinically verified
In order to confirm the affiliated biological processes with the MI response mentioned in the previous section, we searched PubMed for experimental and clinical evidence. In the *BipartiteGraphPathwayGOBP*, the MI pathway, annotated with 11 proteins, was connected with 16 GOBP terms that were linked to 64 proteins, and they shared 10 common proteins. We further verified that among the 54 proteins exclusively annotated by GOBP terms, 11 proteins had been chosen as the seed proteins to construct the MI-specific protein network. We have previously shown that these seed proteins were associated with MI and confirmed by at least 2 citations [11].
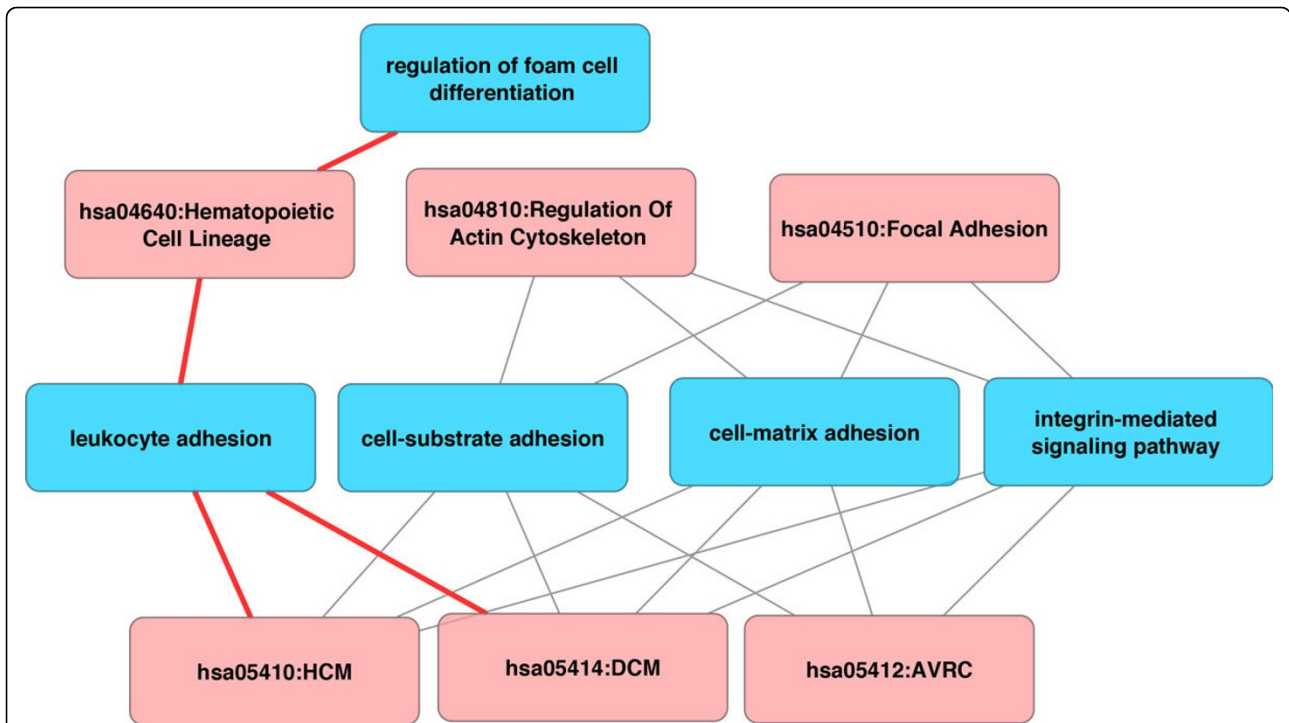
Figure 3 Sub-network of Cardiomyopathy. Pathways were represented in red while GOBPs were represented in blue. Pathways of Hypertrophic, Dilated and Arrhythmogenic Right Ventricular Cardiomyopathy were shown to be connected to biological processes including leukocyte adhesion, cell-substrated adhesion, and cell-matrix adhesion. Integrin-ECM interactions are required for cell adhesion.

To verify that the remaining 43 proteins of the expanded set of proteins for the MI pathway were related to MI, we searched for their official names and aliases on PubMed along with the keyword "myocardial infarction" for publications that confirmed the association between these proteins and MI (Table 3). There were 34 proteins firmly associated with MI by at least 2 publications. There were 3 proteins, namely CD44, SERPIND1 and HNF4A, directly associated with MI by one publication. There were 6 proteins, namely ANXA7, FBLN5, FGF7,
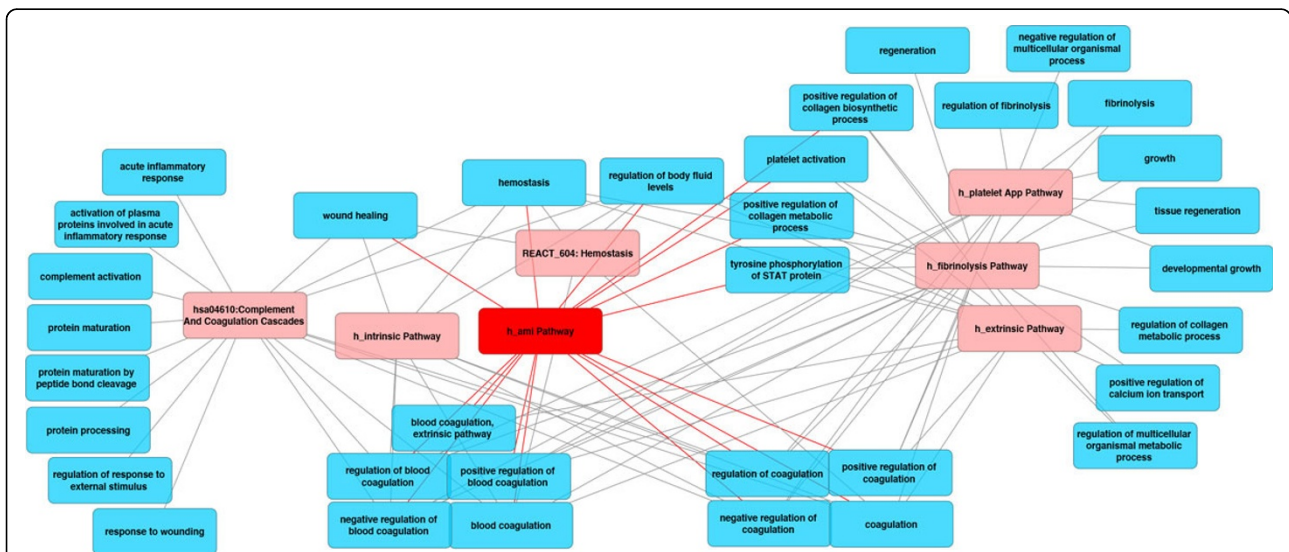

Figure 4 Sub-network of MI. Pathways were represented in while GOBPs were represented in blue. The major underlying processes for MI included coagulation, homeostasis, collagen metabolic/biosynthetic process, calcium ion transport, tissue regeneration, and wound healing.

