Published in final edited form as:

Arch Biochem Biophys. 2016 July 01; 601: 121–132. doi:10.1016/j.abb.2016.02.022.

Gene expression patterns in transgenic mouse models of hypertrophic cardiomyopathy caused by mutations in myosin regulatory light chain☆

Wenrui Huang[a,b], **Katarzyna Kazmierczak**[a], **Zhiqun Zhou**[a], **Vanessa Aguiar-Pulido**[b], **Giri Narasimhan**[b,c], and **Danuta Szczesna-Cordary**[a,*]

[a]Department of Molecular and Cellular Pharmacology, University of Miami Miller School of Medicine, Miami, FL 33136, USA

[b]Bioinformatics Research Group (BioRG), School of Computing and Information Sciences, Florida International University, Miami, FL 33199, USA

[c]Biomolecular Sciences Institute, Florida International University, Miami, FL 33199, USA

Abstract

Using microarray and bioinformatics, we examined the gene expression profiles in transgenic mouse hearts expressing mutations in the myosin regulatory light chain shown to cause hypertrophic cardiomyopathy (HCM). We focused on two malignant RLC-mutations, Arginine 58→Glutamine (R58Q) and Aspartic Acid 166 → Valine (D166V), and one benign, Lysine 104 → Glutamic Acid (K104E)-mutation. Datasets of differentially expressed genes for each of three mutants were compared to those observed in wild-type (WT) hearts. The changes in the mutant vs. WT samples were shown as fold-change (FC), with stringency FC ≥ 2. Based on the gene profiles, we have identified the major signaling pathways that underlie the R58Q-, D166V- and K104E-HCM phenotypes. The correlations between different genotypes were also studied using network-based algorithms. Genes with strong correlations were clustered into one group and the central gene networks were identified for each HCM mutant. The overall gene expression patterns in all mutants were distinct from the WT profiles. Both malignant mutations shared certain classes of genes that were up or downregulated, but most similarities were noted between D166V and K104E mice, with R58Q hearts showing a distinct gene expression pattern. Our data suggest that all three HCM mice lead to cardiomyopathy in a mutation-specific manner and thus develop HCM through diverse mechanisms.

Keywords

Pathological hypertrophy; Transgenic mice; Microarray; Differential gene expression; Network/centrality analysis

☆This article is part of a Special Issue entitled Myofilament Modulation of Cardiac Contraction, edited by Brandon J. Biesiadecki.

*Corresponding author. dszczesna@med.miami.edu (D. Szczesna-Cordary).

1. Introduction

Dominant mutations in the *MYL2* gene encoding for the human ventricular myosin regulatory light chain (RLC) are recognized to cause hypertrophic cardiomyopathy (HCM), a genetic and complex disorder known to be highly heterogeneous with respect to the course of the disease, age of onset, severity of symptoms and risk for sudden cardiac death (SCD) [1,2]. HCM is the most common cause of SCD among young individuals and competitive athletes [3]. The latest genetic studies on *MYL2* associated heart disease have revealed that mutations in *MYL2* are more frequent than previously reported (for review see Refs. [4,5]) and in just the past few years, new mutations were identified [6–8], with some particular *MYL2* variants demonstrating multiple occurrences in different ethnic populations [9,10]. Mutated patients have often no symptoms and live an uneventful life; however, some specific RLC mutations have been associated with heart failure and SCD [10–14]. Despite many studies on the mutated proteins and animal models of HCM, the molecular mechanisms underlying the specific disease phenotypes remain not fully understood. To elucidate the molecular basis of HCM-RLC phenotypes, we have applied the RNA microarray analysis used by many to identify the mechanisms of various genetic diseases [15]. The study was focused on two RLC mutations, Arginine 58 → Glutamine (R58Q) and Aspartic Acid 166 → Valine (D166V), both reported to cause malignant outcomes in humans [11–14], and one benign HCM-RLC mutation, Lysine 104 → Glutamic acid (K104E) [9,16]. Previously generated transgenic (Tg) mice expressing these HCM-RLC mutations in the heart were used and the results were compared with Tg wild-type (WT) mice, expressing the non-mutated human ventricular RLC [17–19]. In particular, we aimed to study the mechanisms that trigger development of malignant vs. benign RLC-HCM phenotypes in humans. In functional studies on Tg mice, the two malignant RLC mutations (R58Q and D166V) were shown to exert similar effects on force generation in skinned and intact papillary muscle fibers, i.e. both significantly increased the Ca^{2+}-sensitivity of contraction, diminished maximal tension and delayed muscle relaxation suggesting the possibility of diastolic dysfunction [18–21], which was in fact confirmed by echocardiography and invasive hemodynamics in Tg mice [22,23]. On the other hand, the K104E RLC mutation, did not significantly alter the Ca^{2+}-sensitivity of force but did cause changes in maximal force generation (reduced) and the ATPase activity (enhanced) [17]. Our histopathology data demonstrated HCM-related changes including myofibrillar disarray and fibrotic lesions in 5–9 month-old female and male K104E, D166V and R58Q animals, and these changes were significantly intensified in senescent (>13 month-old) mutant vs. WT mice [17–19,23,24]. Hypertrophy and significant systolic and/or diastolic abnormalities in D166V and R58Q mice, observed by echocardiography, Doppler and invasive hemodynamics were evident in 5–9 month-old and senescent mice [22,23]. However, the phenotype associated with the K104E mutation was mild in young (3-5 month-old) and intermediate aged (~8 month-old) animals but evident in mice >13 months of age compared to age matched Tg-WT mice [17]. It is worth mentioning that no phenotypic differences were noted between Tg-WT mice expressing the human isoform of ventricular RLC compared with non-transgenic (NTg) mice containing the mouse cardiac RLC [18,19,21,22].

The microarray and bioinformatics analyses of gene expression profiles were performed on the hearts of 5–8.5 month-old mice expressing R58Q, K104E and D166V HCM-RLC mutations (three hearts per group) and the results were compared to Tg-WT mice (Table 1). The data demonstrated that the overall gene expression patterns differed between all mutant samples and WT mice. Interestingly, both malignant mutations (R58Q and D166V) shared certain classes of genes that were up or downregulated, but more similarities were noted between D166V and K104E hearts, with R58Q showing a distinct gene expression profile. We then applied bioinformatics tools (based on network theory) on the gene co-expression network (GCN) to infer the most *central* (i.e., potentially the most influential) genes altered in these three disease phenotypes compared to WT hearts [25–29]. The most common one is building co-occurrence or co-expression networks. A gene co-expression network is a network, where each node corresponds to a gene, and a pair of nodes is connected by an edge if there is a significant co-expression relationship between them [30]. Central nodes in GCNs have been proposed as candidate driver genes [28,29,31]. A network-based approach to perform clustering and centrality analysis of differentially expressed genes in GCNs showed varying expression patterns in the three RLC mutant mice across different HCM phenotypes. We also identified central or the most influential genes that were altered in R58Q, D166V and K104E hearts.

2. Materials and methods

2.1. RNA preparation and microarray assay

All animal studies were conducted in accordance with institutional guidelines. The University of Miami has an Animal Welfare Assurance (A-3224-01, effective November 23, 2011) on file with the Office of Laboratory Animal Welfare (OLAW), National Institutes of Health.

Wild-type or HCM B6SJL mice (listed in Table 1) were euthanized, left ventricles were rapidly harvested and immediately submerged in 25 vol of room-temperature RNAlater RNA stabilization reagent (Qiagen) [32]. After overnight incubation at 4 °C, the samples were stored frozen at −20 °C until used. Total RNA was isolated from RNAlater stabilized tissues using an RNeasy Fibrous Tissue Mini Kit (Qiagen, Valencia, CA) after being homogenized in a Mixer-Mill MM301 (Retsch) according to the manufacturer's protocol. Total RNA samples were hybridized to GeneChip® Mouse Gene 2.0 ST Array (Affymetrix®) at the Center for Genome Technology at the University of Miami Miller School of Medicine. The raw data were stored in CEL format files.

2.2. Principle Component Analysis (PCA) for multiple comparisons of gene expression patterns across datasets

The Principle Component Analysis (PCA) is a common statistical analysis technique that aims to reduce the number of dimensions of the high-dimensional data by extracting key features and their contributions to the variations. Each principal component is a linear combination of the original variables that are usually sorted based on the percentage of variance of the original data that they represent. By performing a linear transformation based on the top three principal components, we effectively map the original data points to points

in 3D space, making it convenient to visualize the clustering and grouping of the high-dimensional original data. All 12 raw datasets (three WT, three R58Q, three D166V and three K104E) were imported into the Partek Genomics Suite and the 3D PCA plot was generated by selecting the "Principle Component Analysis" tag in the QA/QC section (Fig. 1). Each PC1, PC2 and PC3 of the plot is labeled with % of explained variance with each dot representing a transformed original data point. Clusters of points in this plot represent groups with similar gene expression patterns.

