Published in final edited form as:
Circ Cardiovasc Genet. 2008 December ; 1(2): 117–125. doi:10.1161/CIRCGENETICS.108.802652.

SEX AND AGE DIMORPHISM OF MYOCARDIAL GENE EXPRESSION IN NONISCHEMIC HUMAN HEART FAILURE

David R. Fermin, MD[1,2,†], **Ana Barac, MD**[4,†], **Sangjin Lee, MD**[1,2], **Sean P. Polster, BS**[1,2], **Sridhar Hannenhalli, PhD**[5], **Tracy L Bergemann, PhD**[6], **Suzanne Grindle, BS**[1,2], **David B. Dyke, MD**[7], **Francis Pagani, MD, PhD**[7], **Leslie W. Miller, MD**[4], **Sarah Tan, BS**[8], **Cris dos Remedios, PhD, D.Sc.**[8], **Thomas P. Cappola, MD, ScM**[5], **Kenneth B. Margulies, MD**[5], and **Jennifer L. Hall, PhD**[1,2,3,*]

[1] Lillehei Heart Institute, University of Minnesota, Minneapolis, MN, 55455. USA

[2] Division of Cardiology, University of Minnesota, Minneapolis, MN, 55455. USA

[3] Developmental Biology Center, University of Minnesota, Minneapolis, MN, 55455. USA

[4] Washington Hospital Center, Washington DC

[5] University of Pennsylvania, Philadelphia, PA

[6] Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, Minnesota, 55455

[7] University of Michigan, Ann Arbor, MI

[8] Department of Anatomy and Histology University of Sydney Australia

Abstract

BACKGROUND—We report the first comprehensive analysis of gene expression differences by sex and age in left ventricular samples from 102 patients with dilated cardiomyopathy.

METHODS AND RESULTS—Gene expression data (HG-U133A gene chip, Affymetrix) were analyzed from 30 females and 72 males from 3 separate centers. Over 1,800 genes displayed sexual dimorphism in the heart (adjusted p-value <0.05). A significant number of these genes were highly represented in gene ontology pathways involved in ion transport and G-protein-coupled receptor signaling. Localization of these genes revealed enrichment on both the sex chromosomes as well as chromosomes 3, 4, and 14. The second goal of this study was to determine the effect of age on gene expression. Within the female cohort, over 140 genes were differentially expressed in the under 55 age group compared to age group above 55 years of age. These genes were highly represented in gene ontology pathways involved in DNA damage. In contrast, zero genes in the male cohort under age 55 met statistical significance when compared to the group over 55.

CONCLUSIONS—Gene expression in dilated cardiomyopathy displayed evidence of sexual dimorphism similar to other somatic tissues and age dimorphism within the female cohort.

Keywords

Heart failure; Sex; Aging; Genes

[*]Author for correspondence: Jennifer L Hall, Division of Cardiology, Lillehei Heart Institute, 312 Church Street, Minneapolis, MN 55455. Voice: 612 626 4566, Fax: 612 626 4411, jlhall@umn.edu.
[†]These authors contributed equally.

Disclosures: LWM is a consultant for Thoratec Corporation. JH is a consultant with Catholic Health Care West. F. P. received research grant support from Thoratec Corporation and Terumo Corporation.

Cardiovascular diseases are the leading cause of death in both men and women.[1] There are significant differences in the incidence, etiology and outcomes of heart failure in men and women, although the causes of these differences are not well understood. Previous reports have established that women are more likely to have nonischemic heart failure than men, a lower 5-year mortality associated with heart failure, and a phenotype of concentric left ventricular hypertrophy and preserved ejection fraction.[2] Surprisingly, we still do not understand if there are cellular or biological differences between men and women that may be responsible for these differences in function, remodeling, disease, and mortality. Evidence for sex-based differences in myocardial gene expression in humans has been limited by small sample sizes.[3,4] A study testing for sex-based dimorphism in gene expression in the human heart with the power to detect genome-wide significance is a first step toward determining if cellular or biological functions may underlie sex-based differences in phenotype.

In addition to sex, age-related trends in the incidence and prevalence of all cardiovascular diseases have been well established in longitudinal studies.[5] An association between age-related trends in cardiovascular disease and age-dependent changes in the expression of genes has not been tested. Identifying genes whose expression levels are modified by age and associated with cardiovascular phenotypes may lead to the identification of new targets for early detection and prevention of cardiovascular disease.

Thus, the goal of this study was to determine the impact of sex and age on gene expression in the failing human heart. The study population included younger male and female subjects (a cutoff of 55 years of age was used as a proxy for premenopausal status in women), which allowed for identification of additional age-specific gene effects between or amongst the sexes.

METHODS

Statement of Responsibility. The authors had full access to and take full responsibility for the integrity of the data. All authors have read and agree to the manuscript as written.

Human Tissue Samples

Human left ventricular myocardium was obtained from 102 patients with idiopathic cardiomyopathy at the time of heart transplant or left ventricular assist device implantation. The study included 72 males (48.2 +/− 12.7 yrs) and 30 females (50.4 +/− 11.9 yrs) from 3 separate cohorts (Temple University, University of Minnesota, University of Michigan) (Table 1). In the subsequent analysis, the gender groups were also stratified by age < 55 years and >= 55 years. In women, since menopausal status was not known, this age cutoff was meant to serve as an approximation of pre- and post-menopause. Total RNA was then extracted from the samples for microarray hybridization assays as described previously.[6] In addition, a validation cohort was selected from these samples based on the tissue availability for candidate gene reconfirmation by quantitative real-time PCR.

