Identification of Candidate Genes for Histiocytoid Cardiomyopathy (HC) Using Whole Genome Expression Analysis: Analyzing Material from the HC Registry

Bahig Shehata, Emory University
Mark Bouzyk, Emory University
Weining Tang, Emory University
Charlotte K. Steelman, Childrens Healthcare Atlanta
Carlos S Moreno, Emory University

Journal Title: Pediatric and Developmental Pathology
Volume: Volume 14, Number 5
Publisher: SAGE Publications (UK and US) | 2011-09-01, Pages 370-377
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.2350/10-05-0826-OA.1
Permanent URL: https://pid.emory.edu/ark:/25593/trvs3

Final published version: http://dx.doi.org/10.2350/10-05-0826-OA.1

Copyright information:
© 2011 Society for Pediatric Pathology.
Accessed August 16, 2019 11:52 PM EDT
Identification of Candidate Genes for Histiocytoid Cardiomyopathy (HC) Using Whole Genome Expression Analysis: Analyzing Material from the HC Registry

Bahig M. Shehata, M.D.1,2, Mark Bouzyk, Ph.D.3, Weining Tang, Ph.D.3, Charlotte K. Steelman, B.S.1, and Carlos S. Moreno, Ph.D.2,3

1Department of Pathology, Children’s Healthcare of Atlanta, Atlanta, GA 30322
2Department of Pathology & Laboratory Medicine, Emory University School of Medicine, Atlanta, GA 30322
3Winship Cancer Institute, Emory University School of Medicine, Atlanta, GA 30322

Abstract

BACKGROUND—Histiocytoid cardiomyopathy (HC) is a rare but distinctive arrhythmogenic disorder characterized by incessant ventricular tachycardia, cardiomegaly, and often sudden death by age 2 years. The underlying genetic mechanism of HC has eluded researchers for decades. To reveal the molecular-genetic basis of HC, molecular analyses of HC hearts and hearts of age-matched controls were performed.

METHODS—Total RNA and genomic DNA were prepared from formalin-fixed paraffin-embedded cardiac tissue from 12 cases of HC and 12 age-matched controls. To identify genes differentially expressed in HC, whole genome cDNA-mediated Annealing, Selection, extension and Ligation profiling was performed. TaqMan quantitative polymerase chain reaction confirmed changes in RNA expression. DNA copy number changes were measured by TaqMan copy number analysis.

RESULTS—Analysis of differential gene expression in HC cases identified two significantly down regulated gene sets aligned sequentially along the genome. The first gene cluster consisted of genes S100A8, S100A9, and S100A12 at 1q21.3c, and the second cluster consisted of genes IL1RL1 (ST2), IL18R1, and IL18RAP at 2q12.1a. Strong decreases in interleukin 33 expression were also observed. Decreases in copy number of the S100A genes were confirmed by TaqMan CNV assays. S100A genes are downstream of the p38-MAPK pathway that can be activated by interleukin 33 signaling.

CONCLUSIONS—These data suggest a model in which the interleukin 33-IL1RL1/p38-MAPK/S100A8-S100A9 axis is down regulated in HC cardiac tissue and provide several candidate genes on 1q21.3c and 2q12.1a for inherited mutations that may predispose individuals to HC.

Keywords

histiocytoid cardiomyopathy; arrhythmia; SIDS; whole genome DASL
INTRODUCTION

Histiocytoid cardiomyopathy (HC) is a rare, but distinctive arrhythmogenic disorder characterized by hamartomatous lesions of cardiac Purkinje cells. It primarily occurs in the first two years of life and has a female to male ratio of 3:1. Because of the female predominance, it was suggested that HC is an X-linked genetic disorder with prenatal lethality in males. It predominantly affects Caucasians (80%) followed by African-American (15%) and Latin-American infants (3%); it is rare in Oriental infants [1]. Since its first description by Voth (1962), HC has been published under many different synonyms, including arachnoidosis of the myocardium, infantile cardiomyopathy, infantile xanthomatous cardiomyopathy, oncocytic cardiomyopathy, focal lipid cardiomyopathy, isolated cardiac lipidosis, infantile cardiomyopathy with histiocytoid changes, myocardial or conduction system hamartoma, foamy myocardial transformation, and congenital cardiomyopathy [2].

