

# The Whole Genome Expression Analysis using Two Microarray Technologies to Identify Gene Networks That Mediate the Myocardial Phenotype of CD36 Deficiency

Imane Sabaouni<sup>1</sup>, Ahmed Moussa<sup>2\*</sup>, Brigitte Vannier<sup>3</sup>, Oussama Semlali<sup>1</sup>, Terri A Pietka<sup>4</sup>, Nada A Abumrad<sup>4</sup> & Azeddine Ibrahimi<sup>1</sup>

<sup>1</sup>Medical Biotechnology Lab (MedBiotech), Rabat Medical & Pharmacy School, Mohammed V<sup>th</sup> Souissi University, Rabat, Morocco; <sup>2</sup>LabTIC Laboratory, ENSA, Abdelmalek Essaadi University, Tangier, Morocco; <sup>3</sup>Receptors, Regulation and Tumor Cells (2RTC) Laboratory, University of Poitiers, France; <sup>4</sup>Division of Biology and Biomedical Sciences, Washington University, USA; Ahmed Moussa – Email: amoussa@uae.ac.ma; \*Corresponding author

Received January 27, 2013; Accepted September 30, 2013; Published October 16, 2013

## Abstract:

We have previously shown that CD36 is a membrane protein that facilitates long chain fatty acid (FA) transport by muscle tissues. We also documented the significant impact of muscle CD36 expression on heart function, skeletal muscle insulin sensitivity as well as on overall metabolism. To identify a comprehensive set of genes that are differentially regulated by CD36 expression in the heart, we used two microarray technologies (Affymetrix and Agilent) to compare gene expression in heart tissues from CD36 Knock-Out (KO-CD36) versus wild type (WT-CD36) mice. The obtained results using the two technologies were similar with around 35 genes differentially expressed using both technologies. Absence of CD36 led to down-regulation of the expression of three groups of genes involved in pathways of FA metabolism, angiogenesis/apoptosis and structure. These data are consistent with the fact that the CD36 protein binds FA and thrombospondin 1 involved respectively in lipid metabolism and anti-angiogenic activities. In conclusion, our findings led to validate our data analysis workflow and identify specific pathways, possibly underlying the phenotypic abnormalities in CD36 Knock -Out hearts.

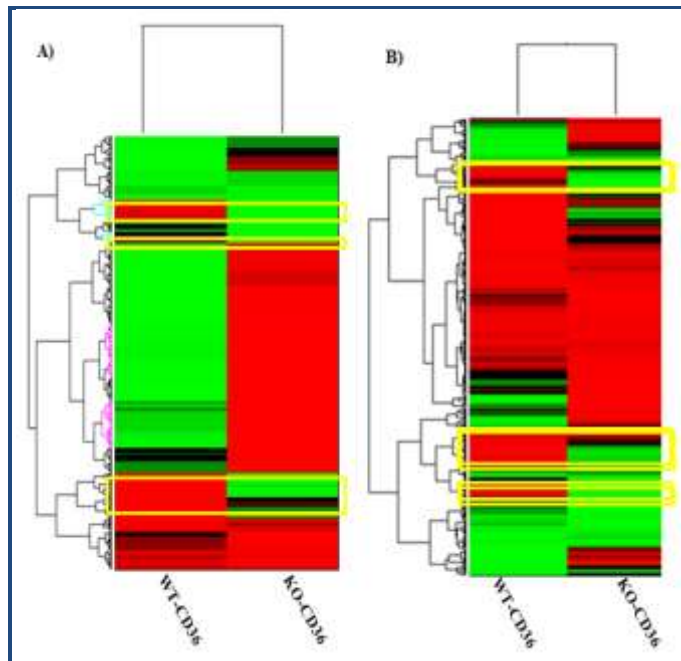
**Keywords:** CD36, Fatty Acid, Microarrays, Metabolism, angiogenesis/apoptosis, Protein interaction, Gene expression.

## Background:

CD36 gene encodes a membrane glycoprotein, and has been identified in wide variety cells types, including platelets, monocytes, and erythroblast, capillary endothelial and mammary epithelial cells [1-6]. CD36 (also known as platelet glycoprotein IV or IIIb) is also a membrane glycoprotein highly expressed in heart tissue. It was shown that CD36 works as receptor/transporter of long chain fatty acids (FA) in muscle tissue and is proposed as one of thrombospondine receptor in endothelial cells [5]. CD36-KO mice (with no expression of

CD36 gene) exhibits defective FA uptake by the heart, which is paralleled by an increase in the heart/body index and by an enlargement of left ventricular space [1]. Two sets of studies were done to identify a comprehensive set of genes that are differentially regulated by CD36 expression in the heart. In 2002 and 2007, we used respectively the Affymetrix and the Agilent technologies to analyze CD36 involvement in the fatty acids uptake and heart hypertrophy [7, 3-5, 8]. We propose to compare results obtained from the two microarrays technologies and investigate the consequences of CD36 absence

on CD36-KO hearts. In this paper, we will describe the methodology used to identify the differentially expressed genes using the Affymetrix and the Agilent technologies. In the second section, we will use the classification results to determine the gene clusters. Finally, these classifications will be used to annotate the functional class of each cluster and characterize the molecular pathways involved in the myocardial phenotype of KO-CD36 mice.