Microarray Analysis

Gene expression arrays from Affymetrix (HG-U133A) were hybridized with the sample RNA as previously described[7] to obtain myocardial gene expression profiles of the 102 subjects. After generation of probe cell intensity data (CEL files), all subsequent statistical analysis was performed using R version 2.5.1[8] with Bioconductor version 2.0 (annotation data from the hgu133a package were built using Entrez gene data as of August 24, 2007 and a Gene Ontology build from February 8, 2007).[9] A number of quality control measures were utilized to assess the quality of each array hybridization as well as institution-to-institution

assay variability (there was no prior randomization or systematic control for variability across institutional laboratories). Quality control metrics included overall intensity distribution plots and Affymetrix 3′/5′ control probe ratios to identify outlier arrays. In addition, positive and negative controls on the periphery of each array were used to generate "center of intensity" plots, which serve as a metric for hybridization uniformity. Each array was evaluated on the basis of all these quality control measures as a whole, and none were found to be consistently out of specification based on commonly accepted metrics for Affymetrix arrays.[10] Utilizing these quality control metrics we did not find bias of hybridization based on the center in which the arrays were processed. Thus, all 102 samples were entered into the analysis and normalized using the robust multiarray average (RMA) method. After normalization, probesets were culled for redundancy if there were multiple probesets per gene (each gene may be represented by multiple unique oligonucleotide probesets on the HG-U133 array). First, all probesets without unique Entrez gene identifiers were removed from the analysis. Next, the median expression of each probeset was calculated, and the probeset with the greatest median expression was selected for each gene, in attempt to best capture the true biological expression while eliminating variance based on noise at low expression levels. This is a variation on probeset selection methods that have been described elsewhere.[11, 12]

Statistical Analysis

Differential expression between the predefined analysis groups (male and female; < 55 and ≥ 55 years old) was modeled using the *limma* linear modeling package in Bioconductor. Contrast matrices were designed as modified analysis of variance (ANOVA) comparisons to detect differences between gender and age subgroups as main effects, as well as the interaction between gender and age effects. After fitting the linear model, the empirical Bayesian method was utilized to calculate test statistics and *P* values. P values were adjusted for multiple testing using the classical false discovery rate (FDR) method described by Benjamini and Hochberg.[13,14,15] This method employed a controlled FDR (set at 0.05) while adjusting for multiple testing simultaneously across multiple subgroup comparisons of age and gender (further multiplying the number of tests adjusted for by the number of comparisons, for a more conservative global adjustment). Resultant gene sets meeting these criteria for differential expression were ascribed genome-wide significance, and were combined with public database annotation data and exported with $\log_2$ transformed expression values and fold changes for further heuristic analysis as described below.

Enrichment Analysis

After genes of interest were identified from the subgroup analyses, a recursive (stepwise grouping) approach was used to reanalyze the gene expression data based on functional classification. This method allows for the elucidation of significant gene groups as opposed to a blind probeset-by-probeset approach, thereby allowing for further discovery based upon gene set "enrichment". Normalized expression data from the subgroup analyses were analyzed using GOstats, a gene set enrichment analysis package in R based on Gene Ontology classifications. This method tests for overrepresentation of functional groups amongst the differentially expressed genes (in this case, the GO:Biological Process Ontology).[16] Testing for enrichment of individual chromosomes based upon the abundance of differentially expressed genes amongst those represented on the HG-U133A (for those that had chromosomal annotation data) was performed using the Fisher exact test.[17]

Transcription factor binding sites in promoters of candidate genes

We first converted the ENSEMBL gene identifiers to RefSeq identifiers using ENSEMBL's biomart tool. Using a previously published, phylogenetic-footprinting approach,[7] we compiled a comprehensive database of putative transcription factor binding sites in 1kb

proximal promoters of all human genes with full length transcripts. Briefly, for each gene, we extracted 1 kb of genomic sequence immediately upstream from the transcription start site from UCSC database (www.genome.ucsc.edu). We searched these promoter regions using the 584 transcription factor binding site motifs obtained from the TRANSFAC database v10.2.[18] A binding site motif is represented as a "Positional Weight Matrix" (PWM), which is a $4 \times k$ matrix for a k bases long binding site and provides, for each of the k positions, the preferences for the four nucleotide bases at that position. Matches between TRANSFAC PWMs and promoter regions of cardiac genes were determined using the tool PWMSCAN.[19] The criterion for a match was p-value cutoff of $2 \cdot 10^{-4}$, corresponding to an expected frequency of 1 random match in 5kb. We filtered these matches further using human-mouse genome sequence alignments to focus on promoter regions that showed evolutionary conservation. For each TRANSFAC match, let c be the fraction of binding site bases that were identical between human and mouse. We retained matches such that either p-value ≤ 0.00002 (expected frequency of 1 in 50 kb) or $c \geq 0.8$. These criteria for matching have been evaluated previously and were shown to accurately detect ~65% of experimentally verified binding sites with a low false positive rate of 1 random match in every 50KB of genomic sequence searched.[19]

Estimating Over-representation transcription factor binding sites in a gene set

Given two sets of gene promoters – a foreground F (with N_F promoters) and a control or background B (with N_B promoters), our goal is to identify motifs that are enriched in F relative to B. For a particular transcription factor binding motif M, let M_F be the number of foreground promoters that have site for motif M. M_B is defined analogously. The quantity $\frac{M_F \bullet N_B}{M_B \bullet N_F}$ represents the enrichment of motif M in F relative to B. We estimated the significance (P value) of the enrichment using a one tailed Fisher exact test. Criteria for calling a family of TF binding motifs enriched in the promoters of differentially expressed genes were based on John Storey's false discovery rates, computed in R using the p-values as input. In order for a family of TF binding motifs to be selected as enriched, the median q-value for PWMs in that family must be < 0.10. **Male vs. Background:** In this experiment male-specific differentially expressed genes were used as the foreground and the non-differentially expressed genes were used as the background. **Female vs. Background:** Female-specific differentially expressed genes were used as the foreground and the non-differentially expressed genes were used as the background. We then compare the result from these two analyses to search for TFs specifically associated with heart failure in men or women with idiopathic dilated cardiomyopathy (IDCM).