Although approximately 100 HC cases have been reported in the literature, the prevalence of this disease may be higher than the reported cases would suggest, as some cases are mistaken for Sudden Infant Death Syndrome (SIDS) [3,4].

A familial tendency was first reported by Bruton, et al. (1977) and Suarez, et al. (1987) [5,6]. Because of these theories and the author’s personal experience of a family with more than one child affected by HC, a world-wide registry of HC was started in 1999. Data obtained from the reported cases and from the HC Registry indicate a familial tendency of approximately 5%.

Most of the published cases report a spectrum of arrhythmias; however, some infants have been misclassified as SIDS, while others experience flu-like symptoms preceding or accompanying the cardiac manifestations. Cardiomegaly is seen in most of the cases.

The etiopathogenesis of HC has eluded researchers for decades, and numerous theories for the molecular/genetic basis of HC have been offered: X-linked dominant (Xp22), cytochrome b mutation, SRY (sex determining region Y)-box 6 (Sox6) mutation, and A8344G (MERRF) mitochondrial DNA mutation [7–9]. However, such theories appear to be sporadic variants, as they have not been found in additional cases despite the collaborative work with authors of the previous series, using HC registry material. Additionally, none of these theories could be confirmed through our molecular analysis in the current study. In an effort to identify the genetic basis of HC, we used whole genome expression analysis to analyze cases from the HC registry.

MATERIALS AND METHODS

The following research protocol was approved by the Emory University Institutional Review Board.

Specimen Collection

Twelve HC autopsy cases were selected from the HC registry at Children’s Healthcare of Atlanta. Cases included nine females and three males, ranging in age from one month to 18 months. Additionally, twelve age-matched controls were collected from patients who died of non-cardiac causes (6 females, 6 males). HC lesions from the twelve cases and subendocardial Purkinje fibers from the twelve controls were isolated from paraffin embedded tissue for whole genome analysis.
RNA Preparation
Formalin-fixed paraffin embedded (FFPE) tissue sections (0.5 µm) of patient and matched control tissues were provided to the Emory Biomarker Service Center (EBSC) for RNA preparation using the Ambion RecoverAll MagMax methodology in a 96-well format on a MagMax Express Liquid Handling Robot. Total RNA was extracted from FFPE sections as previously described [10] using the RecoverAll kit (Applied Biosystems). FFPE RNA was quantitated by nanodrop spectrophotometry and tested for RNA integrity and quality by Taqman analysis of the RPL13a ribosomal protein. Samples with sufficient yield (>500 ng), A_{260}/A_{280} ratio > 1.8, and RPL13a C_T values less than 30 cycles were used for expression profiling.

DNA Preparation
FFPE tissues were deparaffinized in 1 ml of xylene at 50°C for 3 min, washed twice with 100% EtOH, and air-dried for 15 min at room temperature. Cores were digested for 48 hr at 50°C in the presence of Protease and DNA purification was carried out using the Ambion RecoverAll Total Nucleic Acid Isolation Protocol and Kit (Applied Biosystems/Ambion, Austin, TX), following the manufacturer’s instructions. DNA concentration was determined using the Quant-iT PicoGreen dsDNA Kit (Invitrogen, Carlsbad, CA) on a Gemini XPS spectrofluorometer (Molecular Devices, Sunnyvale, CA).