2.3. Bioinformatic analysis of gene expression data

Using bioinformatics analyses, we have examined the gene expression profiles in Tg mouse hearts expressing mutations in the myosin RLC shown by population studies to cause HCM. Raw data (CEL. files) were uploaded into Partek Genomics Suite for normalization and statistical analysis. The CEL files, containing the probe-level data from all arrays were imported into Partek® Genomics Suite™ software (Partek Inc., St. Louis, MO, USA). Robust multichip analysis normalization was applied to yield $\log_2$-transformed expression intensities. One-way analysis of variance (ANOVA) was performed on the genes across the four groups (R58Q, D166V, K104E and WT). Pairwise comparisons were performed between mutations providing fold-change (FC) values. The list of FC of all genes were exported into excel file for further processing. Since during the calculation process, the WT samples were used for normalization, the processed FC results only contain the data of three mutations: R58Q, D166V and K104E.

2.4. Scatter plot of gene expression and comparison between groups

The scatter plots (Fig. 2) were created using SigmaPlot 11.0 software with each dot in the figure representing the FC of one gene with respect to the WT compared under two different conditions. The complete list of 29,726 genes were compared in pairs: R58Q vs. D166V, R58Q vs. K104E and K104E vs. D166V. Each quadrant in the x-y coordinates was divided into two parts (a and b) and divided by a 45° line generating eight parts 1-a, 1-b, 2-a, 2-b, 3-a, 3-b, 4-a and 4-b, indicating an upregulation or downregulation of a gene and the amount of upregulated/downregulated genes in one mutant vs. second compared mutant. The final 3D scatter plot was also generated with the FC in all three R58Q, D166V and K104E mutants as compared to WT.

2.5. Gene ontology analysis

The differentially expressed gene lists with at least a 2-fold change (in either direction) vs. WT were generated and sorted for further analysis. The molecular function and biological processes for each differentially expressed gene were annotated and categorized using the "mygene" package from the statistical software suite R [33]. Pie charts (Fig. 3) of biological processes and molecular functions with at least two differentially expressed genes in each category were plotted as a percentage of the total number of altered genes.

2.6. Signaling pathway analyses

The functional pathway analyses were performed using the Partek® Pathway™ software. Gene set enrichment analysis (GSEA) was carried out for interpreting microarray data to

detect disrupted or influential pathways and genes based on their FC and a biological mechanism. The gene expression results from two comparisons (mutant vs. WT) were ranked by absolute FC (Table 2), and the statistical enrichment of curated pathways was determined from known gene sets. The results of the analysis generated a list of the most significantly enriched pathways ranked by enrichment p value with significance defined as $p \leq 0.05$ (Table 3). The number of differentially expressed genes and total number of genes were also listed for each pathway.

2.7. Applying network-based analysis for gene cluster and centrality analysis

We used normalized expression data from all 12 microarray datasets [(3 mutants+WT) × 3 per group] with rows corresponding to genes and 12 columns for the 12 datasets. The matrix was then separated into four sub matrices including only D166V and R58Q, only D166V and K104E, only R58Q and K104E, and one final group containing all three mutants. Intersections were performed among groups in such a way that only common differentially expressed genes (|FC| ≥ 2-fold) were maintained in the matrices. The following steps were then pursued: (a) calculate the correlation among genes; (b) build co-occurrence or co-expression networks; (c) cluster genes that show strong correlations with each other; and (d) identify “central” genes among each group. The details of the above steps can be found in Fernandez et al. [34] (co-occurrence networks) and Cickovski et al. [25] (centrality).

3. Results

To quantitatively assess the effects of HCM mutations on gene expression in three different RLC disease phenotypes, we used a microarray analysis to profile the mRNA expression patterns in transgenic mouse hearts bearing the R58Q, K104E and D166V mutations in the human ventricular RLC (GenBank accession no. P10916) and the results were compared to those obtained for WT hearts expressing full length, non-mutated human ventricular RLC. Three WT hearts and three hearts from each mutant mice (males and females) were used and they closely represented the population of R58Q, K104E or D166V mice studied previously [17–19,23,24]. With the exception of 7–8.5 month-old K104E animals, the age of D166V, R58Q and WT mice was 5–6 months (Table 1). The reason for the chosen age of the mice subjected to microarray analyses was to assess the gene expression patterns in mutants that were thoroughly characterized earlier including the histopathology data on 8 month-old K104E mice [17], and 5–6 month-old D166V [23] and R58Q [19,22] animals. At the age of microarray analysis, all mice demonstrated hypertrophy and diastolic dysfunction but the severity of the phenotype significantly intensified in senescent mice (>13 month-old) compared with age matched WT animals [17–19,23,24]. No phenotypic differences were noted between Tg-WT mice expressing the human isoform of ventricular RLC compared with non-transgenic (NTg) mice containing the mouse cardiac RLC [18,19,21,22], and therefore no NTg samples were included in the current study.

3.1. Comparisons between R58Q, D166V and K104E vs. WT hearts

The principal component analysis (PCA) plots were used to demonstrate the overall gene expression patterns of all three HCM mutated samples [35] (Fig. 1). The three axes in the PCA plots represent the three principle components (PC1, PC2, PC3) calculated from the

samples with each dot representing the expression profile of a single gene. The plot demonstrates that the overall gene expression profile in WT hearts was distinct from the ones with one of the three mutations (Fig. 1). Surprisingly, the K104E and D166V profiles were similar to each other while the R58Q mutant hearts clearly differed from WT and the other two mutants (Fig. 1). The plots of the fold change (FC) in each mutant vs. WT of 29,726 genes are presented in Fig. 2. Each dot in the figure represents one gene, and its position is determined by the FC on x-axis and y-axis. Note that the genes located in 1-a region are more upregulated in D166V than R58Q hearts, while those located in 1-b are more upregulated in R58Q vs. D166V (Fig. 2A). In region 2-a, genes are upregulated in D166V but downregulated in R58Q.while in region 2-b genes are upregulated in D166V but downregulated in R58Q. Similar analysis of mutant pair's comparisons can be done for other x-y quadrants (Fig. 2B and C). The proximity of a gene to the 45° line indicated the degree of similarity between two mutants. All three FC values were combined in Fig. 2D in the 3D plot. Genes that were highly dysregulated are farther to the left or right, while the highly significant changes appear higher on the plot. Note the largest FC between R58Q and K104E with the majority on the x-axis ranging from −4 to 4. K104E and D166V showed the least FC but were most significant compared with R58Q.

3.2. Regulation of structural genes in three models of cardiac hypertrophy

Table 2 presents a list of differentially expressed genes in the three RLC HCM phenotypes vs. WT hearts. The positive FC values indicate upregulated genes, while the negative values indicate downregulated genes. The * symbol denotes FC ≥ 1.5, and ** denotes FC ≥ 2.0. Genes were categorized based on the biological properties/molecular functions of the corresponding proteins. Among structural genes/proteins, the β-myosin heavy chain (MHC) encoded by *Myh7* was largely upregulated in R58Q vs. WT hearts with an FC value of 2.38. Less upregulated *Myh7* was observed in D166V vs. WT hearts (FC = 1.27), while no changes in β-MHC expression were noted in K104E mice. The upregulation of the β-MHC in mouse myocardium and the switch from the α-MHC to β-MHC has been reported in mouse models of HCM [36] and observed to occur in response to a wide variety of pathological insults (reviewed in Ref. [37]). Similar to upregulated β-MHC in both malignant HCM models, collagen VIII encoded by the *Col8a1* gene was also significantly upregulated in R58Q vs. WT hearts (FC = 2.01) and in D166V vs. WT hearts (FC = 1.57), while no changes were found in the benign K104E mouse model of HCM (Table 2). Due to transgenic overexpression of the human cardiac RLC in the hearts of mice, the *Myl7* gene encoding the atrial isoform of mouse RLC was down-regulated in R58Q and K104E mice, while it did not change in D166V compared with WT hearts (Table 2). Consistently with a largely reduced phosphorylation of the RLC in the hearts of D166V and R58Q mice [18,22,24], the level of cardiac myosin light chain kinase (encoded by *Mylk3*) expression was lower in D166V (FC = −2.13) and in R58Q (FC = −1.40) compared with WT hearts Table 2). It was also reduced in K104E hearts (FC = −1.21), consistently with decreased RLC phosphorylation reported in K104E mice [17]. Likewise, the expression of the *Mylk4* gene, which encodes for the myosin light chain kinase family, was significantly downregulated in R58Q hearts and to a lesser degree in K104E vs. WT hearts (Table 2). The other sarcomere structural related genes, such as *Ttn* (titin), *Mybpc3* (MyBP-C), *Tnnc1*

(TnC) and *Tnni3* (TnI) showed no significant changes in any of HCM heart models compared with WT hearts (Table 2).

