Cardiovascular Pathology 34 (2018) 50–57

Contents lists available at ScienceDirect

Cardiovascular Pathology

Original Article

Cardiomyocytes have mosaic patterns of protein expression☆,☆☆

Tony Y. Wang [a], Dongwong Lee [b], Karen Fox-Talbot [a], Dan E. Arking [b], Aravinda Chakravarti [b], Marc K. Halushka [a,*]

[a] Department of Pathology, Johns Hopkins School of Medicine, Baltimore, MD, USA
[b] McKusick - Nathans Institute of Genetic Medicine, Johns Hopkins School of Medicine, Baltimore, MD, USA

ARTICLE INFO

Article history:
Received 31 January 2018
Received in revised form 4 March 2018
Accepted 19 March 2018
Available online xxxx

Keywords:
Myosin light chain
Proteomics
Myomesin
Cardiomyocyte
Expression

ABSTRACT

Skeletal myocytes have well-established fast and slow twitch fibers with unique gene and protein specific expression patterns. By immunohistochemical staining, these show a mosaic pattern across myocytes. We hypothesized cardiac myocytes may behave similarly where some proteins are differentially expressed between mature cardiomyocytes. We utilized the tool HPASubC on over 52,000 cardiac images of the Human Protein Atlas to identify differential protein expression patterns by immunohistochemistry across the cardiomyocytes. We matched identified proteins to open chromatin and gene expression data. We identified 143 putative proteins with mosaic patterns of expression across the cardiomyocytes. We validated four of these proteins (MYL3, MYL4, PAM, and MYOM1) and demonstrated unique atrial or ventricular patterns of expression for each. Acetylation of histone H3K27 at the promoters of these four genes were consistent with the atrial/ventricular expression patterns. Despite the generally accepted homogeneity of cardiomyocytes, a small subset of proteins varies between cardiomyocytes in a mosaic pattern. This fundamental process has been previously uncharacterized. These changes may inform on different functional and disease-related activities of proteins in individual cardiomyocytes.

1. Introduction

It is well-established that skeletal muscle cells can be simplistically divided into slow-twitch (type I) and fast-twitch (type II) fibers [1]. These fibers each have unique and essential activities that are mediated by different patterns of protein expression. Differentially expressed proteins in these fibers include myosin heavy chains (MYH1, MYH2), tropomyosins (TPM1, TPM2), troponins (TNNI1, TNNI2) and others [2]. Individual skeletal myocyte protein expression differences have been noted by ATPase activity, immunohistochemistry (IHC) and immunofluorescence (IF) studies. Depending on the function of a skeletal muscle, the ratio of the fibers will vary [2]. These same proteins are not known to be differentially expressed in human cardiomyocytes, therefore there has been little effort focused on identifying proteins that might vary between cells in the heart. However, there have been some hints in the literature that cardiac myocytes may also vary in protein expression.

In 1984, the Yazaki group reported isozymic changes of human atrial myosin (referred to as HCα, but currently named MYL4) where they showed IF differences across ventricular myocytes [3]. More recently, the work of the Baker laboratory has demonstrated cellular heterogeneity of responses to α1 adrenergic stimulation in isolated myocytes [4]. The Wang laboratory showed that cardiac-specific deletion of *Trbp* in a mouse model changed the expression of fast and slow-twitch myofiber genes including *Tnni2*, *Tpm2*, *Myl1*, *Myl7*, *Myl9*, *Tnnc1*, *Myh7b* and *Myl3* [5]. The summation of these papers suggests there might be a collection of proteins that differ between cardiac myocytes. However, there has not yet been a way to globally and systematically investigate thousands of proteins for these alternative patterns of expression in cardiac myocytes.