Quantitative Real-Time PCR

Expression levels of candidate genes were reproduced using quantitative real-time PCR. Eleven male and 5 female subjects were selected from the initial array analysis. Four candidate genes were selected for validation based on the mean Affymetrix expression levels as well as based on their potential role in the pathogenesis of heart failure. Tropomyosin 3 (TPM3) and myosin light chain 4 (MLK4) encode proteins implicated in the regulation of cardiac contractility.[20,21] while nitric oxide synthase 1 (NOS1), localized in the sarcoplasmic reticulum of the cardiomyocytes, may modulate calcium handling via its association with ryanodine receptor.[22] Phosphorylase kinase B (PHKB) is involved in the breakdown of glycogen in the heart.

Primer and probe sets were obtained from optimized TaqMan assays (Applied Biosystems, Foster City, CA) designed for the ABI 7900HT Fast Real-Time PCR System. Target amplification and detection were performed on samples and controls in the same thermal cycling reaction in replicated fashion, allowing for minimization of experimental variability

and calculation of ΔCt based on the corresponding reference control, hypoxanthine guanine phosphoribosyl transferase, (i.e. Target$^{\Delta CtHPRT}$).[23] All samples were analyzed in duplicate. The results are shown in Supplemental Figure 1. There was a significant correlation between array and RT-PCR relative values for all of the genes except NOS1 which also had lower expression levels in Affymetrix array (Supplemental Figure 1).

RESULTS

Patient Characteristics

The clinical information for the subjects is included in Table 1. The age distribution, ejection fractions and proportion of inotrope use were not significantly different between the male and female cohorts. Similarly, the male/female distribution, ejection fractions and proportion of inotrope use were not significantly different between the under age 55 vs. the over age 55 cohort.

Sex and Age-Based Differential Expression

The aim of the analysis was to identify the impact of sex and age on gene expression patterns in dilated cardiomyopathy. Using a false discovery rate as described by Benjamini and Hochberg, with a significance level set at 0.05, we found 1837 genes that were differentially expressed between the male and female cohort (for the list of genes see Supplemental Table 1).

Table 2 summarizes the distribution of genes that exhibited differential expression between men and women, stratified by age, at various stringency levels (absolute fold change >1.0, >1.2, >2.0, adjusted *P* value < 0.05). These genes were further characterized as female-biased or male-biased, based on the predominant expression in females or males respectively. As shown in Table 2 the majority of differences in gene expression are modest in size (<1.2 fold), which is consistent with previously published data.[24] Among the differentially expressed genes, there were a significantly larger number of female-biased genes compared to male-biased genes (the greater abundance of >2.0 fold expression in males is largely attributable to Y-linked genes). The bottom portion of Table 2 provides a breakdown of differential expression by age within the entire cohort as well as within the male and female cohorts. In this comparison, there is a clear delineation between age based expression in males and females. Within the male cohort, zero genes met the statistical cutoff for differential expression between the younger and older cohort. In contrast, in the female cohort, 142 genes were differentially expressed in the under 55 cohort compared to the over 55 group.

The overall comparison to detect interaction between age effects and sex effects did not yield any significant genes when using the specified parameters for genome-wide significance, though there were a number of genes with significant unadjusted *P* values (not shown). The age* sex interaction contrast tests for additional deviation in expression ratios that can not be explained solely by the sum of age and sex differences.

Enrichment Analysis—Evidence for chromosomal enrichment is shown in Figure 1. Using the Fisher exact test to compare the frequencies of significant genes on each chromosome against the total number of genes interrogated on each chromosome, chromosomal enrichment in the male versus female comparison was detected. Aside from the sex chromosomes, which segregated as expected, chromosome 4 was found to have more female-biased differential expression, while chromosomes 3 and 14 contained more male-biased genes (Figure 1).

To better understand the functional categories and the networks represented by the genes whose expression was significantly different between males and females we used a program based on Gene Ontology biological processes (the GOstats package in Bioconductor for R). Similar categorical groupings were observed when analyzed using another public domain enrichment analysis tool in the DAVID suite as a replication exercise (data not shown).[17]

Table 3A lists the main functional biological categories of the genes that were significantly different between men and women using Gene Ontology. Both ion transport and G-protein-coupled receptor signaling pathways are included at the P≤0.001 significance level. These two broad categories included some of the genes known to mediate physiological processes in the heart such as G-protein receptor kinases (GRK2 and GRK6), adenosine receptor, chemokine receptor 5 and phosphoinositide-3-kinase as well as sodium and potassium channels (for a full list of the genes within this pathway and all other significant pathways please see Supplemental Table 2).

In the female over- vs. under 55 years comparison, only the DNA damage response pathway met the statistical significance level of P<0.01 (Table 3B and Supplemental Table 2).

Transcription Factor Binding Site Analysis

For the lists of male-biased and female-biased genes, we compared the frequency of transcription factor bindings sites within the 1 KB promoter regions to the frequencies observed in a background set of 5000 genes which showed no evidence of differential expression (see data supplement for complete lists of male-biased, female biased, and background set). We found that 51 TRANSFAC positional weight matrices (PWMs) were enriched in the promoters of female-biased genes with a false-discovery rate of 0.10, whereas only 2 PWMs were enriched in male promoters (both belong to the E2F family) (Supplementary Table 3). These PWMs correspond to the transcription factor families summarized in Table 4. Several transcription factor families previously associated with the pathogenesis of heart failure were identified, including MEF2, GATA, NKX, and FOX.