3.3. Regulation of cytoskeletal, Ca^{2+} handling and HCM-related genes

Abnormalities were found in genes encoding for non-contractile proteins and these may be related to the specific cardiomyopathy phenotypes underlying R58Q, D166V and K104E induced HCM. For example, all three mutations significantly upregulated the *Myot* gene encoding the Z-disc protein myotilin (Table 2), Myotilin location and interactions with other Z-disc proteins are known, but its role in the regulation of muscle structure and function remains unknown [38]. Its upregulation in the three cardiomyopathy models may be related to myofibrillar myopathies characterized by an abnormal accumulation of intrasarcoplasmic proteins and disorganization of the inter-myofibrillar network of the Z-discs [38]. Furthermore, R58Q increased the expression of another Z-disc associated gene, *Pdlim3* encoding the actinin-binding LIM Protein (ALP) (FC = 1.62). Alp is highly expressed in the heart and localizes to Z-discs and intercalated discs. It functions to enhance the crosslinking of actin by alpha actinin-2 and is important for right ventricular chamber formation and contractile function. Its upregulation in R58Q and to a lesser extent in D166V hearts could be compensatory to the left ventricular dysfunction in these mice [22,23]. Strikingly, histone demethylase encoded by the *Uty* gene was largely downregulated in all three cardiomyopathy phenotypes with FC = −2.95, −4.02 and −8.53 in R58Q, D144V and K104E hearts, respectively (Table 2). Histone modifications are known to regulate chromatin structure, transcription and various nuclear processes [39]. Histone demethylases are also important for normal development and are involved in various diseases [39]. Genetic studies have shown that deletion or mutation of several different demethylases can lead to developmental defects in model organisms [40]. Downregulation of *Uty* in our HCM mice may therefore be directly associated with the development of cardiomyopathy phenotypes (Table 2). Interestingly, cysteine dioxygenase 1, cytosolic encoded by *Cdo1* was significantly upregulated in D166V (FC = 2.49) and K104E (FC = 2.39) compared with WT hearts (Table 2). Cysteine dioxygenase 1 initiates several important metabolic pathways and is a critical regulator of cellular cysteine concentrations. Control of cysteine levels by upregulation of *Cdo1* may be necessary to maintain cardiac function in D166V and K104E mice. *Cdo1* is also involved in the biosynthesis of taurine, which is an antioxidant and serves the protective role in the heart during hypoxia [41]. Its increased expression in these hearts may suggest a potential compensatory effect in the heart during hypertrophy/ ischemia. In the group of Ca^{2+} handling genes, *Sln* encoding sarcolipin was significantly downregulated in R58Q vs. WT hearts (FC = −2.25) and even more so in the K104E hearts (FC = −3.68) (Table 2). Sarcolipin is a key regulator of sarco (endo)- plasmic reticulum Ca^{2+}-ATPase and its expression has been found to be altered in diseased atrial myocardium [42]. Downregulation of *Sln* in these mice is most likely associated with increased activity of Ca^{2+}-ATPase pump and abnormal intracellular Ca^{2+} handling.

On the other hand, no abnormalities were found in genes responsible for metabolism (Table 2). As expected, the levels of HCM biomarkers, e.g. atrial natriuretic peptide (ANP) were upregulated in models of malignant HCM (FC = 2.54 in R58Q vs. WT and 2.43 in D166V vs. WT). Brain natriuretic peptide was also upregulated in D166V mice (FC = 1.54). Cardiac

TnI interacting kinase (*Tnni3k*) and potassium channel, voltage gated subfamily E regulatory beta subunit 1 (*Kcne1*) were significantly upregulated in the K104E heart model (FC = 3.58 and 2.19, respectively) and to lesser extent in R58Q mice (FC = 1.62 and 1.70, respectively) compared with WT (Table 2).

3.4. Gene ontology analysis of biological process and molecular functions in R58Q, D166V and K104E mutant heart models

The differentially expressed genes defined as those with an absolute fold change ≥2 in the mutant vs. WT expression levels were identified and subjected to enrichment analyses of Gene Ontology (GO) terms (Fig. 3). Each differentially expressed gene was categorized into corresponding biological processes/molecular functions and only these processes containing at least two differentially expressed genes were included in the pie plots. As shown in Fig. 3A, fewer processes were implicated in D166V when compared to the R58Q and K104E mutants. In the latter two, differentially expressed genes were dispersed in more processes suggesting that they may be less correlated with each other compared with genes involved in biological processes in D166V hearts, which seem to be more strongly correlated with each other (Fig. 3A). All three mutations showed involvement in transcription regulations and oxidation-reduction processes. R58Q and D166V shared involvement in processes such as transport and cell adhesion while R58Q and K104E shared connection in immune response processes (Fig. 3A). It is also worth noting that R58Q hearts were involved in processes such as G-protein coupled receptor signaling pathway, proteolysis and lipid metabolism processes. The processes that were special for the K104E mutation included protein phosphorylation, regulation of heart rate and cardiac conduction and negative regulation of cell growth/proliferation. D166V was clearly the most different among all three mutations and showed the most number of enriched biological process terms related to oxidation-reduction with 29% of all differentially expressed genes. This suggests that the disease-causing mechanism in D166V mice could be largely associated with oxidative stress and/or hypoxia. The other major processes involved in D166V hearts were transcription related (14%), dosage compensation by inactivation of X chromosome (14%) and cell adhesion (11%) processes (Fig. 3A).

Fig 3B shows the involvement of the three HCM mutations, R58Q, D166V and K104E in the regulation of genes responsible for molecular functions. All three mutations shared few major functions including protein binding, nucleotide binding, DNA/RNA binding, metal ion binding and dioxygenase activity. They also shared oxidoreductase activity function with the higher involvement noticed in D166V (14%, Fig. 3B), supporting the notion that D166V may be associated with oxidative stress and/or hypoxia (Fig. 3A). The molecular functions of collagen binding and hydrolase activity were shared by two malignant R58Q and D166V mutations, while protein dimerization was shared by R58Q and K104E hearts (Fig. 3B). Distinct features of R58Q mutation included ATPase (2%) and GTPase (6%) activities, as well as chemokine (2%) and cytokine (4%) activities. The pattern of molecular functions observed across the three HCM mutations suggests that R58Q might be involved in a higher number of different molecular functions than D166V and K104E mutations (Fig. 3B).

3.5. The signaling pathways in three HCM mouse models

The differentially expressed genes (absolute FC ≥ 2) in three HCM models vs. WT were further subjected to the pathway enrichment analysis. All pathways with $p \leq 0.05$ were considered significant and are listed in Table 3. The R58Q mutation was observed to be involved in the most number of pathways compared with D166V and K104E, and they are directly correlated with hypertrophy and/or the heart (Table 3). The dominant pathway observed in R58Q hearts that involved 261 genes was the mitogen-activated protein kinase (MAPK) signaling pathway that consists of a sequence of acting kinases that ultimately results in phosphorylation and activation of terminal kinases, such as c-Jun N-terminal kinases (JNKs) as well as extracellular signal-regulated kinases (ERK) [43]. The MAPK signaling pathway is one of the major pathways in the heart and plays a pivotal role in the development and/or progression of HCM [44]. Other pathways that may be involved in the R58Q mediated hypertrophy include the Toll-like receptor signaling pathway and the Focal adhesion pathway (Table 3). In particular, the Toll-like receptor signaling pathway was found to be critical in ischemia inflammation and injury [45], while any deficiency in the Focal adhesion pathway might be involved in the development of cardiac hypertrophy through the activation of the Focal adhesion kinase [46].

Among seven significantly affected pathways identified in D166V, there were the ECM-receptor interaction pathway ($p = 0.021$), TGF-beta signaling pathway ($p = 0.023$), and the Taurine and hypotaurine metabolism pathway ($p = 0.030$) (Table 3). The ECM-receptor interaction is considered essential in cardiac development and activation may occur in response to pathological signals such as hypertrophic cardiomyopathy [47]. The TGF-beta signaling pathway is associated with cardiac remodeling, fibrosis and thus could be linked to myocardial inflammation and infarction [48]. The Taurine and hypotaurine metabolism pathway that might be related to the compensatory effects of the heart during ischemia was also observed in D166V hearts. Similarly, the latter pathway was significant in K104E hearts ($p = 0.027$) (Table 3). In addition, the tight junction signaling pathway was observed in K104E hearts ($p = 0.042$) and it may be related to cardiac conduction and cell-cell communication [49] in K104E mice (Table 3).

3.6. Cluster and centrality analysis on differentially expressed genes

The gene eco-expression network was computed from the differentially expressed genes lists between two groups or among all three mutants, after which cluster and centrality analyses were performed on the network. Fig. 4A demonstrates how differentially expressed genes that are shared by all three mutants grouped into clusters. There are only ten differentially expressed genes shared by all three mutations, and they are divided into two clusters. The only central node (marked with *) identified was the node: *Xist* (X-chromosome inactive specific transcript). The gene *Xist* (node colored light green) has a strongly negative correlation in its co-expression relationship with the rest of the cluster, which consists of the following four genes (depicted in orange): *Kdm5d*, *Ddx3y*, *Eif2s3y* and *Uty*. Surprisingly, these four nodes do not belong to the same pathway. The other smaller cluster (showed in cyan) between *Cdo1* and *Mir505*, depicts a positive correlation between their expression levels (Fig. 4A). It is of great interest to note that *Mir505* is a microRNA, it cannot be inferred whether or not *Cdo1* induces *Mir505* expression or vice versa. The three isolated

nodes: *Slc6a16*, *Ifi204* and *Lgi1* were not correlated with any other genes shown in the graph (Fig. 4A).