The Human Protein Atlas (HPA) is a web resource of proteomic expression that can be used to investigate this question. The HPA has generated scores of tissue microarrays (TMAs) using 44 normal and 20 cancer tissues. These TMAs have been stained using IHC for >17,000 proteins using >25,000 antibodies, with each antibody generally staining three cores of tissue per organ type. This has resulted in ~50,000 heart images (predominately from ventricles) that can be evaluated for any pattern of staining [6].

To investigate interesting staining patterns of HPA tissues as a means to identify similarly expressed proteins, we developed HPASubC [7]. This is a suite of python tools that allows a user to download all of the images of a particular organ, then sequentially review them for a pattern of interest using a PlayStation-style gamepad controller. Images can be rapidly reviewed at ~1/s in an anonymous fashion. We have

☆ Sources of support: None.
☆☆ The authors declare they have no conflict of interest.
* Corresponding author at: Ross Rm 632B, 720 Rutland Avenue, Baltimore, MD, 212805. Tel.: +1 410 614 8138; fax: +1 410 502 5862.
E-mail address: mhalush1@jhmi.edu (M.K. Halushka).

https://doi.org/10.1016/j.carpath.2018.03.002
1054-8807/

demonstrated its usefulness in describing the patterns of non-cardiomyocyte expression in the heart and in sinusoidal expression in the liver [7,8].

We became interested in the concept of cardiac myocyte mosaicism while reviewing HPA heart images for a separate analysis and began noticing occasional images where only a subset of myocytes were stained. We then hypothesized that, like skeletal muscle, there may be different fiber types of the heart denoted by alternative protein expression.

2. Methods

2.1. The Human Protein Atlas and HPASubC

We downloaded 52,737 heart images representing 12,814 proteins from HPA (v16) on 12/9/2016 utilizing the HPASubC script download_images_from_gene_list.py. We then used the HPASubC image_viewer.py to rapidly cycle through the images at ~3000 per viewing. We captured any image in which the staining intensity varied across the myocytes reasoning it could be the result of a mosaic pattern of protein expression. Staining of other cell types was not interpreted. We then re-evaluated all of the images that corresponded to any protein identified and rescored them as mosaic in a second, stricter screening in which the mosaicism between myocytes was more clear. Although multiple images existed for each protein, the presence of a mosaic pattern on only a single core image was sufficient to call the protein mosaic and be included in this study. Additional scripts were used to download RNA-seq expression data and other protein information from HPA. Skeletal muscle images corresponding to the identified mosaic heart proteins were downloaded along with images corresponding to 1000 random proteins (based on ENSG ID numbers) not identified as mosaic in the heart materials.

2.2. Immunohistochemistry

We obtained multiple human cardiac tissues under an approved Investigational Review Board protocol. This included multiple regions of a human heart explanted for cardiac transplantation and pediatric tissues. Slides were obtained from these formalin-fixed paraffin-embedded blocks and underwent IHC. Formalin-fixed paraffin embedded tissues were sectioned onto plus slides. Paraffin removal and high- temperature antigen retrieval was performed by immersing the slides in Trilogy (Sigma-Aldrich, St. Louis, MO) in a pressure cooker to 126 °C and 18–23 psi. Endogenous peroxidase was blocked by incubating the slides in a dual enzyme block (Dako North America, Carpinteria, CA). Slides were incubated with primary antibodies to either Myomesin 1 (MYOM1; 1:200); Myosin, light chain 3 (MYL3; 1:500); Myosin, light chain 4 (MYL4; 1:500); or peptidylglycine alpha-amidating monooxygenase (PAM; 1:500) (Atlas Antibodies, Stockholm, Sweden) followed by an incubation with a polymer HRP IgG (Leica Biosystems, Pleasanton, CA). The antibody complex was detected with ImmPact DAB (Vector Laboratories, Burlingame, CA) and the slides were counterstained with Hematoxylin (Richard-Allen Scientific, Kalamazoo, MI).