Expression Profiling
The cDNA-mediated Annealing, Selection, extension and Ligation (DASL) assay was performed according to the GoldenGate protocol using supplied reagents and following the manufacturer’s instructions (Illumina, Inc.). The whole genome DASL (WG-DASL) assay was used for the determination of expression of 24,526 transcripts derived from RefSeq (Build 36.2, Release 22). Biotinylated cDNAs were prepared using biotinylated random hexamers and oligo-d(T)18 primers to ensure optimal coverage of the fragmented RNA. Pooled query oligonucleotides were annealed to the cDNA under a controlled hybridization program, and unhybridized query oligonucleotides were removed from biotinylated cDNAs by streptavidin selection of hybridized duplexes. Polymerase chain reaction (PCR) was used to amplify duplexes using with Cy3 and Cy5 labeled universal PCR primers. Double-stranded PCR products were immobilized on streptavidin paramagnetic solid support beads, applied to a filter plate, and washed. Single-stranded fluoro-labeled PCR products were eluted and prepared by denaturation and subsequently hybridized to the BeadChip arrays. Arrays were scanned using an Illumina BeadStation 500 Reader.

Data Analysis
BeadStudio image analysis and data extraction software were used to extract raw data and provide preliminary analysis and quality control. On average, each oligonucleotide probe is represented by 30 beads per hybridized sample. Data were quantile normalized in BeadStudio and exported to spreadsheets for custom analysis. Differentially expressed mRNAs were identified by Significance Analysis of Microarrays (SAM) software [11]. SAM analyses were performed as two-class, unpaired analyses with 500 permutations, a 2-fold change threshold, and a false discovery rate (FDR) < 1% as previously described [12]. To identify candidate deletion regions, probes from the WG-DASL annotation files were sorted along their chromosome locations, and consecutive probes with strong decreases in gene expression (> 2-fold) were identified.

Quantitative PCR (QPCR) Validation
Gene sets identified as candidate deletion regions were validated by expression and copy number TaqMan assays. TaqMan validation was performed using an Applied Biosystems
TaqMan Copy Number Assay

Candidate deletion gene sets were validated by TaqMan Copy Number Assays (Applied Biosystems, Foster City, CA). Briefly, a pair of PCR primers was designed to amplify a short region within the target gene. A fluorescent reporter probe labeled with 6-carboxyfluorescein (FAM) was also synthesized to specifically anneal to the target. Similarly, a pair of PCR primers and a VIC-labeled probe was designed for the RNase P locus, which has a stable copy number and served as a normalization reference. The TaqMan assay was carried out on a 7900HT Fast Real-Time PCR System (Applied Biosystems). Each of the 10 µl duplex real-time PCR reactions consisted of 10 ng of genomic DNA, 2.5 µl of TaqMan Genotyping Mastermix, and 0.5 µl of primer/probe mix each for the target and reference loci. The PCR program was set for 10 min at 95°C, followed by 40 cycles of 15 sec at 95°C and 1 min at 60°C. The FAM and VIC fluorescent signal intensity was followed in real-time throughout the reaction in order to calculate the respective C_T values. Copy number analysis was performed using the SDS v2.1 software package and the CopyCaller v1.0 software (Applied Biosystems).

Decreases in copy number of S100A genes were confirmed by TaqMan copy number variation (CNV) assays. Changes in DNA copy number were measured by TaqMan copy number analysis. Primer/probes were selected so they were close to the probes used on the WG-DASL chip. To determine if the TaqMan copy number assay could be used for FFPE samples, a pair of matching FFPE and fresh-frozen normal control tissue samples were assayed.

Ingenuity Analysis

The most significantly differentially expressed genes were analyzed for gene ontology information using the Ingenuity Pathway analysis software (Ingenuity Systems).

RESULTS

Whole Genome DASL Analysis of HC samples and Validation by TaqMan QPCR

WG-DASL data from four of the samples were not useable, and analysis was performed on the remaining 20 samples (10 cases and 10 controls). SAM analysis of differentially expressed genes identified 1356 probes with at least a 2-fold change and a FDR < 1% (Supplementary Table 2). In the HC cases, 315 up-regulated probes and 1041 down-regulated probes were analyzed by hierarchical clustering (Figure 1A).