**Figure 1:** Hierarchical clustering of selected genes: **A)** Hierarchical clustering using selected genes from Affymetrix Technology; **B)** Hierarchical clustering using selected genes from Agilent Technology

## Methodology:

### Data analysis

In order To search for differentially regulated gene networks in the absence of CD36 gene, we performed a comprehensive gene analysis by hybridizing microarray chips with RNA probes prepared from mouse heart CD36-KO and CD36-WT. Two technologies were used. The Affymetrix GeneChip Murine Genome U74 which contains 36,000 probes (Affymetrix, Santa Clara, CA). Once the probe array had been hybridized, stained, and washed, it was scanned using a GeneArray scanner. A GeneChip Operating System, running on a PC workstation was used to control the functions of the scanner and collect fluorescent intensity data. The second approach used was the Whole Mouse Genome Microarray 4x44K (Agilent Technologies, Santa Clara, CA). Arrays were washed and dried out using a centrifuge according to manufacturer's instructions (One-Color Microarray-Based Gene Expression Analysis, Agilent Technologies). Arrays were scanned at 5 mm resolution on an Agilent DNA Microarray Scanner (GenePix 4000B, Agilent Technologies) using the default settings for 4x44k format one-color arrays. Images provided by the scanner were analyzed using Feature Extraction software v10.1.1.1 (Agilent Technologies). Raw data files were analyzed using the software R associated to packages of "Bioconductor" project [8]. Developed affymetrix workflows begin with data normalization using Robust Multichip Average Method (RMA)

[9], which allows reduction of block effect done at the probset level. The second step consisted in selecting differentially expressed genes between CD36-WT and CD36-KO using the Significance Analysis of Microarrays (SAM) algorithm [10] with the Fold Change and P-value Cutoffs respectively fixed to at 1.5 and 0.002. Agilent workflow begins with data normalization by using Lowess normalization [11] that applied to a two-color array expression dataset. The second step, as in the case of Affymetrix, SAM algorithm was used to identify differentially expressed genes with the same FC and P-value cutoffs. Class discovery analyzed a given set of genes to produce subgroups that share common features. An analysis method often used for class discovery is "cluster analysis" or clustering. It is aimed at dividing the data points (genes or samples) into groups (clusters) using measures of similarity [12, 13] creating hierarchical clustering of co-regulated genes

### Identification of differentially expressed genes

Using Affymetrix technology, we were able to identify 39 differentially expressed genes between CD36-KO and CD36-WT and when using the Agilent technology with the same parameters, we identified 35 differentially expressed genes. The comparison of the two lists of identified genes showed that 30 of them were common to the two technologies. Differentially expressed genes were clustered by hierarchical clustering (Figure 1). This type of classification and class discovery involves analyzing a given set of gene expression profiles with the goal of discovering subgroups that share common features.

### Functional Analysis

After establishing a differentially expressed gene list common to the two technologies (Affymetrix and Agilent), functional annotation allowed the determination of each gene function in the list. Functional classes were extracted using Gene Ontology tools. Associations, biochemical pathway data were retrieved from the Gene Ontology consortium (GO) [14, 15], Kyoto Encyclopedia of Genes and Genomes (KEGG) [16]. Gene sets were arranged and associated with different functional categories. Integrated expression data and evolutionary conservation of proteins were used to predict interacting proteins, protein complexes and proteins functions [17, 18, 19]. Table 1 (See supplementary material) showed the first classification leading to the identification of three gene categories: a gene group involved in metabolism, another group involved in the Angiogenesis Table 2 (see supplementary material) and a third one implicated in cellular remodeling Table 3 (see supplementary material) (structural genes).