Fig. 4B demonstrates the comparison of two malignant mutations: D166V and R58Q. More shared differentially expressed genes were observed, and more clusters were formed. The cluster with *Xist* has one extra node: *Nmrk2*, which is strongly (negatively) correlated with the previously mentioned *Cdo1* and another central node *Mylk4*, which acts as a "bridge" connecting two clusters (orange and yellow clusters). Three out of four genes (*Ltbp2*, *Comp* and *Thbs4*) in the yellow cluster are associated with TGF-beta signaling pathway. Besides a mutually inimical pair of genes (*Prg4* and *Hsph1*), eight isolated nodes were seen, including *Lgi1* and *Ifi204*, which are differentially expressed genes shared by all three mutants. The central nodes identified were: *Nmrk2*, *Xist*, *Mylk4* and *Ltbp2*. Fig. 4C shows the results for D166V vs. K104E. The pattern is similar to the pattern seen in the group of all three mutants in Fig. 4A, but with two different clusters: the cyan clusters with *Slc6a16* correlated with a new gene *Akr1e1*, and the purple cluster with *Mir505* correlated with *Ifi204* and *Lgi1*, but surprisingly without *Cdo1*, presumably due to the interference of R58Q: in which the *Mir505* is no longer positively correlated with *Ifi204* and *Lgi1*, but it is positively correlated with *Cdo1*, and the variance is so large that the algorithm clustered *Cdo1* with *Mir505* rather than *Ifi204* and *Lgi1*. The central nodes identified in this group are *Xist* and *Ifi204*. Fig. 4D shows the comparison of results between K104E and R58Q. The large amount of isolated nodes and smaller clusters suggested the possibility of a different mechanism for the action of the shared differentially expressed genes. The whole gene expression profile between K104E and R58Q that showed the majority of the genes located in the second and fourth quadrant (Fig. 2B), also provided some indirect evidence to support this hypothesis. In addition, except for the common cluster that have been consistently seen through Fig. 4A–C, other clusters are small in size and are not related to any major pathways. This result may indicate distinct gene expression patterns between the K104E and R58Q mutations.

4. Discussion

Different clinical [9,11–14,16] and functional [17–19,22,23] phenotypes of HCM associated with the R58Q, D166V and K104E mutations in the RLC have been previously reported. The HCM is an autosomal dominant disease that can be caused by single or multisite mutations in all major sarcomeric proteins; however, the molecular mechanisms, severity of HCM and progression to heart failure are not known. In particular, overall gene expression patterns and potential signaling pathways are still to be elucidated. In this report we utilized microarray and related bioinformatics analyses to examine the gene expression profiles among two malignant (R58Q and D166V) and one benign (K104E) HCM-causing mutations in the myosin RLC. We have already conducted substantial research on all three mutations using traditional molecular biology, biochemistry and biophysics techniques and provided a detailed characterization of the hypertrophic state of the heart in all tested animals [17–19,23,24]. Histopathological characterization showed hypertrophy and diastolic dysfunction in all mutant mice compared to WT [22,23,50], but the study on senescent animals (not used in the microarray experiments) revealed the progressive nature of the HCM phenotype in all mutant vs. age matched WT mice [17,22,23,50]. The major contractile findings on two malignant R58Q and D166V mutations included a mutation-induced decrease in maximal

pCa 4 force and abnormally increased calcium sensitivity of contraction. Both mutations resulted in largely reduced RLC phosphorylation and these changes subsequently led to systolic and diastolic dysfunction *in vivo* [18,22–24,51,52]. On the other hand, studies on Tg-K104E mice showed that the K104E mutation also led to decreased force generation and reduced RLC phosphorylation; however, myofilament calcium sensitivity was not affected and the phenotype was relatively benign in younger mice [17,50]. However, mice older than 13 months of age showed a wide range of HCM abnormalities *in vitro* and *in vivo* [17,50]. These findings suggested that the K104E mutation may display a distinct pattern of differentially expressed genes compared to the other two mutations, D166V and R58Q. The patterns of differentially expressed genes in D166V and R58Q hearts may be similar but it is also possible that changes in different genes/pathways may result in similar changes in phenotypes.

To our surprise, as shown in the PCA plot in Fig. 1, the absolute gene expression pattern of D166V was more similar to the K104E mutation, while the profile of R58Q was distinct. However, after normalizing to WT control, the relative gene expression distribution across all three mutations showed that the major contribution to the difference between D166V and R58Q was in genes that simultaneously increased or decreased in both phenotypes but with different FC, while expression of genes in K104E did not change compared to WT. This was true for an upregulated β-MHC in both malignant HCM models and collagen VIII, while no changes were found in the benign K104E mouse model of HCM (Table 2). Likewise, the level of HCM biomarkers were significantly upregulated in malignant HCM models vs. WT while no change vs. WT was observed in benign K104E hearts (Table 2). These results suggested that the potential molecular mechanisms underlying the D166V and R58Q phenotypes lie in the number of shared gene clusters or pathways that are upregulated or downregulated simultaneously, while for K104E, fewer clusters or significantly affected pathways are present (Fig. 4). Interestingly, in biological processes, 29% of genes in D166V were related to oxidation-reduction processes, taking 29% of total differentially expressed genes (Fig. 3). This suggested that the disease causing mechanism of D166V could be largely associated with oxidative stress. The molecular functions of collagen binding and hydrolase activity were shared by two malignant R58Q and D166V mutations (Fig. 3B). The pathway analysis based on differentially expressed gene lists (Table 3) showed a few cardiovascular related pathways in R58Q that are important in ischemia injury and myocardial inflammation (Toll-like receptors signaling pathway) [53], in regulating development of eccentric hypertrophy under hypertrophic stimulation (Focal adhesion pathway) [54], and in the development and/or progression of HCM through the MAPK signaling pathway. The MAPK signaling pathway is one of the major pathways in the heart playing a pivotal role in HCM-dependent heart remodeling [44]. For D166V, two pathways were of potential interest, the TGF-beta signaling pathway and Taurine and hypotaurine metabolism. The TGF-beta plays an important role in the development of cardiac fibrosis and its expression is enhanced during hypertrophy [55]. Consistently, the TGF-beta signaling pathway, which is associated with cardiac remodeling in response to myocardial inflammation and infarction [48] may contribute to D166V induced HCM. Likewise, the Taurine and hypotaurine metabolism pathway may be related to the compensatory effects of

the heart during ischemic episodes in D166V hearts. This pathway was also significant in K104E hearts (Table 3).

The clustering and centrality analysis [29,31] applied to the three mutations together, showed that the common clusters are scarce and that each of the three phenotypes may have its own distinct disease causing mechanism. The only related information is for the *Cdo1* gene, which positively correlated with an unknown microRNA: *Mir505*, suggesting hypoxia conditions may occur in all three mutants. The cluster with *Uty*, *Eif2s*, *Ddx3y* and *Kdm5d* is linked with its antagonistic gene *Xist* and may be to some extent associated with Down syndrome; however, its potential effects in HCM remain unknown. *Mir505* was also upregulated in comparison of K104E and D166V, regulating another gene *Ifi204* (old name: P204, interferon activated gene 204), which may be involved in cardiomyocyte proliferation and differentiation [56]. The analysis of the results for two malignant mutations, D166V and R58Q agrees with what was shown in Fig 2A, where many genes were positively correlated suggesting substantial shared gene clusters in these two mutant hearts (Fig. 4B). It is also worth noting that the central node that represents the *Mylk4* gene forms a bridge between the TGF-beta cluster and *Nmrk2* (Nicotinamide Riboside Kinase 2), which contributes to NAD^+ salvage and regulates skeletal muscle adaption. This may also imply that *Mylk4* does not only phosphorylate targets during skeletal muscle adaption, but may also affect TGF-beta signaling pathway in a direct or indirect manner. The results for D166V vs. K104E (Fig. 4C) showed a pattern similar to one seen in the group of all three mutants (Fig. 4A), while the results for K104E and R58Q group indicated a large amount of unconnected differentially expressed genes (Fig. 4D). This could be due to the same direction of changes observed for genes in K104E and in R58Q hearts, where potential “positive” or “negative” correlations between the gene clusters have been canceled.

5. Conclusions

The comparison of differentially expressed gene patterns for all analyzed RLC hearts showed that all three HCM-RLC heart models were clearly distinct from WT hearts. Each mutation displayed distinctive gene expression profiles indicating mutation-specific disease mechanisms. The R58Q mutation was observed to be involved in the most number of signaling pathways compared with D166V and K104E, and they were directly correlated with the development and progression of cardiomyopathy disease. The most abundant biological processes observed in D166V hearts were related to the oxidation-reduction pathways important in hypoxia and/or myocardial inflammation/infarction. The effect of K104E on the signaling pathways was least pronounced, which agrees with its benign phenotype in humans. Our results suggest that the potential molecular mechanisms underlying two malignant D166V and R58Q phenotypes lie in the number of shared gene clusters or pathways that are upregulated or downregulated simultaneously, while for K104E, fewer clusters or significantly affected pathways were present.