DISCUSSION

Surprisingly little is known regarding the underlying mechanisms behind sex and age related differences in cardiovascular disease. To begin to explore this, we analyzed gene expression profiles in a cohort of 70 men and 32 women with dilated cardiomyopathy. We identified 1837 differentially expressed genes between the sexes. The majority of these genes demonstrated modest absolute fold changes of 1.2–2.0 (74.9%). These sexually dimorphic genes showed enrichment on autosomes 3, 4 and 14 as well as the sex chromosomes. A breakdown of these genes into biological processes revealed an overrepresentation of genes involved in ion transport activity and G-protein-coupled receptor signaling pathways. Age-related differences in gene expression were identified in the female cohort (less than 55 years of age vs. older than 55 years of age). These genes were highly represented in gene ontology pathways involved in DNA damage and signal transduction. In contrast, zero genes in the male cohort under age 55 met statistical significance when compared to the group over 55, despite the larger sample size of the male cohort. In conclusion, gene expression in the heart of patients with idiopathic cardiomyopathy displayed evidence of sexual dimorphism similar to other somatic tissues and age dimorphism within the female cohort.

Our findings are in agreement with a recent report on sexually dimorphic genes present in other somatic tissues in mice, such as liver, adipose tissue, muscle and brain, where >70% of differentially expressed genes were found to display <1.2-fold difference in expression between males and females.[25] Greater fold changes (>2.0 fold) were seen in sex-chromosome linked genes as would be expected. (Supplemental Table 4). A recent study on

a smaller set of human cardiac samples obtained from healthy organ donors reported 16 sexually dimorphic genes that displayed >2-fold differences and nearly all were linked to sex chromosomes.[3] Six of these sexually dimorphic genes identified in the hearts of healthy donors were identified in the failing hearts in our cohort (EIF1AY, RPS4Y1, DDX3Y, JARID1D, USP9Y, CYorf14) (Supplemental Table 4). JARID1D is a member of the Jumonji family. Several members of this Jumonji family were recently identified to act as transcriptional repressors and modulators of chromatin.[26–30] JARDID1D and JARID1C have been identified as histone demethylases.[26] Thus JARID1D may be an important sex-linked gene that regulates sex-linked changes in gene expression through transcriptional repression or modification of chromatin leading to ultrastructural and functional changes in the heart. Early work showed that mice with a homozygous knockout of the jumonji gene showed abnormal heart development.[31] Jumonji has also been shown to repress the transcriptional activities of GATA4 and Nkx2.5.[31] Taken together, these findings suggest that X-linked expression of JARID1D may be a critical differentiator underlying sex-related differences in cardiovascular function. X and Y linked genes with a priori biological evidence for a role in genetic, molecular, and cellular processes linked to heart structure and function are likely first candidates for targets that lead to differences between men and women.

To increase our understanding of the functional categories highly represented by the genes whose expression was significantly different between men and women, we used an integrative analysis based on GO biological process terms. Two of the major findings of this study are that sexually dimorphic genes are significantly overrepresented in ion transport and G protein coupled receptor signaling pathways.

The sexually dimorphic genes encoding sodium channels such as SCN3B, SCN8A, SCN10A, SCN11A and SCN3, as well as potassium channels (KCNA2, KCNB2, KCNJ3, KCNJ6, KCNJ14, KCNMB3, KCNMB4, KCNV1) clustered in the ion transport categories. Pre-clinical and clinical data suggest a possible role of ion channels in sex specific cardiovascular phenotypes. For example, female predominance was found in both the acquired and congenital long-QT syndromes.[32] In addition, a recent report by Albert et al. identified an association between functionally significant mutations and rare polymorphisms in SCN5A and sudden cardiac risk in women.[33] *In vitro* work with human mammary arteries showed that arterial vasorelaxation in response to levosimendan, a calcium sensitizer and a novel vasodilatory agent that is used in the therapy of heart failure, was also quantitatively and qualitatively different in males and females.[34] In our analysis, we found that men and women with end-stage heart failure had significantly different expression levels of RNA encoding for sodium and potassium channels. In addition, we identified significant overrepresentation of the transcription factor binding site Nkx2.5 in the female biased genes. Mutations in Nkx2.5 result in conduction abnormalities.[35] Furthermore, many sodium and potassium channel family members harbor Nkx2.5 sites in their promoters. Future work will be needed to explore the biological regulation of these ion channel families and the role and regulation of Nkx.2.5 in male and female pre-clinical models.

Several transcription factors in addition to Nkx2.5 were identified that were overrepresented in the differentially expressed genes in the female hearts, including factors such as MEF2 and GATA that have been previously associated with heart failure along with novel candidates. Unexpectedly, we detected enrichment of the transcription factor binding site that recognizes Sry (a gene found on the Y chromosome) in the genes differentially expressed in females. Since SRY is only expressed in men, this finding cannot imply an increase in SRY activity in the female failing heart. Instead, it may be a marker of target-genes that are regulated in a different manner in women compared to men via other mechanisms. Alternatively, this finding could be a false-positive association, or it may it

may reflect limitations in the specificity of TF binding motifs in TRANSFAC. For example, the binding motif for Sry (**AAACAAA**) is similar to that for So x-9 (GT**AAACAAT**AGA) and a few FOXO transcription factors (TATGT**AAACAAA**CAA). The overexpression of transcription factor binding sites for GATA and MEF are particularly intriguing given their known role in hypertrophy and cardiac remodeling.[31, 36] The only transcription factor binding site overrepresented in the genes up-regulated in the male hearts was E2F. This transcription factor binding site is present in many genes that control cell cycle, gene transcription, and cell death. Future work in the laboratory will be needed to test the mechanistic role of these transcription factors.