To determine candidate regions of genomic deletions, we identified sets of probes down-regulated by at least 2-fold and sequentially aligned along the human genome. Thirteen regions with at least 2 consecutive probes, 10 of which had 3 consecutive probes, were identified (Table 1). Of the 13 regions, two regions stood out due to their large fold decreases in expression. The first gene cluster consisted of three S100A genes at 1q21.3c: S100A8 (~7.6-fold), S100A9 (~4.5-fold) and S100A12 (~10.7-fold). The second gene cluster consisted of three genes involved in cytokine signaling at 2q12.1a: IL1RL1 (or ST2; ~6.9-fold), IL18R1 (~2.6-fold), and IL18RAP (~3.2-fold). Changes in RNA expression for these gene clusters were confirmed by TaqMan QPCR (Figure 1B). We also confirmed reduced expression of IL1RL1 by TaqMan assay (~52-fold), but changes in IL18R1 and IL18RAP expression were insignificant by TaqMan QPCR.
Sox6, cytochrome b, and MyoD1 were not differentially expressed by WG-DASL in HC cases. However, in the Xp22 region, a total of seven genes were differentially expressed in HC cases. Six genes (i.e., NLGN4X, P1GA, HCCS, SCML1, SMPX, PRDX4) were down regulated by more than 2-fold, and one gene (i.e., GPM6B) was up regulated by more than two-fold. These differences were likely not due to differences in male/female ratios because only 39 of 975 (4%) probes on the X chromosome were significantly changed, and the majority of those (29 of 39, or 74%) were decreased in the cases, which had fewer males than the controls.

**DNA Copy Number Analysis of HC samples**

Equal copy numbers between FFPE and fresh-frozen normal control tissue were obtained for all assays tested. This pair of samples (unrelated to HC) was included in all assays and the fresh sample set as a calibrator (copy number = 2). Due to the limited quantity of DNA recovered from the FFPE samples, three of the down regulated genes (S100A12, IL18R1 and IL1RL1) and a nearby control locus, PGLYRP3 that did not show decreased expression, were chosen for the TaqMan assay. The calculated gene copy numbers for the three candidate genes were reduced in the HC samples when compared to the matching control samples (Figure 1C). This result suggested that cases were haploinsufficient for these genes.

**Ingenuity Pathway Analysis of HC samples**

To gain insight into the biology of the observed changes in gene expression and DNA copy number in HC cases, we performed an Ingenuity Pathway Analysis. Analysis of the complete set of altered genes identified several significantly changed biological functions (Table 2). The most significantly enriched functional annotation was for cell death (p = 5.40E-09, number of genes = 247). We also observed enrichment of ten genes annotated for Cardiovascular System Development and Function (p = 7.47E-03), namely CCL2, CD44, CX3CL1, DEFA3, HAS2, IL6, ITGB1, ITGB3, SERPINE1, and VEGFA.

We next constructed networks of genes altered in HC, using the sets of genes with alterations in gene expression and DNA copy number (Figure 1D). This network identified the signal transduction pathway connecting IL-33, IL1RL1, p38MAPK, and the S100A8/S100A9 complex. Although it is not a part of one of the thirteen regions with consecutive down regulated probes shown in Table 1, we did observe a 2.4-Fold decrease in expression of interleukin (IL)-33 (Supplementary Table 2).

**DISCUSSION**

Based on minimal evidence and extrapolation, several theories of HC pathogenesis (e.g., cytochrome b mutation, SOX6 mutation, A8344G mDNA mutation) have been hypothesized [7–9]; however, sufficient evidence has not been found to incriminate a specific etiology and pathogenesis for HC. Moreover, our current whole genome expression analysis failed to support such previous theories.

The theory of Xp22 proposed by Bird et al (1994) was carefully evaluated during our whole genome DASL because of the prevalence of HC in female infants [14]. We found seven genes to be differentially expressed in HC patients compared to control patients. However, after thorough investigation, none of those showed any correlation to cardiovascular pathology. Although we were unable to associate this region to the molecular-genetic basis of HC, the enigma of the female predominance still remains.