5.1. Limitations of the microarray study

One has to acknowledge the limitations of the current microarray study in drawing conclusions on the disease mechanisms. Even though genes with high FC values were

considered, more than 3 mice per group, as used in this study, would admittedly lead to inferences with higher statistical power. Note that Partek® Genomics Suite™ software uses ANOVA with Least Square means to compute fold change values. Future studies with a larger sample size should also include separate groups of male and female animals to assess potential gender related mechanisms. Although phenotype differences between the different models were taken into consideration, animal age, sex, transgenic line and potentially other contributing disease factors (such as inflammation) could also play a role in differential gene expression to cause disease and future studies considering these factors would further strengthen mechanistic understanding. Importantly, the results from this microarray study warrant a phenotypic verification at the protein expression level using molecular proteomics.

Acknowledgments

This work was supported by U.S. National Institutes of Health (NIH) grants HL-123255 and HL-108343 (DSC) and the American Heart Association grants 12PRE12030412 (WH) and 15POST25080302 (ZZ). The efforts of VAP and GN were supported by a grant from the Alpha-One Foundation and by funding from the College of Engineering and Computing at Florida International University.

Abbreviations

HCM	hypertrophic cardiomyopathy
RLC	myosin regulatory light chain
R58Q	Arginine 58→Glutamine mutation in myosin RLC
D166V	Aspartic Acid 166→Valine mutation in myosin RLC
K104E	Lysine 104→Glutamic acid mutation in myosin RLC
WT	wild-type human ventricular RLC
GO	Gene Ontology
FC	fold-change
PCA	Principle Component Analysis

References

1. Alcalai R, Seidman JG, Seidman CE. J Cardiovasc Electrophysiol. 2008; 19:104–110. [PubMed: 17916152]
2. Maron BJ, Doerer JJ, Haas TS, Tierney DM, Mueller FO. Circulation. 2009; 119:1085–1092. [PubMed: 19221222]
3. Maron BJ. Card Electrophysiol Rev. 2002; 6:100–103. [PubMed: 11984027]
4. Muthu, P., Huang, W., Kazmierczak, K., Szczesna-Cordary, D. Cardiomyopathies – from Basic Research to Clinical Management. Veselka, J., editor. InTech; Croatia: 2012. p. 383-408.Ch. 17
5. Szczesna D. Current drug targets, Cardiovasc Haematol Disord. 2003; 3:187–197.
6. Olivotto I, Girolami F, Ackerman MJ, Nistri S, Bos JM, Zachara E, Ommen SR, Theis JL, Vaubel RA, Re F, Armentano C, Poggesi C, Torricelli F, Cecchi F. Mayo Clin Proc. 2008; 83:630–638. [PubMed: 18533079]
7. Santos S, Marques V, Pires M, Silveira L, Oliveira H, Lanca V, Brito D, Madeira H, Fonseca E, Freitas A, Carreira I, Gaspar I, Monteiro C, Fernandes A. BMC Med Gen. 2012; 13:17.

8. Alvarez-Acosta L, Mazzanti A, Fernández X, Ortí M, Barriales-Villa R, García D, Maneiro E, Rebolo P, Álvarez E, Monserrat L. J Cardiovasc Dis. 2014; 2:82–90.
9. Andersen PS, Havndrup O, Hougs L, Sorensen KM, Jensen M, Larsen LA, Hedley P, Thomsen AR, Moolman-Smook J, Christiansen M, Bundgaard H. Hum Mutat. 2009; 30:363–370. [PubMed: 19035361]
10. Garcia-Pavia P, Vazquez ME, Segovia J, Salas C, Avellana P, Gomez-Bueno M, Vilches C, Gallardo ME, Garesse R, Molano J, Bornstein B, Alonso-Pulpon L. Eur J Heart Fail. 2011; 13:1193–1201. [PubMed: 21896538]
11. Flavigny J, Richard P, Isnard R, Carrier L, Charron P, Bonne G, Forissier JF, Desnos M, Dubourg O, Komajda M, Schwartz K, Hainque B. J Mol Med. 1998; 76:208–214. [PubMed: 9535554]
12. Kabaeva ZT, Perrot A, Wolter B, Dietz R, Cardim N, Correia JM, Schulte HD, Aldashev AA, Mirrakhimov MM, Osterziel KJ. Eur J Hum Genet. 2002; 10:741–748. [PubMed: 12404107]
13. Morner S, Richard P, Kazzam E, Hellman U, Hainque B, Schwartz K, Waldenstrom A. J Mol Cell Card. 2003; 35:841–849.
14. Richard P, Charron P, Carrier L, Ledeuil C, Cheav T, Pichereau C, Benaiche A, Isnard R, Dubourg O, Burban M, Gueffet J-P, Millaire A, Desnos M, Schwartz K, Hainque B, Komajda M. EUROGENE Heart Failure Project, Circulation 107 (2003) 2227–2232 and erratum. Circulation. 2004; 109:3258.
15. Liu Y, Ringner M. R77, Genome Biol. 2007; 8
16. Andersen PS, Havndrup O, Bundgaard H, Moolman-Smook JC, Larsen LA, Mogensen J, Brink PA, BÃŗglum AD, Corfield VA, Kjeldsen K, Vuust J, Christiansen M. J Med Genet. 2001; 38:e43. [PubMed: 11748309]
17. Huang W, Liang J, Kazmierczak K, Muthu P, Duggal D, Farman GP, Sorensen L, Pozios I, Abraham T, Moore JR, Borejdo J, Szczesna-Cordary D. J Mol Cell Cardiol. 2014; 74:318–329. [PubMed: 24992035]
18. Kerrick WGL, Kazmierczak K, Xu Y, Wang Y, Szczesna-Cordary D. FASEB J Off Publ Fed Am Soc Exp Biol. 2009; 23:855–865.
19. Wang Y, Xu Y, Kerrick WGL, Wang Y, Guzman G, Diaz-Perez Z, Szczesna-Cordary D. J Mol Biol. 2006; 361:286–299. [PubMed: 16837010]
20. Wang L, Muthu P, Szczesna-Cordary D, Kawai M. J Mol Cell Cardiol. 2013; 62:153–163. [PubMed: 23727233]
21. Greenberg MJ, Watt JD, Jones M, Kazmierczak K, Szczesna-Cordary D, Moore JR. J Mol Cell Cardiol. 2009; 46:108–115. [PubMed: 18929571]
22. Abraham TP, Jones M, Kazmierczak K, Liang H-Y, Pinheiro AC, Wagg CS, Lopaschuk GD, Szczesna-Cordary D. Cardiovasc Res. 2009; 82:84–92. [PubMed: 19150977]
23. Yuan CC, Muthu P, Kazmierczak K, Liang J, Huang W, Irving TC, Kanashiro-Takeuchi RM, Hare JM, Szczesna-Cordary D. Proc Natl Acad Sci U S A. 2015; 112:E4138–E4146. [PubMed: 26124132]
24. Muthu P, Kazmierczak K, Jones M, Szczesna-Cordary D. J Cell Mol Med. 2012; 16:911–919. [PubMed: 21696541]
25. Cickovski T, Peake E, Aguiar-Pulido V, Narasimhan G. Proceedings of ICCABS. 2015
26. Newman, MEJ. Networks : an Introduction. Oxford University Press; Oxford; New York: 2010.
27. Narang V, Ramli MA, Singhal A, Kumar P, de Libero G, Poidinger M, Monterola C. PLoS Comput Biol. 2015; 11:e1004504. [PubMed: 26393364]
28. Bussemaker HJ, Foat BC, Ward LD. Annu Rev Biophys Biomol Struct. 2007; 36:329–347. [PubMed: 17311525]
29. Fernandez-Romero JA, Deal C, Herold BC, Schiller J, Patton D, Zydowsky T, Romano J, Petro CD, Narasimhan M. Trends Microbiol. 2015; 23:429–436. [PubMed: 25759332]
30. Stuart JM, Segal E, Koller D, Kim SK. Science. 2003; 302:249–255. [PubMed: 12934013]
31. Cickovski T, Flor T, Irving-Sachs G, Novikov P, Parda J, Narasimhan G. IEEE/ACM transactions on computational biology and bioinformatics/IEEE, ACM. 2015; 12:445–454.
32. Kazmierczak K, Xu Y, Jones M, Guzman G, Hernandez OM, Kerrick WGL, Szczesna-Cordary D. J Mol Biol. 2009; 387:706–725. [PubMed: 19361417]