Table 2. Pathways and GOBP in the MI functional pathway-process network

GOBP names	Pathways	F	C	A	E	P	I	H
activation of plasma proteins involved in acute inflammatory response			√					
acute inflammatory response			√					
blood coagulation		√	√	√	√		√	√
blood coagulation, extrinsic pathway				√	√			
coagulation		√	√	√	√		√	√
complement activation			√					
developmental growth		√				√		
fibrinolysis		√				√		
growth		√				√		
hemostasis		√	√	√	√		√	√
negative regulation of blood coagulation		√	√	√		√	√	
negative regulation of coagulation		√	√	√		√	√	
negative regulation of multicellular organismal process						√		
platelet activation		√		√	√			
positive regulation of blood coagulation		√		√	√	√		
positive regulation of calcium ion transport		√			√			
positive regulation of coagulation		√		√	√	√		
positive regulation of collagen biosynthetic process		√		√	√			
positive regulation of collagen metabolic process		√		√	√			
protein maturation			√					
protein maturation by peptide bond cleavage			√					
protein processing			√					
regeneration		√						
regulation of blood coagulation		√	√	√		√	√	
regulation of body fluid levels		√	√	√	√		√	√
regulation of coagulation		√	√	√		√	√	
regulation of collagen metabolic process		√			√			
regulation of fibrinolysis						√		
regulation of multicellular organismal metabolic process		√			√			
regulation of response to external stimulus			√					
response to wounding			√					
tissue regeneration		√				√		
tyrosine phosphorylation of STAT protein		√		√	√			
wound healing			√	√			√	√
Number of connected GOBPs		22	17	16	14	12	9	5

KLF6, FR2RL2 and GGCX indirectly linked to MI. Among 16 MI-associated GOBP terms, 11 biological processes were fully associated with the MI pathway as all of their member proteins were associated with MI and confirmed by at least 2 publications. The remaining 5 GOBP terms had 90% of the member proteins associated with the MI pathway, confirmed by at least 1 publication, and 80% or more of the member proteins were confirmed to be associated with MI by at least 2 publications. Therefore, we showed that the associations between MI pathway and biological processes in the *BipartiteGraphPathwayGOBP* have been experimentally and clinically verified. We also expanded the coverage of the original MI pathway by adding 54 new proteins. Further research will be needed to address the intermediate steps within the MI pathway and

develop more extensive description of the MI pathway that covers a longer time scale.

Phosphorylation of STAT protein, coagulation and regulation of collagen process are required to activate the MI pathway

We further explored the possibility of representing pathways as Boolean functions of biological processes. This idea originates from the fact that proteins within biological system typically act in concert. Biological processes are processed through protein-protein or molecular interactions, which usually have similar functions. The establishment of the bipartite graph of pathways and GOBP yielded Boolean functions to determine the state of pathways based on biological processes. We illustrated the MI
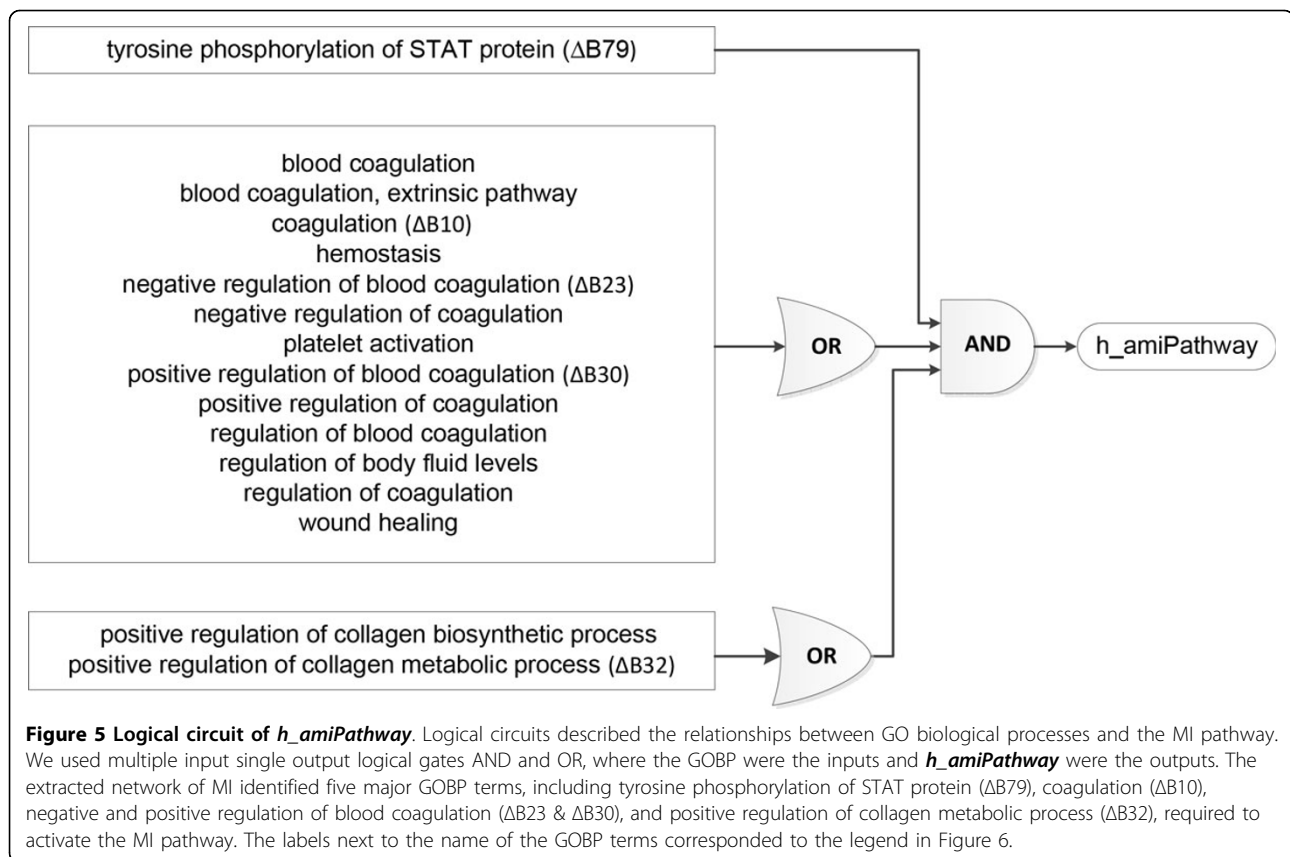
Table 3. Proteins of MI pathway-associated GOBP terms with cited publications