2.3. String analysis, gene ontology, and mitochondrial localization

The identified mosaic proteins were analyzed using String 10.0 [9] for interactions using the web-based portal at https://string-db.org/. Gene Ontology, a method to find enrichment of functional relationships within gene lists, was ascertained using the Gene Ontology Consortium web-based portal at http://geneontology.org/ [10]. All proteins listed in the Human Mitochondrial Protein Database (http://bioinfo.nist.gov/hmpd/index.html) were obtained and a final list of 813 unique protein names was generated.

2.4. Histone modification ChIP-seq analysis

We downloaded mapped reads, fold change over control, and stable peaks for H3K27ac ChIP-seq experiments for eight heart tissues from the ENCODE Project Consortium [11] website (www.encodeproject.org) (Table 1). For each of the 143 genes that showed mosaic expression patterns, we calculated normalized ChIP-seq read ratio (R_i) in the promoter regions (5000 basepair of DNA centered at transcription start sites) for each sample using the following formula: $R_i = (h_i / H) / (c_i / C)$, where h_i and c_i are the read counts from histone and control ChIP-seq for gene i, and T and C are the total read counts from histone and control ChIP-seq, respectively. We used bedtools (v2.24.0) 'multicov' command [12] for counting reads from the bam files. These values were then $\log_2$ transformed and averaged for each tissue type (atria vs ventricles). The log ratio between atria and ventricles were calculated lastly. For the gene expression data, GTEx V7 median gene transcripts per million (TPM) values [13] for the heart tissues (atria and ventricles) were used. Median TPMs were further $\log_2$ transformed with 0.001 pseudo-count and quantile normalized. We developed in-house scripts for the analysis, which are available upon request. We visualized enhancer and promoter activities around the four genes (MYL4, MYL3, MYOM1, and PAM) using the ChIP-seq signal data (fold change over control) with the Integrative Genomics Viewer (IGV) [14]. Hg19 was used as a reference genome for the analysis. The range of signals for visualization was set between 0 and 30 for all tracks.

3. Results

3.1. One hundred forty-three proteins have possible mosaic patterns of expression

We viewed 52,737 heart images covering 12,814 proteins from the HPA using HPASubC. From these we identified 319 possible mosaic proteins in an initial screen. A second, more stringent screening of the images reduced this number to 152 proteins. We then obtained matched transcripts per million (TPM) RNA-seq data from HPA and removed any protein for which the TPM value was 0. This left 143 putative mosaic proteins (1.1% of all proteins reviewed) (Table 2, Supplemental Table 1). For some proteins (such as FAM3C), the pattern was clear, while for other proteins, the differences were subtler and not always consistent across cores (Fig. 1, Supplemental Fig. 1). Even within a given myocyte, the staining could vary across the cell. The identified proteins included well-known cardiac proteins such as the myosin light chains (MYL3 & MYL4) as well as lesser explored proteins such as zinc finger CCHC-type containing 10 (ZCCHC10) and U-box domain containing 5 (UBOX5). The genes for seven of these proteins reside on the X chromosome. In comparison to the general mitochondrial proteome (813 proteins), three possible mosaic proteins were on that list (CASQ1, PDP1, and RTN4).

3.2. Roughly one-quarter of heart mosaic proteins are also mosaic in skeletal muscle

We then inquired if there was a relationship between cardiac mosaic patterns and skeletal mosaic patterns. To do this, we reviewed all skeletal muscle images at HPA that corresponded to the 143 putative mosaic proteins. Of these, 39 proteins (27%) were mosaic in both cardiac and skeletal muscle (Supplemental Fig. 2). As a comparison group, we evaluated 1000 random proteins in the skeletal muscle images and found 758 to have some degree of staining. Of these, 216 (28.5%) were mosaic. Thus, there was no enrichment in skeletal muscle.

3.3. Mosaic patterns differ by protein across atria and ventricles

We selected four of the putative mosaic proteins for further review (MYL4, MYL3, MYOM1, PAM). Three of these (MYL4, MYL3, and