The second major pathway identified from a pathways analysis of statistically significant sexually dimorphic genes was G protein coupled receptor signaling. The dimorphic genes from the GPCR functional group included several targets with previously demonstrated roles in ventricular remodeling and function; G-protein receptor kinases (GRK2, GRK6) and phosphoinositide-3-kinase have been reported as critical regulators of inotropic and lusitropic properties of the heart,[37] and nonspecific ablation of the GRK2 (also known as beta adrenergic receptor kinase 1) gene was found to cause intrauterine heart failure and death in mice.[38] Importantly, increasing evidence for sex-related differences in ventricular remodeling is emerging from different animal models.[39–41] Our data provides the basis to further explore the relationship between the observed dimorphism in the signaling pathways and sex-related differences in discrete cardiac phenotypes, such as response to pressure and volume overload or remodeling after myocardial infarction.[41]

When divided into age groups below and above 55 years of age, zero genes were differentially expressed by age in the male cohort. In contrast, when the female cohort was analyzed separately, over 142 genes were significantly different in the age group below 55 yrs compared to above 55 yrs. These genes revealed only one significantly overrepresented functional category, the DNA damage response pathway, with 3 representative genes including v-abl Abelson murine leukemia viral oncogene homolog 1, GPI anchored molecule like protein and BRCA1 interacting protein C-terminal helicase 1.

Perspective and Study Limitations

The elucidation of sex-related differences in prognosis and outcomes in chronic heart failure (HF) has been limited by the smaller proportion of women enrolled in clinical trials, thus leading to inconsistent reports about the relationship between sex and survival in heart failure.[42, 43,44,45] The investigators of Candesartan in Heart Failure: Assessment of Reduction in Mortality and Morbidity (CHARM) Program examined the effect of sex on clinical outcomes in 5199 men and 2400 women with HF and found that women had lower risk of both fatal and nonfatal cardiovascular outcomes that could not be explained by LVEF or origin of the heart failure,[46] thus suggesting that an alternative explanation should be sought for the observed differential survival. In addition to increasing the number of women in clinical trials and the number of female animals in pre-clinical studies, we must also continue to test for potential differences at the level of the chromosome as well as mRNA and protein to help define potential mechanisms for differences in clinical trial outcome. Surprisingly little is known about the role of biology in sex and age related differences in cardiovascular disease.[47] Uncovering gene expression differences between men and women, and in response to age, will provide new steps towards understanding the interacting cause of disease, prevention of disease, and treatment of disease.

Microarray analyses have been utilized to identify new candidate genes potentially implicated in cardiovascular disease.[4,7] The results of these studies, including our own, suggest that many factors must be considered when examining gene expression in human tissue including different cell types within the tissue, social, diet, and other environmental

factors, and the interaction of other on-going biological events in the body including but not limited to inflammation, and metabolic disease. Boheler et al identified differences in gene expression based on age and sex in failing and nonfailing hearts in 2003.[4] However, these studies have all been limited in sample size. The analysis presented here represents the largest human cohort analyzed to date for sex and age-related differences in cardiac gene expression. However, in order to test or confirm these findings larger better curated databases will need to be developed. These types of analyses will likely be important moving forward as sex linked cardiovascular traits are mapped to loci.[33] A fair amount of heterogeneity existed in the clinical characteristics of the patients. In addition, other pertinent data including all cardiac medications, exact menopausal status of female subjects and use of hormone replacement therapy were not readily available for most subjects. The use of inotropes was similar in all groups, thus eliminating many of the known clinical confounders that could be associated with changes in gene expression. However, any of these factors may influence the gene expression profile. Furthermore, such heterogeneity along with a modest sample size (though relatively large for human microarray studies) does not allow for more discrete analysis by variables, such as investigating age as a continuous variable with respect to differential gene expression. In addition to the limitations, it should also be pointed out that an analysis of this scope does permit a robust examination of cell specific markers. For example, we identified unique leukocyte CD (cluster of differentiation) markers for both T and NK cells present in these hearts (CD3, 896, 160, 172g, and 247).

Conclusions

In conclusion, we have found substantial evidence for sex and age dimorphisms of cardiac gene expression in nonischemic human heart failure by performing a comprehensive study in a large number of patients. A long term goal is to link chromosomal information with gene expression and intermediate phenotypic endpoints in the heart as well as within different cell types within the heart including cardiac myocytes, fibroblasts, vascular smooth muscle cells and endothelial cells. This will require further observational and clinical trials that will be able to clearly distinguish and identify phenotypes, as well as further cellular and molecular studies that will identify tissue- and disease-specific functions of the pathways involved.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Sources: These studies were supported by the National Institutes of Health, Bethesda, MD (AG17022 to K.B.M.and HL092379 to T.P.C and S.H.) and philanthropic funds from the Lillehei family (JH).

References

1. Lloyd-Jones DM, Larson MG, Beiser A, Levy D. Lifetime risk of developing coronary heart disease. Lancet. 1999; 353:89–92. [PubMed: 10023892]
2. Levy D, Kenchaiah S, Larson MG, Benjamin EJ, Kupka MJ, Ho KK, Murabito JM, Vasan RS. Long-term trends in the incidence of and survival with heart failure. N Engl J Med. 2002; 347:1397–1402. [PubMed: 12409541]
3. Isensee J, Witt H, Pregla R, Hetzer R, Regitz-Zagrosek V, Ruiz Noppinger P. Sexually dimorphic gene expression in the heart of mice and men. J Mol Med. 2008; 86:61–74. [PubMed: 17646949]