Proteins	Gene names	Official Names	Supporting Articles
A1AT_HUMAN	SERPINA1	Alpha-1-antitrypsin	[27,28]
ACVL1_HUMAN	ACVRL1	Serine/threonine-protein kinase receptor R3	[29,30]
ADA17_HUMAN	ADAM17	Disintegrin and metalloproteinase domain-containing protein 17	[31,32]
ANPRA_HUMAN	NPR1	Atrial natriuretic peptide receptor 1	[33,34]
APOA_HUMAN	LPA	Apolipoprotein(a)	[35,36]
CAV1_HUMAN	CAV1	Caveolin-1	[37,38]
CBPB2_HUMAN	CPB2	Carboxypeptidase B2	[39,40]
CD36_HUMAN	CD36	Platelet glycoprotein 4	[41,42]
EGLN_HUMAN	ENG	Endoglin	[29,30]
F13A_HUMAN	F13A1	Coagulation factor XIII A chain	[43,44]
FA11_HUMAN	F11	Coagulation factor XI	[45,46]
FA5_HUMAN	F5	Coagulation factor V	[47,48]
FA8_HUMAN	F8	Coagulation factor VIII	[43,45]
FA9_HUMAN	F9	Coagulation factor IX	[45,46]
FIBG_HUMAN	FGG	Fibrinogen gamma chain	[49,50]
FINC_HUMAN	FN1	Fibronectin	[51,52]
GPV_HUMAN	GP5	Platelet glycoprotein V	[53,54]
HIF1A_HUMAN	HIF1A	Hypoxia-inducible factor 1-alpha	[55,56]
IC1_HUMAN	SERPING1	Plasma protease C1 inhibitor	[57,58]
IFNG_HUMAN	IFNG	Interferon gamma	[59,60]
ITA5_HUMAN	ITGA5	Integrin alpha-5	[61,62]
KNG1_HUMAN	KNG1	Kininogen-1	[63,64]
LYOX_HUMAN	LOX	Protein-lysine 6-oxidase	[65,66]
PAR2_HUMAN	F2RL1	Proteinase-activated receptor 2	[67,68]
PAR4_HUMAN	F2RL3	Proteinase-activated receptor 4	[67,69]
PGFRA_HUMAN	PDGFRA	Platelet-derived growth factor receptor alpha	[70,71]
PLF4_HUMAN	PF4	Platelet factor 4	[72,73]
PROZ_HUMAN	PROZ	Vitamin K-dependent protein Z	[74,75]
SMAD3_HUMAN	SMAD3	Mothers against decapentaplegic homolog 3	[30,76]
TGFB2_HUMAN	TGFB2	Transforming growth factor beta-2	[77,78]
TGFR2_HUMAN	TGFBR2	TGF-beta receptor type-2	[79,80]
TRBM_HUMAN	THBD	Thrombomodulin	[81,82]
TSP1_HUMAN	THBS1	Thrombospondin-1	[83,84]
UROK_HUMAN	PLAU	Urokinase-type plasminogen activator	[85,86]
CD44_HUMAN	CD44	CD44 antigen	[87]
HEP2_HUMAN	SERPIND1	Heparin cofactor 2	[88]
HNF4A_HUMAN	HNF4A	Hepatocyte nuclear factor 4-alpha	[89]
ANXA7_HUMAN	ANXA7	Annexin A7	[90]
FBLN5_HUMAN	FBLN5	Fibulin 5	[91]
FGF7_HUMAN	FGF7	Fibroblast growth factor 7	[92]
KLF6_HUMAN	KLF6	Krueppel-like factor 6	[93]
PAR3_HUMAN	F2RL2	Proteinase-activated receptor 3	[94]
VKGC_HUMAN	GGCX	Vitamin K-dependent gamma-carboxylase	[95]

Proteins with indirect association with MI were contained in shaded box.

pathway *h_amiPathway* as logic circuits with multiple input single output logic gates (Figure 5). The MI pathway requires tyrosine phosphorylation of STAT protein, either positive regulation of collagen biosynthetic process or metabolic process, and one or more of the GOBP terms in the large group for activation. We later extracted the MI network, and identified the five major GOBP terms that contributed to the activation of *h_amiPathway* (Figure 6; see Additional file 1 for names of all pathways and GOBPs in the MI network). Tyrosine phosphorylation of STAT protein, negative and positive regulation of blood coagulation, coagulation and positive regulation of collagen

Figure 5 Logical circuit of *h_amiPathway*. Logical circuits described the relationships between GO biological processes and the MI pathway. We used multiple input single output logical gates AND and OR, where the GOBP were the inputs and ***h_amiPathway*** were the outputs. The extracted network of MI identified five major GOBP terms, including tyrosine phosphorylation of STAT protein (ΔB79), coagulation (ΔB10), negative and positive regulation of blood coagulation (ΔB23 & ΔB30), and positive regulation of collagen metabolic process (ΔB32), required to activate the MI pathway. The labels next to the name of the GOBP terms corresponded to the legend in Figure 6.

metabolic process are required to activate the MI pathway. By displaying pathways as logic circuits, we could observe the involvement of multiple functional groups, thus providing an intuitive way to understand associated pathways.

Critical routes of the extracted MI network

The complete network of pathways and GOBP contains a huge amount of information although it could be overwhelming. We extracted the MI network and only retained the backbone to explore additional features that might not have been covered. Figure 6 showed the routes from MI-related pathways, represented as light red circles, to the *h_amiPathway*, whose color was in red, through biological processes in as blue circles. The graph was undirected, meaning some routes could be bidirectional. A small branch of the network inside the blue rectangle was zoomed out for illustration purpose. The complete list of pathways and GOBP can be viewed in Supplemental Table 1. We observed that all 7 pathways in those 6 branches needed to pass through coagulation to be connected to *h_amiPathway*. We found the cell cycle pathway, *hsa04110:CellCycle*, particularly interesting since the pathway was linked to *h_fibrinolysisPathway*, through *cell growth*. Heissig et. al (2007) showed that by deleting plasminogen, a classical fibrinolytic

factor that controls hematopoietic stress response, in mice, hematopoietic stem cells were prevented from entering the cell cycle and undergoing multilineage differentiation after myelosuppression, leading to the death of the mice [24]. In other words, the plasminogen fibrinolytic pathway is crucial for hematopoietic regeneration. In another study, Heidt et al. (2014) showed that hematopoetic stem cells in the bone marrow could be activated by chronic stress, and further differentiated into increasing number of leukocytes. These leukocytes travel into the blood circulation and participate in the development of cardiovascular diseases [25]. Incidentally, fibrinolytic therapies have been used to enhance restoration of myocardial flow in the epcicardial infarct-related coronary artery [26]. Thus, it will be interesting to investigate the role of fibrinolysis and the increasing number of leukocytes in the cardiac remodelling post-MI and heart failures.