Here, we identified two functionally clustered groups of genes on 1q21.3c (S100A8, S100A9, S100A12) and 2q12.1a (ST2, IL18R1, IL18RAP) as well as a single gene on 9p24.1b (II-33) with reduced expression in HC lesions. Furthermore, DNA copy number

*Pediatr Dev Pathol*. Author manuscript; available in PMC 2012 March 6.
decreases were confirmed for S100A12, ST2, and IL18R1. To evaluate direct and indirect gene networks, we used Ingenuity Pathway Analysis to map biological pathways that linked these genes. All genes were directly or indirectly linked within one common network. These results suggest that the IL-33/ST2/p38 MAPK/S100A8&9 axis is down regulated in HC cardiac tissue.

Transmembrane receptor ST2 is a member of the IL-1-receptor superfamily that serves as a gateway to many proinflammatory signaling pathways [15]. IL-33 was recently identified as the functional ligand for ST2 [16], and this discovery has led to many new insights into cardiovascular pathology.

The IL-33/ST2 axis has been identified as a crucial cardioprotective pathway in several studies. IL-33/ST2 signaling is capable of activating p38 MAPK and nuclear factor (NF)-κB signaling pathways [16]. Zechner et al. (2008) recently reported that P38-dependent NF-κB activation aids in protecting myocardial cells from undergoing apoptosis [17]. Additionally, IL-33 is produced by cardiac fibroblasts in response to biomechanical stress, and its signaling through ST2 negatively regulates cardiomyocyte hypertrophy and cardiac fibrosis [18]. Activation of this pathway has also been shown to prevent cardiomyocyte apoptosis and improve cardiac function post myocardial infarction [15]. Finally, Yin et al. (2010) identified a potential role for IL-33/ST2 signaling in the induction and maintenance of cardiac transplant tolerance [19].

S100 calcium-binding proteins S100A8, S100A9 and S100A12 were also down regulated in HC. S100A8 and S100A9 are found both intracellularly and extracellularly and are, therefore, capable of exerting both autocrine and intracellular effects [20,21]. Together, S100A8 and S100A9 form the ubiquitous heterodimer calprotectin, a proinflammatory cytokine, and this protein complex has been shown to exert a pro-apoptotic effect [22]. Furthermore, like IL-33/ST2 signaling, there is strong evidence that extracellular S100A8/A9 signaling via the RAGE (Receptor of Advanced Glycation End products) receptor results in p38 MAPK activation of NFκB [23]. Furthermore, S100A8 and S100A9 are known to be important in heart function [24]. In the heart, inflammatory-challenged cardiac cells secrete S100A8/A9 and likely signal through RAGE to suppress cardiac contractility [24]. Finally, although its function is still poorly understood, S100A12 is also implicated in cardiovascular pathology, and its downregulation has been shown to suppress inflammatory and apoptosis pathways [22,25].

IL18R1 and IL18RAP are the receptor and accessory protein, respectively, for the pro-hypertrophic cytokine IL-18. In a mouse model, signaling of IL-18 via IL18R1 and IL18RAP has been shown to induce cardiomyocyte hypertrophy.

In conclusion, our data suggest a potential role of the IL-33/ST2/p38-MAPK/S100A8-S100A9 signaling cascade in the pathogenesis of HC. The net effect is reduced NF-κB activation and reduced survival of cardiac myocytes, ultimately resulting in cardiac failure during development. These studies provide the starting point for examination of several candidate genes on 1q21.3c and 2q12.1a for inherited mutations that predispose patients to HC.

Future plans are to analyze remaining HC registry cases to confirm the present findings. Additionally, we plan to collect and analyze blood from the siblings and parents of HC patients to better understand the pathogenesis and inheritance pattern.

**Supplementary Material**

Refer to Web version on PubMed Central for supplementary material.
Acknowledgments

This work was supported in part by the Emory Biomarker Service Center. The authors also thank all contributors and supporters of the Histiocytoid Cardiomyopathy Registry at Children’s Healthcare of Atlanta.