4. Boheler KR, Volkova M, Morrell C, Garg R, Zhu Y, Margulies K, Seymour AM, Lakatta EG. Sex- and age-dependent human transcriptome variability: implications for chronic heart failure. Proc Natl Acad Sci U S A. 2003; 100:2754–2759. [PubMed: 12601168]
5. Cupples, L.; D'Agostino, R. Section 34: some risk factors related to the annual incidence of cardiovascular disease and death in pooled repeated biennial measurements. In: Kannel, W.; Wolf, P.; Garrison, R., editors. Series Section 34: some risk factors related to the annual incidence of cardiovascular disease and death in pooled repeated biennial measurements. Bethesda, MD: US Department of Health and Human Services; 1987.
6. Hall JL, Grindle S, Han X, Fermin D, Park S, Chen Y, Bache RJ, Mariash A, Guan Z, Ormaza S, Thompson J, Graziano J, de Sam Lazaro SE, Pan S, Simari RD, Miller LW. Genomic profiling of the human heart before and after mechanical support with a ventricular assist device reveals alterations in vascular signaling networks. Physiol Genomics. 2004; 17:283–291. [PubMed: 14872006]
7. Hannenhalli S, Putt ME, Gilmore JM, Wang J, Parmacek MS, Epstein JA, Morrisey EE, Margulies KB, Cappola TP. Transcriptional genomics associates FOX transcription factors with human heart failure. Circulation. 2006; 114:1269–1276. [PubMed: 16952980]
8. R: A Language and Environment for Statistical Computing [computer program]. Version. Vienna, Austria: 2007.
9. Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. 2004; 5:R80. [PubMed: 15461798]
10. Parman, C.; Halling, C.; Gentleman, R. affyQCReport: QC Report Generation for affyBatch objects.
11. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A. 2005; 102:15545–15550. [PubMed: 16199517]
12. Jiang, Z.; Gentleman, R. Extensions to gene set enrichment. Bioconductor Project working Papers. Working paper 12. http://www.bepress.com/bioconductor/paper12
13. Hochberg Y, Benjamini Y. More powerful procedures for multiple significance testing. Stat Med. 1990; 9:811–818. [PubMed: 2218183]
14. Klipper-Aurbach Y, Wasserman M, Braunspiegel-Weintrob N, Borstein D, Peleg S, Assa S, Karp M, Benjamini Y, Hochberg Y, Laron Z. Mathematical formulae for the prediction of the residual beta cell function during the first two years of disease in children and adolescents with insulin-dependent diabetes mellitus. Med Hypotheses. 1995; 45:486–490. [PubMed: 8748093]
15. Reiner A, Yekutieli D, Benjamini Y. Identifying differentially expressed genes using false discovery rate controlling procedures. Bioinformatics. 2003; 19:368–375. [PubMed: 12584122]
16. Falcon S, Gentleman R. Using GOstats to test gene lists for GO term association. Bioinformatics. 2007; 23:257–258. [PubMed: 17098774]
17. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol. 2003; 4:P3. [PubMed: 12734009]
18. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, Reuter I, Chekmenev D, Krull M, Hornischer K, Voss N, Stegmaier P, Lewicki-Potapov B, Saxel H, Kel AE, Wingender E. TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res. 2006; 34:D108–110. [PubMed: 16381825]
19. Levy S, Hannenhalli S. Identification of transcription factor binding sites in the human genome sequence. Mamm Genome. 2002; 13:510–514. [PubMed: 12370781]
20. Pieples K, Arteaga G, Solaro RJ, Grupp I, Lorenz JN, Boivin GP, Jagatheesan G, Labitzke E, DeTombe PP, Konhilas JP, Irving TC, Wieczorek DF. Tropomyosin 3 expression leads to hypercontractility and attenuates myofilament length-dependent Ca(2+) activation. Am J Physiol Heart Circ Physiol. 2002; 283:H1344–1353. [PubMed: 12234784]

21. Olsson MC, Patel JR, Fitzsimons DP, Walker JW, Moss RL. Basal myosin light chain phosphorylation is a determinant of Ca2+ sensitivity of force and activation dependence of the kinetics of myocardial force development. Am J Physiol Heart Circ Physiol. 2004; 287:H2712–2718. [PubMed: 15331360]
22. Liu VW, Huang PL. Cardiovascular roles of nitric oxide: a review of insights from nitric oxide synthase gene disrupted mice. Cardiovasc Res. 2008; 77:19–29. [PubMed: 17658499]
23. Fleige S, Walf V, Huch S, Prgomet C, Sehm J, Pfaffl MW. Comparison of relative mRNA quantification models and the impact of RNA integrity in quantitative real-time RT-PCR. Biotechnol Lett. 2006; 28:1601–1613. [PubMed: 16900335]
24. Margulies KB, Matiwala S, Cornejo C, Olsen H, Craven WA, Bednarik D. Mixed messages: transcription patterns in failing and recovering human myocardium. Circ Res. 2005; 96:592–599. [PubMed: 15718504]
25. Yang X, Schadt EE, Wang S, Wang H, Arnold AP, Ingram-Drake L, Drake TA, Lusis AJ. Tissue-specific expression and regulation of sexually dimorphic genes in mice. Genome Res. 2006; 16:995–1004. [PubMed: 16825664]
26. Iwase S, Lan F, Bayliss P, de la Torre-Ubieta L, Huarte M, Qi HH, Whetstine JR, Bonni A, Roberts TM, Shi Y. The X-linked mental retardation gene SMCX/JARID1C defines a family of histone H3 lysine 4 demethylases. Cell. 2007; 128:1077–1088. [PubMed: 17320160]
27. Metzger E, Yin N, Wissmann M, Kunowska N, Fischer K, Friedrichs N, Patnaik D, Higgins JM, Potier N, Scheidtmann KH, Buettner R, Schule R. Phosphorylation of histone H3 at threonine 11 establishes a novel chromatin mark for transcriptional regulation. Nat Cell Biol. 2008; 10:53–60. [PubMed: 18066052]
28. Hong S, Cho YW, Yu LR, Yu H, Veenstra TD, Ge K. Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases. Proc Natl Acad Sci U S A. 2007; 104:18439–18444. [PubMed: 18003914]
29. Chang B, Chen Y, Zhao Y, Bruick RK. JMJD6 is a histone arginine demethylase. Science. 2007; 318:444–447. [PubMed: 17947579]
30. Takeuchi T, Watanabe Y, Takano-Shimizu T, Kondo S. Roles of jumonji and jumonji family genes in chromatin regulation and development. Dev Dyn. 2006; 235:2449–2459. [PubMed: 16715513]
31. Kim Y, Phan D, van Rooij E, Wang DZ, McAnally J, Qi X, Richardson JA, Hill JA, Bassel-Duby R, Olson EN. The MEF2D transcription factor mediates stress-dependent cardiac remodeling in mice. J Clin Invest. 2008; 118:124–132. [PubMed: 18079970]
32. Imboden M, Swan H, Denjoy I, Van Langen IM, Latinen-Forsblom PJ, Napolitano C, Fressart V, Breithardt G, Berthet M, Priori S, Hainque B, Wilde AA, Schulze-Bahr E, Feingold J, Guicheney P. Female predominance and transmission distortion in the long-QT syndrome. N Engl J Med. 2006; 355:2744–2751. [PubMed: 17192539]
33. Albert CM, Nam EG, Rimm EB, Jin HW, Hajjar RJ, Hunter DJ, MacRae CA, Ellinor PT. Cardiac sodium channel gene variants and sudden cardiac death in women. Circulation. 2008; 117:16–23. [PubMed: 18071069]
34. Akar F, Manavbasi Y, Parlar AI, Ulus AT, Katircioglu SF. The gender differences in the relaxation to levosimendan of human internal mammary artery. Cardiovasc Drugs Ther. 2007; 21:331–338. [PubMed: 17703356]
35. Dupays L, Jarry-Guichard T, Mazurais D, Calmels T, Izumo S, Gros D, Theveniau-Ruissy M. Dysregulation of connexins and inactivation of NFATc1 in the cardiovascular system of Nkx2-5 null mutants. J Mol Cell Cardiol. 2005; 38:787–798. [PubMed: 15850572]
36. Hasegawa K, Lee SJ, Jobe SM, Markham BE, Kitsis RN. cis-Acting sequences that mediate induction of beta-myosin heavy chain gene expression during left ventricular hypertrophy due to aortic constriction. Circulation. 1997; 96:3943–3953. [PubMed: 9403619]
37. Matkovich SJ, Diwan A, Klanke JL, Hammer DJ, Marreez Y, Odley AM, Brunskill EW, Koch WJ, Schwartz RJ, Dorn GW 2nd. Cardiac-specific ablation of G-protein receptor kinase 2 redefines its roles in heart development and beta-adrenergic signaling. Circ Res. 2006; 99:996–1003. [PubMed: 17008600]