Discussion

In this study, we established a network by integrating GO biological processes and pathways from BioCarta, KEGG, and REACTOME enriched for MI-specific proteins using statistical measures and hierarchical structures. We examined the similarities between pathways and biological
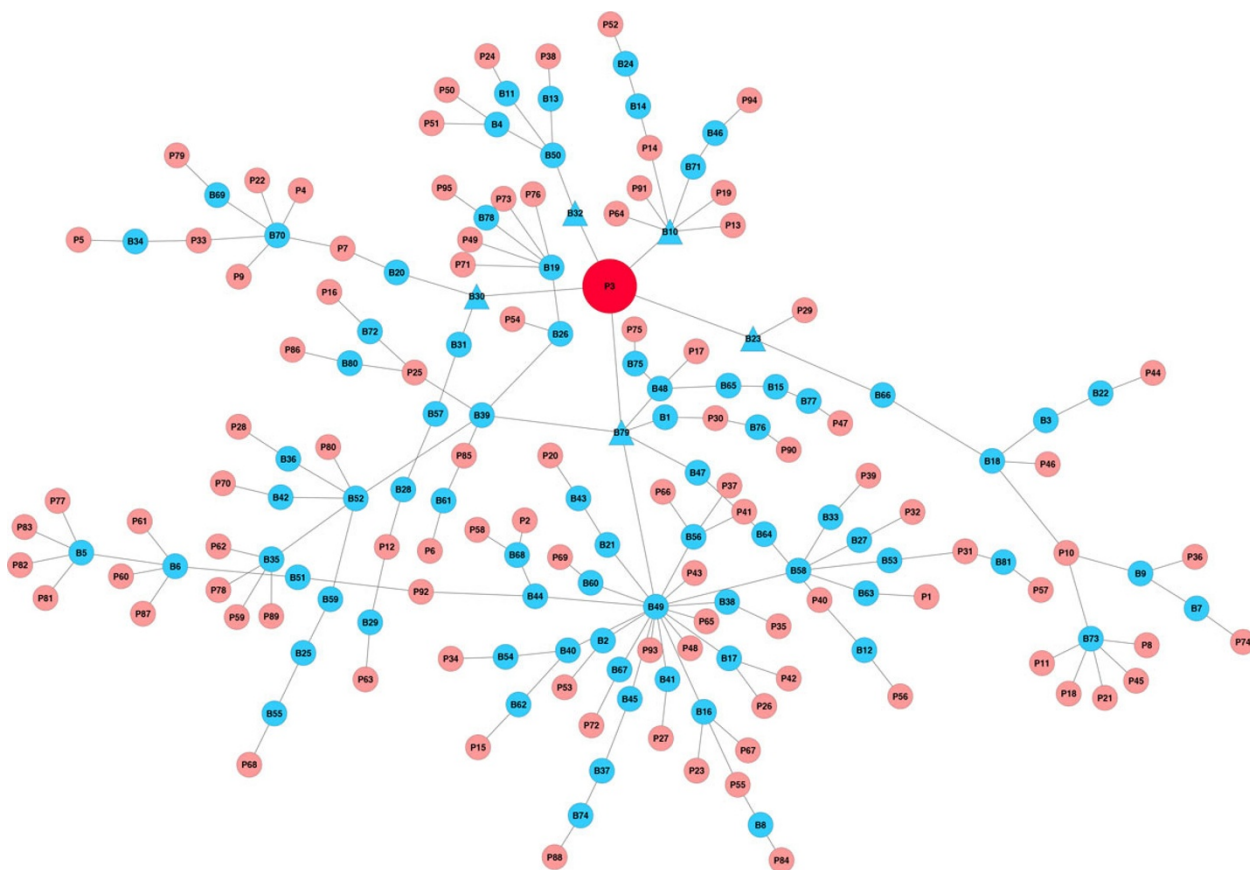
Figure 6 The extracted MI network. The acute MI pathway was colored in red while other pathways were colored in light red. Biological processes were represented in blue circles. GOBPs having direct impact on h_amiPathways were represented as blue triangles. A small branch of the network inside the blue rectangle involving coagulation was zoomed out for demonstration. Below are legends for selected pathways and processes (for the complete list of pathways and processes, see Supplemental Table 1). P3: h_amiPathway. P40: h_tgfbPathway. P58: hsa04350:TGF-betaSignalingPathway. P92: REACT_6844:Signaling by TGF beta. B10: coagulation. B30: positive regulation of blood coagulation. B32: positive regulation of collagen metabolic process. B44: positive regulation of protein kinase B signaling cascade. B58: regulation of kinase activity. B49: protein kinase cascade. B79: tyrosine phosphorylation of STAT protein.

processes, and derived Boolean models of pathways in terms of standardized vocabulary with GOBP terms. This network can be used to explore critical routes that connect pathways and biological processes to the development of diseases or conditions. To demonstrate a functional interaction network, we started from the proteins in an MI-specific protein-protein interaction network we had previously constructed, acquired the enriched GO biological processes and pathways, constructed the GOBP graph and the functional pathway-process network, and determined the logical circuitry representing the involvement of GOBPs in pathways. The approach could be used with any set of genes or proteins, specific to any conditions or diseases, to develop additional features and visualizations.

This study presented three important results. First, we established a MI-specific functional biological pathway-process network, with demonstrated sub-networks shown in Figures 2 and 3. We standardized pathway descriptions

by their connected GOBP terms, making it easier to compare differences and similarities between pathways, especially those with similar descriptions from different databases. We provided an example in section 3.2 with TGF-beta signaling pathways and pointed out the common and exclusive biological properties from BioCarta, KEGG and REACTOME. Second, we derived the relationships between GOBP terms based on the hierarchical structure defined in the GO Consortium and organized these terms into functional groups that could contribute differently to the pathways. For each pathway, GOBP terms that belonged to different functional groups should act simultaneously to activate the pathway, whereas only one process in a functional group was needed initiate the function. We used multiple input single output logical gates AND and OR, where the GOBP were the inputs and pathways were the outputs. We built two logic circuits corresponding to the MI and fibrinolysis pathways. It was

shown that tyrosine phosphorylation of STAT protein, coagulation and regulation of collagen process were required to activate the MI pathway. We also provided experimental and clinical evidence for the association between the MI pathway and biological processes. Third, we illustrated a centralized version of the complete network of pathways and GOBP, providing insights of critical routes from and to the main pathway, *h_amiPathway*. Because MI was the major theme of this study, this extracted network allowed us to quickly visualize the connection between pathways before and after MI and their involvement in the changes in the post-MI myocardium.

Our results illustrated that using the functional biological pathway-process network is a promising method to identify biological properties of pathways under specific conditions. Pathways having similar descriptions encompassed both similar and diverse biological processes, indicating variation in their ability to share similar functional characteristics. The coverages of biological pathways can be increased with the incorporation of more biological processes and protein members, promoting more comprehensive pathways. As we discover and understand more about genes and proteins, the network helps to expand the participating genes or proteins in the pathways through the introduction of related genes in the GOBP. Pathways will be more comprehensive, leading to better knowledge of diseases. However, functional groups of GOBP terms based on hierarchical structures might need to be further evaluated for coherence. Moreover, GOBP functional groups might not have the same amount of contribution to the corresponding pathways; probabilistic Boolean models would allow more robustness in the face of uncertainty. In conclusion, we report here the establishment of the network of pathways and biological processes that can be used as a foundation to identify biological properties of pathways, providing interaction and visualization of biological systems at pathway level.