33. Reed E, Nunez S, Kulp D, Qian J, Reilly MP, Foulkes AS. Stat Med. 2015; 34:3769–3792. [PubMed: 26343929]
34. Fernandez M, Riveros JD, Campos M, Mathee K, Narasimhan G. BMC Genom. 2015 Nov 10.16(Suppl. 11):S6. http://dx.doi.org/10.1186/1471-2164-16-S11-S6. Epub 2015 Nov 10.
35. Luo J, Schumacher M, Scherer A, Sanoudou D, Megherbi D, Davison T, Shi T, Tong W, Shi L, Hong H, Zhao C, Elloumi F, Shi W, Thomas R, Lin S, Tillinghast G, Liu G, Zhou Y, Herman D, Li Y, Deng Y, Fang H, Bushel P, Woods M, Zhang J. Pharmacogenomics J. 2010; 10:278–291. [PubMed: 20676067]
36. Barefield D, Kumar M, de Tombe PP, Sadayappan S. Am J Physiol Heart Circ Physiol. 2014; 306:H807–H815. [PubMed: 24464755]
37. Harvey PA, Leinwand LA. J Cell Biol. 2011; 194:355–365. [PubMed: 21825071]
38. Ochala J, Carpén O, Larsson L. Upsala J Med Sci. 2009; 114:235–241. [PubMed: 19878039]
39. Kooistra SM, Helin K. Nat Rev Mol Cell Biol. 2012; 13:297–311. [PubMed: 22473470]
40. Pedersen MT, Helin K. Trends Cell Biol. 2010; 20:662–671. [PubMed: 20863703]
41. Karar J, Dolt KS, Mishra MK, Arif E, Javed S, Pasha MA. FEBS Lett. 2007; 581:4577–4582. [PubMed: 17825300]
42. Xie LH, Shanmugam M, Park JY, Zhao Z, Wen H, Tian B, Periasamy M, Babu GJ. Am J Physiol Cell Physiol. 2012; 302:C1762–C1771. [PubMed: 22496245]
43. Garrington TP, Johnson GL. Curr Opin Cell Biol. 1999; 11:211–218. [PubMed: 10209154]
44. Heineke J, Molkentin JD. Nat Rev Mol Cell Biol. 2006; 7:589–600. [PubMed: 16936699]
45. Timmers L, Sluijter JP, van Keulen JK, Hoefer IE, Nederhoff MG, Goumans MJ, Doevendans PA, van Echteld CJ, Joles JA, Quax PH, Piek JJ, Pasterkamp G, de Kleijn DP. Circ Res. 2008; 102:257–264. [PubMed: 18007026]
46. Clemente CF, Xavier-Neto J, Dalla Costa AP, Consonni SR, Antunes JE, Rocco SA, Pereira MB, Judice CC, Strauss B, Joazeiro PP, Matos-Souza JR, Franchini KG. J Mol Cell Cardiol. 2012; 52:493–501. [PubMed: 22056317]
47. Ross RS, Borg TK. Circ Res. 2001; 88:1112–1119. [PubMed: 11397776]
48. Bujak M, Frangogiannis NG. Cardiovasc Res. 2007; 74:184–195. [PubMed: 17109837]
49. Lisewski U, Shi Y, Wrackmeyer U, Fischer R, Chen C, Schirdewan A, Juttner R, Rathjen F, Poller W, Radke MH, Gotthardt M. J Exp Med. 2008; 205:2369–2379. [PubMed: 18794341]
50. Huang W. Open Access Dissertations. 2015 Paper 1373.
51. Greenberg MJ, Kazmierczak K, Szczesna-Cordary D, Moore JR. Proc Natl Acad Sci U S A. 2010; 107:17403–17408. [PubMed: 20855589]
52. Karabina A, Kazmierczak K, Szczesna-Cordary D, Moore JR. Arch Biochem Biophys. 2015
53. Chao W. Am J Physiol Heart Circ Physiol. 2009; 296:H1–H12. [PubMed: 19011041]
54. Peng X, Kraus MS, Wei H, Shen TL, Pariaut R, Alcaraz A, Ji G, Cheng L, Yang Q, Kotlikoff MI, Chen J, Chien K, Gu H, Guan JL. J Clin Investig. 2006; 116:217–227. [PubMed: 16374517]
55. Dobaczewski M, Chen W, Frangogiannis NG. J Mol Cell Cardiol. 2011; 51:600–606. [PubMed: 21059352]
56. Luan Y, Lengyel P, Liu CJ. Cytokine & Growth Factor Rev. 2008; 19:357–369. [PubMed: 19027346]

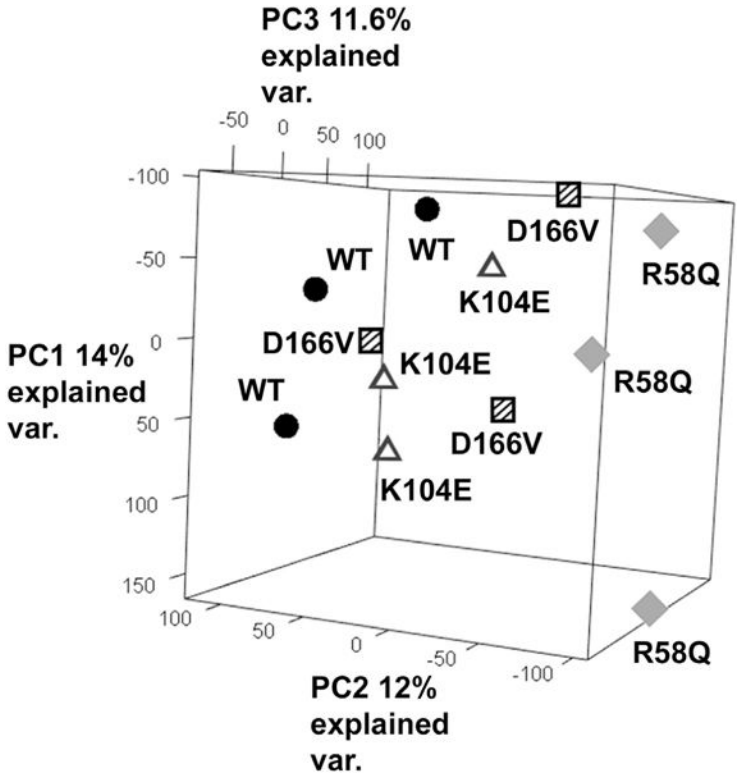

Fig. 1.

The Principle Component Analysis (PCA) plots of overall gene expression patterns in three different HCM-RLC phenotypes and WT hearts. Note that the K104E (open triangles) and D166V (squares) profiles are similar to each other while those present in R58Q (diamonds) differ from WT (black circles) and two mutants.

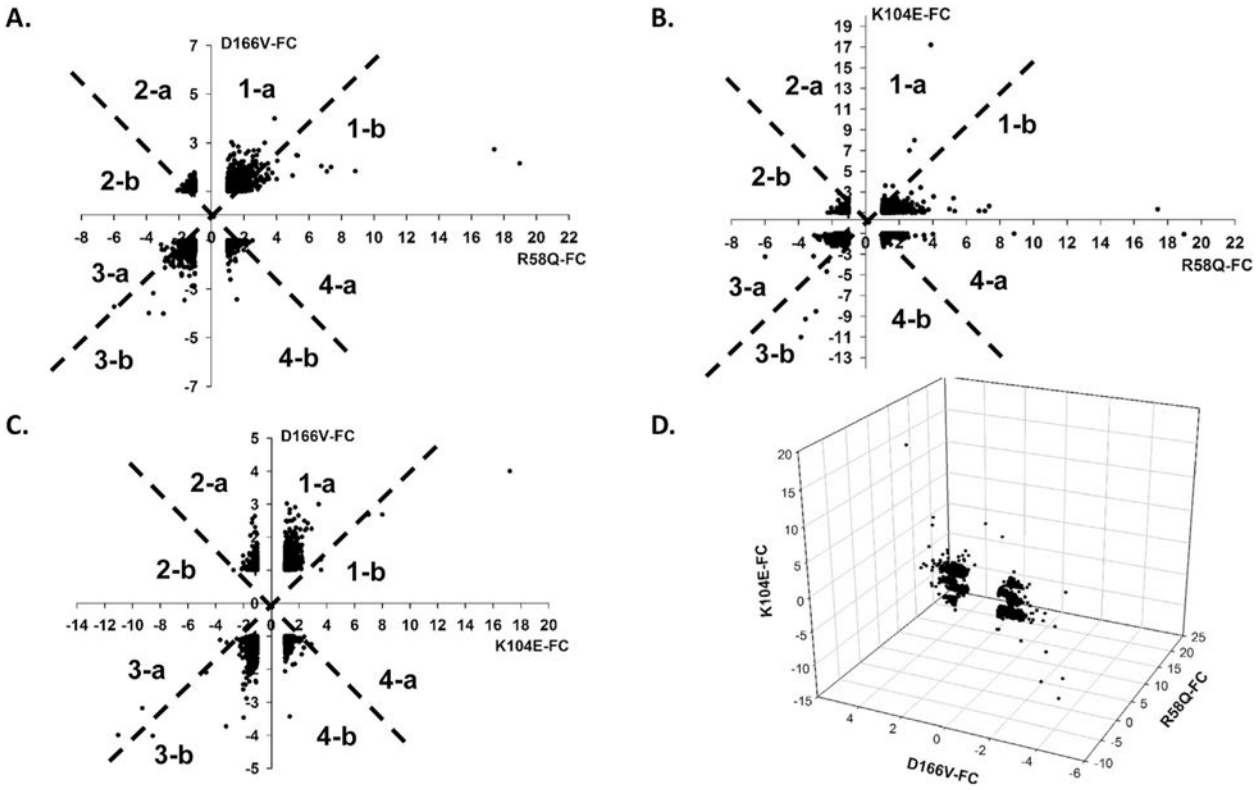

Fig. 2.