38. Jaber M, Koch WJ, Rockman H, Smith B, Bond RA, Sulik KK, Ross J Jr, Lefkowitz RJ, Caron MG, Giros B. Essential role of beta-adrenergic receptor kinase 1 in cardiac development and function. Proc Natl Acad Sci U S A. 1996; 93:12974–12979. [PubMed: 8917529]

39. Douglas PS, Katz SE, Weinberg EO, Chen MH, Bishop SP, Lorell BH. Hypertrophic remodeling: gender differences in the early response to left ventricular pressure overload. J Am Coll Cardiol. 1998; 32:1118–1125. [PubMed: 9768741]

40. Weinberg EO, Thienelt CD, Katz SE, Bartunek J, Tajima M, Rohrbach S, Douglas PS, Lorell BH. Gender differences in molecular remodeling in pressure overload hypertrophy. J Am Coll Cardiol. 1999; 34:264–273. [PubMed: 10400020]

41. Deschepper CF, Llamas B. Hypertensive cardiac remodeling in males and females: from the bench to the bedside. Hypertension. 2007; 49:401–407. [PubMed: 17242305]

42. Adams KF Jr, Sueta CA, Gheorghiade M, O'Connor CM, Schwartz TA, Koch GG, Uretsky B, Swedberg K, McKenna W, Soler-Soler J, Califf RM. Gender differences in survival in advanced heart failure. Insights from the FIRST study. Circulation. 1999; 99:1816–1821. [PubMed: 10199877]

43. Simon T, Mary-Krause M, Funck-Brentano C, Jaillon P. Sex differences in the prognosis of congestive heart failure: results from the Cardiac Insufficiency Bisoprolol Study (CIBIS II). Circulation. 2001; 103:375–380. [PubMed: 11157688]

44. Ghali JK, Pina IL, Gottlieb SS, Deedwania PC, Wikstrand JC. Metoprolol CR/XL in female patients with heart failure: analysis of the experience in Metoprolol Extended-Release Randomized Intervention Trial in Heart Failure (MERIT-HF). Circulation. 2002; 105:1585–1591. [PubMed: 11927527]

45. Ghali JK, Krause-Steinrauf HJ, Adams KF, Khan SS, Rosenberg YD, Yancy CW, Young JB, Goldman S, Peberdy MA, Lindenfeld J. Gender differences in advanced heart failure: insights from the BEST study. J Am Coll Cardiol. 2003; 42:2128–2134. [PubMed: 14680739]

46. O'Meara E, Clayton T, McEntegart MB, McMurray JJ, Pina IL, Granger CB, Ostergren J, Michelson EL, Solomon SD, Pocock S, Yusuf S, Swedberg K, Pfeffer MA. Sex differences in clinical characteristics and prognosis in a broad spectrum of patients with heart failure: results of the Candesartan in Heart failure: Assessment of Reduction in Mortality and morbidity (CHARM) program. Circulation. 2007; 115:3111–3120. [PubMed: 17562950]

47. Denton K, Baylis C. Physiological and molecular mechanisms governing sexual dimorphism of kidney, cardiac, and vascular function. Am J Physiol Regul Integr Comp Physiol. 2007; 292:R697–699. [PubMed: 17095645]

Figure 1.