Additional material

Additional File 1: Pathways and GOBPs of MI network. This file contains names of pathways and GOBPs in the extracted network of MI with labels as displayed in Figure 6. Pathways were ordered alphabetically with prefix "P". GOBPs were ordered alphabetically with prefix "B".

List of abbreviations used

AVRC: Arrhythmogenic Right Ventricular Cardiomyopathy
DCM: Dilated Cardiomyopathy
GO: Gene Ontology
GOBP: Gene Ontology Biological Process
HCM: Hypertrophic Cardiomyopathy
KEGG: Kyoto Encyclopaedia of Genes and Genomes
MI: Myocardial Infarction
MIPIN: Myocardial Infarction-Specific Protein-Protein Interaction Network
PANTHER: Protein Analysis Through Evolutionary Relationships
STAT: Signal Transducer and Activator of Transcription
TGF-beta: Transforming Growth Factor Beta

Authors' details
[1]Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, TX 78249, USA. [2]San Antonio Cardiovascular Proteomics Center, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA. [3]Mississippi Center for Heart Research, University of Mississippi Medical Center, Jackson, MS 39216, USA. [4]Research Service, G.V. (Sonny) Montgomery Veterans Affairs Medical Center, Jackson, MS 39216, USA.

Published: 11 June 2015

References
1. Kanehisa M, Goto S: KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000, **28(1)**:27-30.
2. Nishimura D: BioCarta. *Biotech Software & Internet Report* 2001, **2(3)**:117-120.
3. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, Caudy M, Garapati P, Gillespie M, Kamdar MR, *et al*: The Reactome pathway knowledgebase. *Nucleic Acids Res* 2014, **42(Database)**:D472-477.
4. Thomas PD, Campbell MJ, Kejariwal A, Mi H, Karlak B, Daverman R, Diemer K, Muruganujan A, Narechania A: PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res* 2003, **13(9)**:2129-2141.
5. Caspi R, Altman T, Billington R, Dreher K, Foerster H, Fulcher CA, Holland TA, Keseler IM, Kothari A, Kubo A, *et al*: The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res* 2014, **42(Database)**:D459-471.
6. Eichner J, Rosenbaum L, Wrzodek C, Haring HU, Zell A, Lehmann R: Integrated enrichment analysis and pathway-centered visualization of metabolomics, proteomics, transcriptomics, and genomics data by using the InCroMAP software. *J Chromatogr B Analyt Technol Biomed Life Sci* 2014, **966**:77-82.
7. Gruden K, Hren M, Herman A, Blejec A, Albrecht T, Selbig J, Bauer C, Schuchardt J, Or-Guil M, Zupancic K, *et al*: A "crossomics" study analysing variability of different components in peripheral blood of healthy caucasoid individuals. *PLoS One* 2012, **7(1)**:e28761.
8. Amiour N, Imbaud S, Clement G, Agier N, Zivy M, Valot B, Balliau T, Armengaud P, Quillere I, Canas R, *et al*: The use of metabolomics integrated with transcriptomic and proteomic studies for identifying key steps involved in the control of nitrogen metabolism in crops such as maize. *Journal of Experimental Botany* 2012, **63(14)**:5017-5033.

9. Babur O, Dogrusoz U, Cakir M, Aksoy BA, Schultz N, Sander C, Demir E: Integrating biological pathways and genomic profiles with ChiBE 2. *BMC Genomics* 2014, **15**:642.

10. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, *et al*: Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* 2000, **25**(1):25-29.

11. Nguyen NT, Zhang X, Wu C, Lange RA, Chilton RJ, Lindsey ML, Jin YF: Integrative computational and experimental approaches to establish a post-myocardial infarction knowledge map. *PLoS Comput Biol* 2014, **10**(3): e1003472.

12. Huang dW, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009, **4**(1):44-57.

13. Cohen J: A Coefficient of Agreement for Nominal Scales. *Educ Psychol Meas* 1960, **20**(1):37-46.

14. Carlson M: GO.db: A set of annotation maps describing the entire Gene Ontology 2013.

15. Csardi G NT: The igraph software package for complex network research. *InterJournal,* Complex Systems(1695); 2006.

16. Gould L, Reddy CV, Swamy CR, Oh KC, Kim SG: Decline of serum phosphorus in acute myocardial infarction. *Angiology* 1979, **30**(4):219-222.

17. Aronson D, Kapeliovich M, Hammerman H, Dragu R: The Relation between Serum Phosphorus Levels and Clinical Outcomes after Acute Myocardial Infarction. *PLoS One* 2013, **8**(3).

18. Vaidyanathan D, Venkatesan S, Ramadesikan VK: Serum phosphate in acute myocardial infarction. *Indian J Physiol Pharmacol* 2000, **44**(2):225-228.

19. Knapp M, Baranowski M, Czarnowski D, Lisowska A, Zabielski P, Gorski J, Musial W: Plasma sphingosine-1-phosphate concentration is reduced in patients with myocardial infarction. *Med Sci Monitor* 2009, **15**(9): Cr490-Cr493.

20. Darsee JR, Heymsfield SB, Nutter DO: Hypertrophic cardiomyopathy and human leukocyte antigen linkage: differentiation of two forms of hypertrophic cardiomyopathy. *N Engl J Med* 1979, **300**(16):877-882.

21. Nutter DO, Heymsfield SB, Glenn JF: Retraction. Darsee JR, Heymsfield SB, Nutter DO. Hypertrophic cardiomyopathy and human leukocyte antigen linkage: differentiation of two forms of hypertrophic cardiomyopathy. N Engl J Med 1979;300:877-82. *N Engl J Med* 1983, **308**(23):1400.

22. Lozano MD, Rubocki RJ, Wilson JE, McManus BM, Wisecarver JL: Human leukocyte antigen class II associations in patients with idiopathic dilated cardiomyopathy. Myocarditis Treatment Trial Investigators. *J Card Fail* 1997, **3**(2):97-103.

23. Li X, Luo R, Jiang R, Chen R, Hua W: Human leukocyte antigen-DQ beta 1 chain (DQB1) gene polymorphisms are associated with dilated cardiomyopathy: a systematic review and meta-analysis. *Heart Lung* 2012, **41**(4):360-367.

24. Heissig B, Lund LR, Akiyama H, Ohki M, Morita Y, Romer J, Nakauchi H, Okumura K, Ogawa H, Werb Z, *et al*: The plasminogen fibrinolytic pathway is required for hematopoietic regeneration. *Cell Stem Cell* 2007, **1**(6):658-670.

25. Heidt T, Sager HB, Courties G, Dutta P, Iwamoto Y, Zaltsman A, von Zur Muhlen C, Bode C, Fricchione GL, Denninger J, *et al*: Chronic variable stress activates hematopoietic stem cells. *Nat Med* 2014, **20**(7):754-758.