LITERATURE CITED


Figure 1.
(A) Hierarchical clustering of 1356 up- and down-regulated probes in HC cases relative to controls. Probes shown had a minimum of 2-fold change and FDR < 1%. (B) TaqMan qPCR validation of mRNA changes observed by whole-genome DASL assay in candidate HC genes. (C) DNA Copy Number Analysis Results. The TaqMan assay revealed reduced copy numbers for genes S100A12, IL18R1, and IL1RL1 in HC samples (blue bars) compared to control samples (red bars). Nearby control locus, PGLYRP3 did not show significant changes in copy number. (D) Network of altered candidate genes in HC. Solid lines represent direct interactions. Dotted lines represent indirect interactions. The node indicates the class of the molecule. The intensity of the node color indicates the degree of up (red)- or down (green)-regulation.
<table>
<thead>
<tr>
<th>Cytoband</th>
<th>HUGO Symbols</th>
</tr>
</thead>
<tbody>
<tr>
<td>1p22.1b-1p22.1a</td>
<td>FNBP1L, BCAR3, DNTTIP2</td>
</tr>
<tr>
<td>1q21.3c</td>
<td>S100A9, S100A12, S100A8</td>
</tr>
<tr>
<td>2q11.2c</td>
<td>TXNDC9, EIF5B, REV1</td>
</tr>
<tr>
<td>2q12.1a</td>
<td>IL1RL1, IL1RL1, IL18R1, IL18RAP</td>
</tr>
<tr>
<td>4q32.3c–q33a</td>
<td>PALLD, CBR4, SH3RF1, NEK1</td>
</tr>
<tr>
<td>5q33.3c</td>
<td>EBF1, UBLCP1</td>
</tr>
<tr>
<td>6p22.3e</td>
<td>RBM24, CAP2</td>
</tr>
<tr>
<td>8q24.22b–8q24.22c</td>
<td>PHF20L1, SLA</td>
</tr>
<tr>
<td>9q22.33c–9q31.1a</td>
<td>SEC61B, NR4A3</td>
</tr>
<tr>
<td>9q31.2a–q31.2b</td>
<td>TMEM38B, ZNF462</td>
</tr>
<tr>
<td>10q21.3e</td>
<td>DDX21, KIAA1279, SRGN</td>
</tr>
<tr>
<td>13q14.2a–13q14.2b</td>
<td>HTR2A, SUCLA2, NUDT15</td>
</tr>
<tr>
<td>17q23.1a</td>
<td>TMEM49, TUBD1, RPS6KB1</td>
</tr>
</tbody>
</table>
Table 2
Subset of significantly enriched Gene Ontologies for Biological Functions identified by Ingenuity Pathway Analysis.

<table>
<thead>
<tr>
<th>Category</th>
<th>P-value</th>
<th># Molecules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cell Death</td>
<td>5.40E-09</td>
<td>247</td>
</tr>
<tr>
<td>Infectious Disease</td>
<td>1.35E-08</td>
<td>28</td>
</tr>
<tr>
<td>Connective Tissue Disorders</td>
<td>7.03E-06</td>
<td>54</td>
</tr>
<tr>
<td>Cell Cycle</td>
<td>1.10E-05</td>
<td>120</td>
</tr>
<tr>
<td>Cellular Development</td>
<td>1.13E-05</td>
<td>23</td>
</tr>
<tr>
<td>Hematological System Development and Function</td>
<td>1.95E-05</td>
<td>17</td>
</tr>
<tr>
<td>Cellular Growth and Proliferation</td>
<td>4.29E-05</td>
<td>39</td>
</tr>
<tr>
<td>Gene Expression</td>
<td>8.72E-05</td>
<td>55</td>
</tr>
<tr>
<td>Endocrine System Disorders</td>
<td>8.87E-05</td>
<td>17</td>
</tr>
<tr>
<td>Cardiovascular System Development and Function</td>
<td>7.47E-03</td>
<td>10</td>
</tr>
</tbody>
</table>