Scatter plots of fold change (FC with respect to WT) of the 29,726 genes. A. Overall gene expression pattern of R58Q vs. D166V. B. R58Q vs. K104E. C. K104E vs. D166V. D. 3D plot of overall gene expression in all three HCM-RLC mutants: R58Q vs D166V vs. K104E. Each quadrant in the x-y coordinates has been divided into two parts (a and b) divided by a 45 line generating regions: 1-a, 1-b, 2-a, 2-b, 3-a, 3-b, 4-a and 4-b. The proximity of a gene to the 45 line indicates the similarity between the profiles. Positive FC indicates gene upregulation while the negative FC, downregulation.

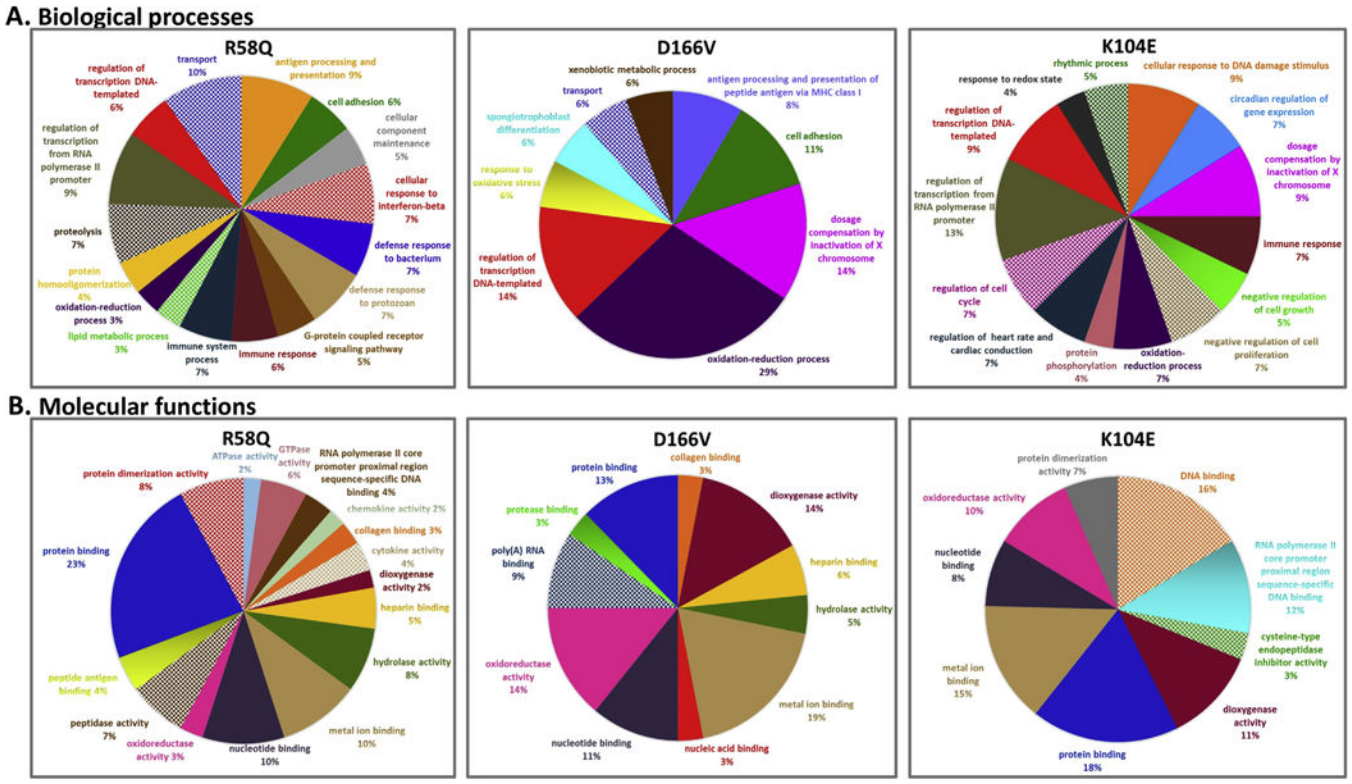

Fig. 3.

Biological processes (A) and molecular functions (B) associated with differentially expressed genes in the hearts of R58Q, D166V and K104E vs. WT mice. The pie charts were drawn based on the percentage of enrichment of differentially expressed genes involved in specific biological processes and molecular functions.

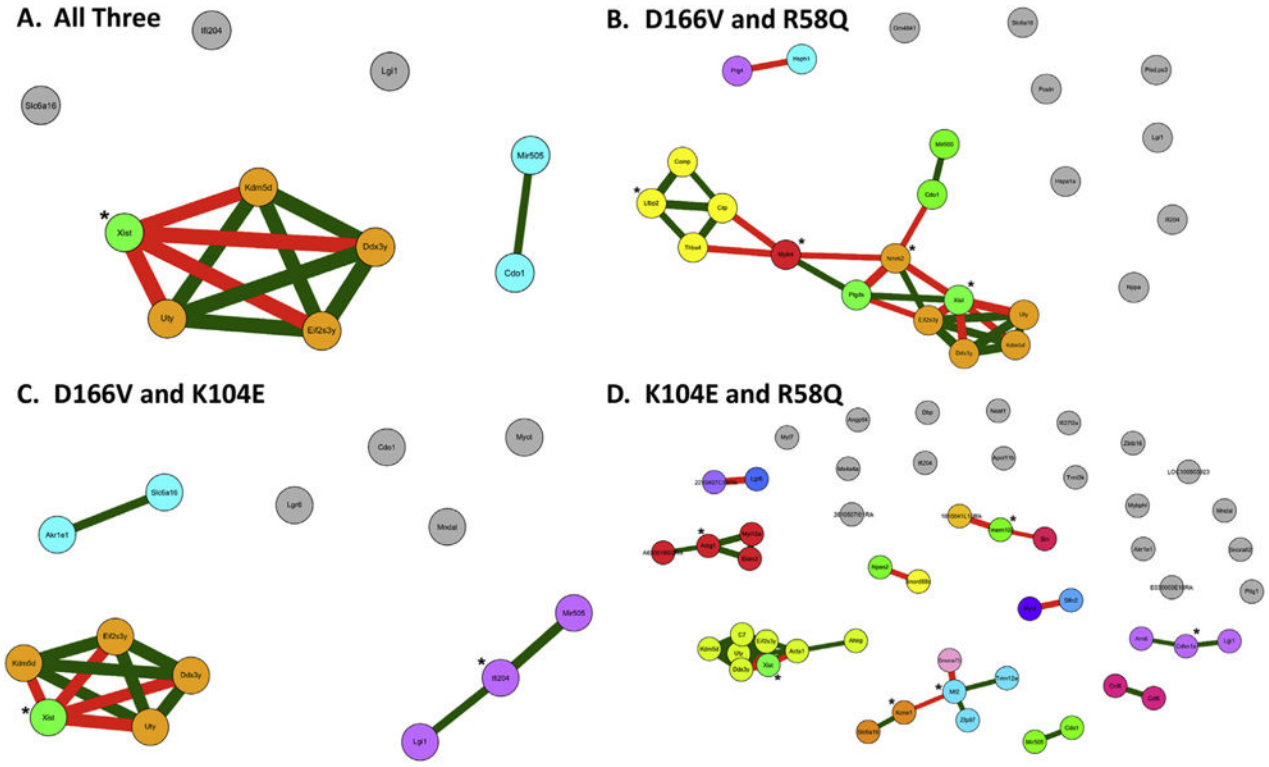

Fig. 4.

Cluster and centrality analysis of differentially expressed genes. A. Co-occurrence network with only the differentially expressed genes between all three mutants; B. between D166V and R58Q. C. between D166V and K104E, and D. between K104E and R58Q. For each graph, each node represents one gene, and the nodes of same color belong to a cluster of genes that show strong correlation in their expression levels. The isolated nodes depicted in grey color suggest that these nodes are not belonging to any cluster and thus are not significantly correlated with other genes shown in the graph. The node marked with "*" indicates *central* nodes, namely, the most potentially influential node within its neighborhood. The edges in the graph represent correlations between genes with green edges suggesting positive correlations of their expression levels and red edges suggesting negative correlations. The positive correlation indicates the pair of genes are upregulated and downregulated simultaneously, while the negative correlation indicates one gene is upregulated and the other gene is downregulated.

Table 1

List of WT or HCM –RLC (B6SJL) mice used for mRNA microarray.

Mice	R58Q			D166V			K104E			WT		
	Sex	Line	Age [months]	Sex	Line	Age [months]	Sex	Line	Age [months]	Sex	Line	Age [months]
1	M	9	5	M	4	5	M	7	8	M	2	5
2	F	8	5	F	4	6	F	2	7	M	2	5
3	M	9	5	M	4	5	F	3	8.5	M	2	5

Three hearts per group were used (n = 3).

Table 2

Genes differentially expressed in three RLC HCM phenotypes.