26. Armstrong PW, Collen D, Antman E: Fibrinolysis for acute myocardial infarction: the future is here and now. *Circulation* 2003, **107**(20):2533-2537.

27. Gilutz H, Siegel Y, Paran E, Cristal N, Quastel MR: Alpha 1-antitrypsin in acute myocardial infarction. *Br Heart J* 1983, **49**(1):26-29.

28. Toldo S, Seropian IM, Mezzaroma E, Van Tassell BW, Salloum FN, Lewis EC, Voelkel N, Dinarello CA, Abbate A: Alpha-1 antitrypsin inhibits caspase-1 and protects from acute myocardial ischemia-reperfusion injury. *J Mol Cell Cardiol* 2011, **51**(2):244-251.

29. Gonzalez-Nunez M, Munoz-Felix JM, Lopez-Novoa JM: The ALK-1/Smad1 pathway in cardiovascular physiopathology. A new target for therapy? *Biochim Biophys Acta* 2013, **1832**(10):1492-1510.

30. Tian F, Zhou AX, Smits AM, Larsson E, Goumans MJ, Heldin CH, Boren J, Akyurek LM: Endothelial cells are activated during hypoxia via endoglin/ ALK-1/SMAD1/5 signaling in vivo and in vitro. *Biochem Biophys Res Commun* 2010, **392**(3):283-288.

31. Shimoda Y, Satoh M, Nakamura M, Akatsu T, Hiramori K: Activated tumour necrosis factor-alpha shedding process is associated with in-hospital

32. Satoh M, Ishikawa Y, Itoh T, Minami Y, Takahashi Y, Nakamura M: The expression of TNF-alpha converting enzyme at the site of ruptured plaques in patients with acute myocardial infarction. *Eur J Clin Invest* 2008, **38**(2):97-105.

33. Nakayama T, Soma M, Saito S, Honye J, Sato M, Aoi N, Kosuge K, Haketa A, Kanmatsuse K, Kokubun S: Missense mutation of exon 3 in the type A human natriuretic peptide receptor gene is associated with myocardial infarction. *Medical science monitor : international medical journal of experimental and clinical research* 2003, **16**(12):CR505-510.

34. Nakanishi M, Saito Y, Kishimoto I, Harada M, Kuwahara K, Takahashi N, Kawakami R, Nakagawa Y, Tanimoto K, Yasuno S, *et al*: Role of natriuretic peptide receptor guanylyl cyclase-A in myocardial infarction evaluated using genetically engineered mice. *Hypertension* 2005, **46**(2):441-447.

35. McQueen MJ, Hawken S, Wang X, Ounpuu S, Sniderman A, Probstfield J, Steyn K, Sanderson JE, Hasani M, Volkova E, *et al*: Lipids, lipoproteins, and apolipoproteins as risk markers of myocardial infarction in 52 countries (the INTERHEART study): a case-control study. *Lancet* 2008, **372**(9634):224-233.

36. Kamstrup PR, Tybjaerg-Hansen A, Nordestgaard BG: Lipoprotein(a) and risk of myocardial infarction–genetic epidemiologic evidence of causality. *Scand J Clin Lab Invest* 2011, **71**(2):87-93.

37. Jasmin JF, Rengo G, Lymperopoulos A, Gupta R, Eaton GJ, Quann K, Gonzales DM, Mercier I, Koch WJ, Lisanti MP: Caveolin-1 deficiency exacerbates cardiac dysfunction and reduces survival in mice with myocardial infarction. *Am J Physiol Heart Circ Physiol* 2011, **300**(41): H1274-1281.

38. Shivshankar P, Halade GV, Calhoun C, Escobar GP, Mehr AJ, Jimenez F, Martinez C, Bhatnagar H, Mjaatvedt CH, Lindsey ML, *et al*: Caveolin-1 deletion exacerbates cardiac interstitial fibrosis by promoting M2 macrophage activation in mice after myocardial infarction. *J Mol Cell Cardiol* 2014, , **76C**: 84-93.

39. Juhan-Vague I, Morange PE, Aubert H, Henry M, Aillaud MF, Alessi MC, Samnegard A, Hawe E, Yudkin J, Margaglione M, *et al*: Plasma thrombin-activatable fibrinolysis inhibitor antigen concentration and genotype in relation to myocardial infarction in the north and south of Europe. *Arterioscler Thromb Vasc Biol* 2002, **22**(5):867-873.

40. Kraft P, Schwarz T, Meijers JC, Stoll G, Kleinschnitz C: Thrombin-activatable fibrinolysis inhibitor (TAFI) deficient mice are susceptible to intracerebral thrombosis and ischemic stroke. *PLoS One* 2010, **5**(7):e11658.

41. Pellikka M, Narhi L, Perola M, Penttila A, Karhunen PJ, Mikkelsson J: Platelet GPIbalpha, GPIV and vWF polymorphisms and fatal pre-hospital MI among middle-aged men. *Journal of thrombosis and thrombolysis* 2008, **26**(2):91-96.

42. Knowles JW, Wang H, Itakura H, Southwick A, Myers RM, Iribarren C, Fortmann SP, Go AS, Quertermous T, Hlatky MA: Association of polymorphisms in platelet and hemostasis system genes with acute myocardial infarction. *Am Heart J* 2007, **154**(6):1052-1058.

43. Siegerink B, Algra A, Rosendaal FR: Genetic variants of coagulation factor XIII and the risk of myocardial infarction in young women. *Br J Haematol* 2009, **146**(4):459-461.

44. Rallidis LS, Politou M, Komporozos C, Panagiotakos DB, Belessi CI, Travlou A, Lekakis J, Kremastinos DT: Factor XIII Val34Leu polymorphism and the risk of myocardial infarction under the age of 36 years. *Thromb Haemost* 2008, **99**(6):1085-1089.

45. Doggen CJ, Rosendaal FR, Meijers JC: Levels of intrinsic coagulation factors and the risk of myocardial infarction among men: Opposite and synergistic effects of factors XI and XII. *Blood* 2006, **108**(13):4045-4051.

46. Minnema MC, Peters RJ, de Winter R, Lubbers YP, Barzegar S, Bauer KA, Rosenberg RD, Hack CE, ten Cate H: Activation of clotting factors XI and IX in patients with acute myocardial infarction. *Arterioscler Thromb Vasc Biol* 2000, **20**(11):2489-2493.

47. Mannucci PM, Asselta R, Duga S, Guella I, Spreafico M, Lotta L, Merlini PA, Peyvandi F, Kathiresan S, Ardissino D: The association of factor V Leiden with myocardial infarction is replicated in 1880 patients with premature disease. *J Thromb Haemost* 2010, **8**(10):2116-2121.

48. Dowaidar M, Settin A: Risk of myocardial infarction related to factor V Leiden mutation: a meta-analysis. *Genet Test Mol Biomarkers* 2010, **14**(4):493-498.

complication in patients with acute myocardial infarction. *Clinical science* 2005, **108**(4):339-347.