Bar graphs illustrate the chromosomal distribution of differentially expressed genes in the whole male versus female cohort; lighter bars represent the significant genes on each chromosome; the number of unique genes assayed on each chromosome is represented by the total height of each bar. The table below the graph lists chromosomes that passed the test for enrichment by Fisher exact test with Benjamini correction, $P < 0.05$.

Table 1

Clinical data summary for patients for whom these data were available. P values are calculated from Fisher exact test comparing male vs. female within age category (row) or comparing under or over age 55 within male or female cohort (column).

	Male	Female	*P* value
Age, years	48.2 ± 13.1 (n=72)	50.4 ± 11.9 (n=30)	
< 55 years	41.6 ± 10.4 (n=49)	44.7 ± 10.2 (n=20)	
≥ 55 years	62.1 ± 4.6 (n=23)	61.9 ± 3.84 (n=10)	
EF, %	12.1 ± 5.6 (n = 66)	12 ± 5.5 (n = 30)	
< 55 years	11.1 ± 4.9 (n = 43)	12.9 ± 6.3 (n = 20)	
≥ 55 years	14 ± 6.4 (n = 23)	10.3 ± 3.4 (n = 10)	
P value			
Inotropes, n (%)	60 (90.9%) (n = 66)	26 (86.7%) (n = 30)	0.53
< 55 years	39 (90.7%)(n = 43)	16 (80%)(n = 20)	0.24
≥ 55 years	21 (91.3%)(n =23)	10 (100%)(n = 10)	0.35
P value	0.94	0.14	

Table 2

Age and sex stratified comparisons of differentially expressed genes (adjusted *P* value < 0.05). The top half of the table shows male and female biased genes, defined by the sex in which the gene is more highly expressed. The bottom half of the table shows age dependent expression, also stratified by sex. The significant genes are further categorized by fold change.

	Absolute Fold Change	Differentially Expressed Genes (adjusted *P* < 0.05)	Female-Biased Genes	Male-Biased Genes
Sexually Dimorphic (M vs. F)				
All (n =102)	> 1.0	1837	1272	565
	> 1.2	1377	842	535
	> 2.0	11	2	9
< 55 y (n = 69)	> 1.0	12	3	9
	> 1.2	11	2	9
	> 2.0	6	1	5
≥ 55 y (n = 33)	> 1.0	728	572	156
	> 1.2	381	244	137
	> 2.0	5	1	4
Age dependent (within gender)				
All (n =102)	> 1.0	66		
	> 1.2	60		
	> 2.0	1		
Male (n = 72)	> 1.0	0		
	> 1.2	0		
	> 2.0	0		
Female (n = 30)	> 1.0	142		
	> 1.2	91		
	> 2.0	1		

Table 3

Gene Ontology—Biological Process classification groupings overrepresented in A) Male versus female subjects, and B) Female, Over versus Under 55 years old ($P \leq 0.01$, group size ≥ 5). Included are raw P values, odds ratio of group overrepresentation, expected (Exp. Count) versus actual (Count) number of significant genes, and total genes (Size) in the enriched categories. Biological processes that are labeled as "sub-processes" of a Functional Category are indented underneath the Category.

A) All male versus female subjects

Gene Ontology ID	P	Odds Ratio	Exp. Count	Count	Size	GO: Biological Process Functional Category Description
GO:0006811	4.50E-04	1.473	76	103	563	ion transport
GO:0006812	1.26E-04	1.663	49	73	361	cation transport
GO:0015672	7.71E-03	2.126	31	57	233	monovalent inorganic cation transport
GO:0030001	1.81E-02	1.690	40	61	297	metal ion transport
GO:0006814	6.86E-03	2.577	12	25	88	sodium ion transport
GO:0006813	7.10E-04	1.798	15	25	115	potassium ion transport
GO:0006821	3.71E-05	2.582	6	12	42	chloride transport
GO:0050877	1.23E-03	1.346	81	102	598	neurological process
GO:0007600	3.98E-03	1.626	47	70	352	sensory perception
GO:0007606	5.71E-03	3.321	9	23	68	sensory perception of chemical stimulus
GO:0007608	3.55E-04	2.975	7	17	54	sensory perception of smell
GO:0050909	4.21E-04	5.524	2	6	13	sensory perception of taste
GO:0007601	3.67E-03	1.719	23	36	172	visual perception
GO:0007186	4.50E-04	1.496	73	100	540	G-protein coupled receptor protein signaling pathway
GO:0007565	2.02E-06	2.843	8	18	59	female pregnancy
GO:0042430	1.26E-04	12.877	1	4	6	indole and derivative metabolic process

B) Female, over versus under 55 years old

Gene Ontology ID	P	Odds Ratio	Exp. Count	Count	Size	Functional Category Description
GO:0042770	7.00E-03	8.498	0	3	38	DNA damage response, signal transduction

Table 4

Transcription factor families with significant binding site enrichment in promoters of male- and female-biased genes.

Group	TF family	Median Enrichment Score	Median q-value
Female-biased	TTF1	1.57	0.005
	PBX	1.83	0.005
	SRY	1.64	0.007
	HP1	1.78	0.008
	GATA	1.64	0.014
	Cart-1	1.80	0.019
	HOXA4	1.63	0.022
	APOLYA	1.54	0.023
	HNF	1.55	0.023
	POU	1.64	0.033
	NKX	1.83	0.035
	IPF1	1.57	0.036
	Alx-4	1.59	0.036
	OCT	1.53	0.041
	LEF1	1.51	0.050
	MEF2	1.48	0.058
	FOX	1.52	0.068
	CDX	1.59	0.070
	TFIIA	1.25	0.077
	C/EBP	1.36	0.081
	TATA	1.33	0.085
	BLIMP1	1.29	0.088
	ZNFN1A	1.42	0.091
	PLZF	1.57	0.091
	Pit-1	1.56	0.096
Male-biased	E2F	1.23	0.0008