Protein	Gene symbol	Fold-change (FC)		
		R58Q vs. WT	D166V vs. WT	K104E vs. WT
Structural				
β-myosin	Myh7	2.38**	1.27	−1.11
cMLCK	Mylk3	−1.40	2.13*	−1.21
Atrial ELC	Myl4	−1.35	1.25	−1.76*
Collagen VIII	Col8a1	2.01**	1.57*	−1.01
Skeletal alpha actin	Acta1	−1.07	1.03	−2.13**
Matrix metallopeptidase 2	Mmp2	2.62**	−1.02	1.13
Titin	Ttn	1.10	1.09	1.02
MyBP-C	Mybpc3	−1.06	1.06	−1.14
Troponin C, cardiac/slow skeletal	Tnnc1	1.07	1.13	1.01
Troponin I, cardiac	Tnni3	−1.02	1.05	1.02
Myosin binding protein H like	Mybphl	0.59	0.84	−2.07**
Myosin regulatory light chain 2, atrial form	Myl7	−2.00**	0.79	−2.21**
Myosin light chain kinase family member 4	Mylk4	−2.85**	0.47	−1.46
Cytoskeletal				
Four-and-a-half LIM protein 1	Fhl1	1.26	−1.08	−1.00
Skeletal muscle α 2 actinin	Actn2	−1.06	1.01	−1.06
Nebulin-related protein	Nrap	−1.07	1.01	−1.12
Desmin	Des	1.22	1.20	1.00
Muscle LIM Protein(MLP)	Csrp3	−1.02	−1.00	−1.07
Telethonin	Tcap	−1.05	1.13	−1.63*
Calsarcin1	Myoz2	1.11	1.11	1.00
LIM domain-binding protein 3	Ldb3	−1.25	−1.18	−1.15
Actinin-binding LIM Protein (ALP)	Pdlim3	1.62*	1.36	1.15
Myotilin	Myot	1.97*	2.23**	2.10**
Nebulette	Nebl	−1.19	−1.24	−1.09
Myopalladin	Mypn	−1.51	−1.19	1.00
Nexilin	Nexn	1.06	−1.01	1.09
Ankyrin repeat domain 1	Ankrd1	−1.29	−1.10	−1.28
Obscurin	Obscn	−1.02	1.07	−1.00
Histone demethylase	Uty	−2.95**	−4.02**	−8.53**
Protein synthesis				
Transcription elongation factor 1α1	Eef1a1	1.44	1.25	1.09
A+U-rich element RNA binding factor	Hnrpdl	1.22	1.04	−1.02
Mitochondrial ribosomal protein L52	Mrpl52	1.55*	1.34	1.29
Ribosomal protein L10	Rpl10	1.37	1.17	1.27

Protein	Gene symbol	Fold-change (FC)		
		R58Q vs. WT	D166V vs. WT	K104E vs. WT
Ribosomal protein L13	Rpl13	1.69 *	1.20	1.30
Redox system				
NADH ubiquinone oxidoreductase	Ndufb3	1.27	1.34	1.18
Ubiquinol-cytochrome c reductase hinge protein	Uqcrh	1.25	1.23	1.18
NADH dehydrogenase Fe–S protein 6	Ndufs6	1.22	1.10	1.05
Aldo-keto reductase family 1, member E1	Akr1e1	1.72 *	2.00 **	2.22 **
Cysteine dioxygenase 1, cytosolic	Cdo1	1.34	2.49 **	2.39 **
Ca^{2+} handling				
SERCA2a	Atp2a2	−1.13	1.00	−1.00
RyR2	Ryr2	−1.12	−1.02	−1.04
NCX	Slc8a1	−1.23	−1.17	−1.08
L-type Channel Channel, cardiac	Cacna1c	−1.03	1.00	1.03
Phospholamban	Pln	−1.03	−1.01	1.01
Calcineurin 1	Rcan1	−1.17	−1.02	1.38
Calmodulin	Calm1	−1.02	−1.06	1.03
Plasma membrane Ca^{2+} ATPase (PMCA4)	Atp2b4	1.28	1.31	1.3
G protein coupled receptor	Agtr1b	−1.27	−1.12	−1.16
Calcium/Calmodulin-Dependent Protein Kinase II Alpha	Camk2a	−1.34	−1.05	−1.05
c-Myc	Myc	1.37	1.18	1.24
Mitogen-Activated Protein Kinase 8 (JNK)	Mapk8	−1.01	−1.00	1.06
Protein tyrosine phosphatases (PTP) 22	Ptpn22	1.5 *	−1.19	1.13
MAP Kinase Phosphatase 5	Dusp10	1.21	1.08	1.07
Transforming protein p21	Hras1	1.26	1.17	1.01
Epidermal growth factor	Egf	−2.29 **	−1.71 *	−1.49
Stromal Interaction Molecule 1	Stim1	−1.22	−1.07	−1.04
Protein kinase C, alpha	Prkca	−1.21	−1.16	−1.11
RAS guanyl-releasing protein 3	RasGRP3	−1.21	−1.11	−1.1
Serum Response Factor	Srf	−1.24	−1.12	−1.01
Fibroblast growth factor-1 (FGF-1)	Fgf1	−1.45	−1.41	−1.35
Elk-1	Elk1	−1.21	−1.1	−1.11
Sarcolipin (associated with sarcoplasmic reticulum calcium ion transport)	Sln	−2.25 **	0.90	−3.68 **
Metabolism				
Mitochondria related				
Apoptosis-inducing factor, mitochondrion-associated, 1 (For MT content quantification)	Aifm1	1.07	1.05	1.19
Porin	Vdac1	−1.04	−1.01	−1.03
Fatty acid oxidation related				
Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	Acadm	−1.32	−1.20	−1.11
PPARalpha	Ppara	−1.10	−1.04	−1.01
Acetoacetyl-CoA synthetase	Aacs	−1.08	−1.03	−1.05

Protein	Gene symbol	Fold-change (FC)		
		R58Q vs. WT	D166V vs. WT	K104E vs. WT
Glucose metabolism related				
cAMP-dependent protein kinase catalytic subunit alpha (AMPK)	Prkaca	−1.23	−1.18	−1.19
Glucose transporter 4 (GLUT4)	Slc2a4	−1.22	−1.02	−1.07
L-type pyruvate kinase (muscle)	Pkm	−1.10	1.01	−1.12
HCM biomarkers, others				
Atrial natriuretic peptide (ANP)	Nppa	2.54**	2.43**	−1.38
Brain natriuretic peptide (BNP)	Nppb	1.20	1.54*	1.26
Serine protease inhibitor A3A	Serpina3a	1.13	1.39	1.26
Heat Shock 70kD protein 8	Hspa8	1.44	−1.18	1.05
Cardiac Troponin I interacting kinase	Tnni3k	1.62*	1.32	3.58**
Potassium channel, voltage gated subfamily E regulatory beta subunit 1 (regulate heart rate and contractility)	Kcne1	1.70*	1.43	2.19**

Positive FC values indicate genes that were upregulated vs. WT, while negative FC values depict genes that were downregulated vs. WT in three HCM mouse models. Symbol

*denotes FC ≥ 1.5 and

**FC ≥ 2.0. Genes were selected through the complete gene list and categorized based on biological properties/functions.

Table 3

Number of up- or down-regulated mRNAs in signaling pathways.

	Signaling pathways	Number of genes that hit	Total number of genes	Enrichment score	*p* value
R58Q					
1.	Antigen processing and presentation	4	68	8.87	0.0013
2.	Prion diseases	3	35	12.92	0.0020
3.	Bladder cancer	3	42	10.77	0.0033
4.	Toll-like receptor signaling pathway	4	99	6.09	0.0048
5.	Amyotrophic lateral sclerosis	3	55	8.22	0.0066
6.	NOD-like receptor signaling pathway	3	55	8.22	0.0066
7.	Focal adhesion	5	194	3.88	0.010
8.	Arachidonic acid metabolism	3	75	6.03	0.015
9.	Cytokine-cytokine receptor interaction	5	231	3.26	0.020
10.	Chemokine signaling pathway	4	175	3.45	0.031
11.	MAPK signaling pathway	5	261	2.89	0.032
12.	Lysosome	3	120	3.77	0.048
D166V					
1.	Malaria	2	47	15.56	0.0080
2.	Arachidonic acid metabolism	2	75	9.75	0.019
3.	ECM receptor interaction	2	79	9.26	0.021
4.	TGF-beta signaling pathway	2	83	8.81	0.023
5.	Ribosome	2	91	8.04	0.027
6.	Taurine and hypotaurine metabolism	1	10	36.57	0.030
7.	Steroid biosynthesis	1	17	21.51	0.048
K104E					
1.	Circadian rhythm	2	13	62.40	0.0006
2.	Taurine and hypotaurine metabolism	1	10	40.56	0.027
3.	Leukocyte trans-endothelial migration	2	115	7.05	0.033
4.	Cell cycle	2	124	6.54	0.039
5.	Tight junction	2	130	6.24	0.042

Only signaling pathways with enrichment *p*-value <0.05 are listed.