49. Mannila MN, Lovely RS, Kazmierczak SC, Eriksson P, Samnegard A, Farrell DH, Hamsten A, Silveira A: Elevated plasma fibrinogen gamma' concentration is associated with myocardial infarction: effects of variation in fibrinogen genes and environmental factors. J Thromb Haemost 2007, 5(4):766-773.

50. Jacquemin B, Antoniades C, Nyberg F, Plana E, Muller M, Greven S, Salomaa V, Sunyer J, Bellander T, Chalamandaris AG, et al: Common genetic polymorphisms and haplotypes of fibrinogen alpha, beta, and gamma chains affect fibrinogen levels and the response to proinflammatory stimulation in myocardial infarction survivors: the AIRGENE study. J Am Coll Cardiol 2008, 52(11):941-952.

51. Ulrich MM, Janssen AM, Daemen MJ, Rappaport L, Samuel JL, Contard F, Smits JF, Cleutjens JP: Increased expression of fibronectin isoforms after myocardial infarction in rats. J Mol Cell Cardiol 1997, 29(9):2533-2543.

52. van Dijk A, Niessen HW, Ursem W, Twisk JW, Visser FC, van Milligen FJ: Accumulation of fibronectin in the heart after myocardial infarction: a putative stimulator of adhesion and proliferation of adipose-derived stem cells. Cell Tissue Res 2008, 332(2):289-298.

53. Aleil B, Mossard JM, Wiesel ML, Lanza F, Cazenave JP: Increased plasma levels of soluble platelet glycoprotein V in patients with acute myocardial infarction. J Thromb Haemost 2003, 1(8):1846-1847.

54. Morel O, Hugel B, Jesel L, Lanza F, Douchet MP, Zupan M, Chauvin M, Cazenave JP, Freyssinet JM, Toti F: Sustained elevated amounts of circulating procoagulant membrane microparticles and soluble GPV after acute myocardial infarction in diabetes mellitus. Thromb Haemost 2004, 91(2):345-353.

55. Tekin D, Dursun AD, Xi L: Hypoxia inducible factor 1 (HIF-1) and cardioprotection. Acta pharmacologica Sinica 2010, 31(9):1085-1094.

56. Kido M, Du L, Sullivan CC, Li X, Deutsch R, Jamieson SW, Thistlethwaite PA: Hypoxia-inducible factor 1-alpha reduces infarction and attenuates progression of cardiac dysfunction after myocardial infarction in the mouse. J Am Coll Cardiol 2005, 46(11):2116-2124.

57. Wouters D, Wagenaar-Bos I, van Ham M, Zeerleder S: C1 inhibitor: just a serine protease inhibitor? New and old considerations on therapeutic applications of C1 inhibitor. Expert opinion on biological therapy 2008, 8(8):1225-1240.

58. Rennie JA, Crawford GP, Ogston D: Changes in protease inhibitors after acute myocardial infarction. J Clin Pathol 1976, 29(7):639-641.

59. Szkodzinski J, Hudzik B, Osuch M, Romanowski W, Szygula-Jurkiewicz B, Polonski L, Zubelewicz-Szkodzinska B: Serum concentrations of interleukin-4 and interferon-gamma in relation to severe left ventricular dysfunction in patients with acute myocardial infarction undergoing percutaneous coronary intervention. Heart and vessels 2011, 26(4):399-407.

60. Yang ZQ, Xu YQ, Linden J, Kron IL, French BA: Reduced Myocardial Infarct Size in Interferon-Gamma Knock-out Mice Implicates CD4+T Cells in Reperfusion Injury. Circulation 2009, 120(18):S1165-S1166.

61. Nawata J, Ohno I, Isoyama S, Suzuki J, Miura S, Ikeda J, Shirato K: Differential expression of alpha 1, alpha 3 and alpha 5 integrin subunits in acute and chronic stages of myocardial infarction in rats. Cardiovascular Research 1999, 43(2):371-381.

62. Sahul Z, Dione DP, Dobrucki L, Kalinowski L, Brennan M, Mekkaoui C, Cavaliere P, Hawley C, Hu X, Haramis H, et al: Targeted alpha-v integrin imaging defines spatial and temporal changes in the angiogenic process post myocardial infarction. Circulation 2006, 114(18):499-499.

63. Siegerink B, Rosendaal FR, Algra A: High-molecular-weight kininogen and the risk of a myocardial infarction and ischemic stroke in young women: the RATIO case-control study. Journal of Thrombosis and Haemostasis 2012, 10(11):2409-2412.

64. Ito H, Hayashi I, Izumi T, Majima M: Bradykinin inhibits development of myocardial infarction through B2 receptor signalling by increment of regional blood flow around the ischaemic lesions in rats. Br J Pharmacol 2003, 138(1):225-233.

65. Xie Y, Chen J, Han P, Yang P, Hou J, Kang YJ: Immunohistochemical detection of differentially localized up-regulation of lysyl oxidase and down-regulation of matrix metalloproteinase-1 in rhesus monkey model of chronic myocardial infarction. Exp Biol Med (Maywood) 2012, 237(7):853-859.

66. Lerman RH, Apstein CS, Kagan HM, Osmers EL, Chichester CO, Vogel WM, Connelly CM, Steffee WP: Myocardial healing and repair after experimental infarction in the rabbit. Circ Res 1983, 53(3):378-388.

67. Antoniak S, Pawlinski R, Mackman N: Protease-activated receptors and myocardial infarction. IUBMB life 2011, 63(6):383-389.

68. Zhong B, Wang DH: Protease-activated receptor 2-mediated protection of myocardial ischemia-reperfusion injury: role of transient receptor potential vanilloid receptors. Am J Physiol Regul Integr Comp Physiol 2009, 297(6):R1681-1690.

69. Seqqat R, Rafiq K, Hanscom M, Kunapuli SP, Steinberg SF, Houser SR, Sabri A: Protease activated receptor-4 regulates post-infarction ventricular remodeling and cardiac function. Circulation 2007, 116(16):45-45.

70. Zymek P, Bujak M, Chatila K, Cieslak A, Thakker G, Entman ML, Frangogiannis NG: The role of platelet-derived growth factor signaling in healing myocardial infarcts. J Am Coll Cardiol 2006, 48(11):2315-2323.

71. Edelberg JM, Lee SH, Kaur M, Tang L, Feirt NM, McCabe S, Bramwell O, Wong SC, Hong MK: Platelet-derived growth factor-AB limits the extent of myocardial infarction in a rat model: feasibility of restoring impaired angiogenic capacity in the aging heart. Circulation 2002, 105(5):608-613.

72. Dymicka-Piekarska V, Kemona H, Mantur M, Stogowski A, Kemona-Chetnik I, Bychowski J: Platelet factor 4 as a marker of platelet activation in patients with acute myocardial infarction. Roczniki Akademii Medycznej w Bialymstoku 2000, 45:96-103.

73. Kuijpers PM, Hamulyak K, Strik JJ, Wellens HJ, Honig A: Beta-thromboglobulin and platelet factor 4 levels in post-myocardial infarction patients with major depression. Psychiatry research 2002, 109(2):207-210.

74. Fedi S, Sofi F, Brogi D, Tellini I, Cesari F, Sestini I, Gazzini A, Comeglio M, Abbate R, Gensini GF: Low protein Z plasma levels are independently associated with acute coronary syndromes. Thromb Haemost 2003, 90(6):1173-1178.

75. Le Cam-Duchez V, Soria C, Sollier CBD, Borg JY, Coudert M, Montalescot G, Esposito G, Drouet L, Collet JP: Rare genotypes of protein Z gene are a risk factor for premature myocardial infarction but not protein Z plasma level. Thromb Haemostasis 2009, 102(1):131-136.

76. Dobaczewski M, Bujak M, Li N, Gonzalez-Quesada C, Mendoza LH, Wang XF, Frangogiannis NG: Smad3 Signaling Critically Regulates Fibroblast Phenotype and Function in Healing Myocardial Infarction. Circulation Research 2010, 107(3):418-U176.

77. Deten A, Holzl A, Leicht M, Barth W, Zimmer HG: Changes in extracellular matrix and in transforming growth factor beta isoforms after coronary artery ligation in rats. Journal of Molecular and Cellular Cardiology 2001, 33(6):1191-1207.

78. Singla DK, Singla RD, Lamm S, Glass C: TGF-beta 2 treatment enhances cytoprotective factors released from embryonic stem cells and inhibits apoptosis in infarcted myocardium. Am J Physiol-Heart C 2011, 300(4):H1442-H1450.

79. Bujak M, Frangogiannis NG: The role of TGF-beta signaling in myocardial infarction and cardiac remodeling. Cardiovasc Res 2007, 74(2):184-195.

80. Okada H, Takemura G, Kosai KI, Li YW, Takahashi T, Esaki M, Yuge K, Miyata S, Maruyama R, Mikami A, et al: Postinfarction gene therapy against transforming growth factor-beta signal modulates infarct tissue dynamics and attenuates left ventricular remodeling and heart failure. Circulation 2005, 111(19):2430-2437.

81. Ireland H, Kunz G, Kyriakoulis K, Stubbs PJ, Lane DA: Thrombomodulin gene mutations associated with myocardial infarction. Circulation 1997, 96(1):15-18.

82. Chao TH, Li YH, Chen JH, Wu HL, Shi GY, Tsai WC, Chen PS, Liu PY: Relation of thrombomodulin gene polymorphisms to acute myocardial infarction in patients <= 50 years of age. American Journal of Cardiology 2004, 93(2):204-207.

83. Sezaki S, Hirohata S, Iwabu A, Nakamura K, Toeda K, Miyoshi T, Yamawaki H, Demircan K, Kusachi S, Shiratori Y, et al: Thrombospondin-1 is induced in rat myocardial infarction and its induction is accelerated by ischemia/reperfusion. Exp Biol Med 2005, 230(9):621-630.

84. Zwicker JI, Peyvandi F, Palla R, Lombardi R, Canciani MT, Cairo A, Ardissino D, Bernardinelli L, Bauer KA, Lawler J, et al: The thrombospondin-1 N700S polymorphism is associated with early myocardial infarction without altering von Willebrand factor multimer size. Blood 2006, 108(4):1280-1283.

85. Stavropoulou A, Philippou A, Halapas A, Sourla A, Pissimissis N, Koutsilieris M: uPA, uPAR and TGFbeta(1) expression during early and

late post myocardial infarction period in rat myocardium. *In vivo* 2010, **24**(5):647-652.

86. Minami E, Castellani C, Malchodi L, Deem J, Bertko K, Meznarich J, Dishmon M, Murry CE, Stempien-Otero A: **The role of macrophage-derived urokinase plasminogen activator in myocardial infarct repair: urokinase attenuates ventricular remodeling.** *J Mol Cell Cardiol* 2010, **49**(3):516-524.

87. Huebener P, Abou-Khamis T, Zymek P, Bujak M, Ying X, Chatila K, Haudek S, Thakker G, Frangogiannis NG: **CD44 is critically involved in infarct healing by regulating the inflammatory and fibrotic response.** *J Immunol* 2008, **180**(4):2625-2633.

88. Huang SS, Huang PH, Chen YH, Sung SH, Chiang KH, Chen JW, Lin SJ: **Plasma heparin cofactor II activity is an independent predictor of future cardiovascular events in patients after acute myocardial infarction.** *Coron Artery Dis* 2008, **19**(8):597-602.

89. Voight BF, Peloso GM, Orho-Melander M, Frikke-Schmidt R, Barbalic M, Jensen MK, Hindy G, Holm H, Ding EL, Johnson T, *et al*: **Plasma HDL cholesterol and risk of myocardial infarction: a mendelian randomisation study.** *Lancet* 2012, **380**(9841):572-580.

90. Camors E, Monceau V, Charlemagne D: **Annexins and Ca2+ handling in the heart.** *Cardiovasc Res* 2005, **65**(4):793-802.

91. Spencer JA, Hacker SL, Davis EC, Mecham RP, Knutsen RH, Li DY, Gerard RD, Richardson JA, Olson EN, Yanagisawa H: **Altered vascular remodeling in fibulin-5-deficient mice reveals a role of fibulin-5 in smooth muscle cell proliferation and migration.** *Proc Natl Acad Sci USA* 2005, **102**(8):2946-2951.

92. Galli D, Innocenzi A, Staszewsky L, Zanetta L, Sampaolesi M, Bai A, Martinoli E, Carlo E, Balconi G, Fiordaliso F, *et al*: **Mesoangioblasts, vessel-associated multipotent stem cells, repair the infarcted heart by multiple cellular mechanisms: a comparison with bone marrow progenitors, fibroblasts, and endothelial cells.** *Arterioscler Thromb Vasc Biol* 2005, **25**(4):692-697.

93. Sawaki D, Suzuki T, Aizawa K, Matsumura T, Munemasa Y, Ishida J, Fridman SL, Nagai R: **KLF6 Modulates Recruitment and Polarization of Inflammatory Cells Through Cardiomyocytes in Initiation of Cardiac Fibrosis.** *Circulation* 2011, **124**(21).

94. Cornelissen I, Palmer D, David T, Wilsbacher L, Concengco C, Conley P, Pandey A, Coughlin SR: **Roles and interactions among protease-activated receptors and P2ry12 in hemostasis and thrombosis.** *Proc Natl Acad Sci USA* 2010, **107**(43):18605-18610.

95. Brenner B, Sanchez-Vega B, Wu SM, Lanir N, Stafford DV, Solera J: **A missense mutation in gamma-glutamyl carboxylase gene causes combined deficiency of all vitamin K-dependent blood coagulation factors.** *Blood* 1998, **92**(12):4554-4559.