The results appear to be consistent with the role of the CD36 protein in cardiac muscle cells. The identification of the metabolism genes could be explained by the role of CD36 as a receptor / transporter for long-chain fatty acids in heart cells [3]. Indeed, Randle *et al.* [20] showed that fatty acids are the main energy source for the heart and have brought to light from in vitro experiments that long chain fatty acids is preferentially metabolized. He also demonstrated the existence of competition between glucose and fatty acids as heart fuel [20]. Moreover, Dyck *et al.* showed that CD36 plays an important role in the choice of substrate in the heart [21]. In the CD36-KO hearts, the shift to glucose substrate since less FA is available led to the expression and down regulation of genes

involved in both metabolisms. Secondly, in endothelial heart cells; CD36 has been described as a thrombospondin membrane receptor (TSP-1) and plays a role in the inhibition of endothelial cell migration and apoptosis induction. The absence of CD36 led to down regulation of the expression of its ligand (Tsp1) and the expression of new signaling genes. Finally and in order for CD36-KO to go through heart hypertrophy, a set of remodeling genes is expressed and can be grouped into the category of structural genes as shown in **Table 1**.

## Discussion:

In this study, we compared results obtained from two technologies using the same analysis workflow. We first evaluated variance among replicates within each of the platforms and found low levels of variance and high correlation among the two platforms eventhough the two technologies were used at different labs and time. Agilent oligonucleotide technology was used in 2007 at Georges Washington University (SL, MO) and the Affymetrix U74Av2 technology was used in 2002 at Stony Brook University (Stony Brook, NY). Using SAM, we were not able to find any significant differences among the two platforms looking at their ability to detect differential gene expression between WT-CD36 and KO-CD36. Technological differences may influence the results of transcriptional profiling and are important to consider while using published results. However, and based on our study and given high-quality arrays and the appropriate normalization, the primary factor determining variance is biological rather than technological. The biological conditions of the two experiments could explain the small differences in the obtained results. Questions remain regarding the importance of technology choice in evaluating the data generated and comparing among experiments from different laboratories. One of the objective of this comparative study was to elucidate whether gene expression profiles are more influenced by biology or by technological artifacts. Eventhough, the two platforms are based on two distinct manufacturing technologies; a two-color cDNA spotted arrays (Agilent) and in situ synthesized oligonucleotide chips (Affymetrix), our results showed comparable results.

## Conclusion:

Our comparative study led to the validation of a data analysis workflow and the identification of, at least, 30 genes involved in the phenotype of CD36 heart hypertrophy. More biological studies are needed to validate the expression of the identified genes using the qRT-PCR. These studies will be complemented with a modeling project based on constructing a bioinformatic platform. It will be reproducing the behavior dynamic of the system under normal conditions and automatically predicting

the involvement of different gene networks in the development of pathologies such as cardiomyopathy related to the absence of CD36 protein.

## Acknowledgment:

This work was carried out under intramural funding from the Stony Brook University and Mohammed the V<sup>th</sup> Souissi University to AI. We acknowledge support from Volubilis (French-Moroccan Grant) to AI, BV & AM. This work was also supported by a grant from the NIH for H3Africa BioNet to AI & AM. We would like to thank Microarrays facilities in Stony Brook and George Washington Universities for conducting the arrays hybridization experiments.

## References:

- [1] Abumrad NA *et al.* *J Biol Chem.* 1993 **268**: 17665 [PMID: 7688729]
- [2] Kashiwagi H *et al.* *Blood.* 1994 **83**: 3545 [PMID: 7515716]
- [3] Abumrad N *et al.* *J Lipid Res.* 1998 **39**: 2309 [PMID: 9831619]
- [4] Ibrahimi A & Abumrad NA, *Curr Opin Clin Nutr Metab Care.* 2002 **5**: 139 [PMID: 11844979]
- [5] Ibrahimi A *et al.* *Proc Nat Acad Sci U S A.* 1996 **93**: 2646 [PMID: 8610095]
- [6] Asch As *et al.* *J Clin Invest.* 1987 **79**: 1054 [PMID: 2435757]
- [7] Hajri T *et al.* *J Biol Chem.* 2001 **276**: 23661 [PMID: 11323420]
- [8] Bioconductor Project [<http://www.bioconductor.org/>]
- [9] Irizarry RA *et al.* *Nucleic Acids Res.* 2003 **31**: e15 [PMID: 12582260]
- [10] Tusher VG *et al.* *Proc Nat Acad Sci U S A.* 2001 **98**: 5116 [PMID: 11309499]
- [11] Grosu P *et al.* *Genome Res.* 2002 **12**: 1121 [PMID: 12097350]
- [12] Aach J *et al.* *Genome Res.* 2000 **10**: 431 [PMID: 10779484]
- [13] Bethin KE *et al.* *Mol Endocrinol.* 2003 **17**: 1454 [PMID: 12775764]
- [14] Doniger SW *et al.* *Genome Biol.* 2003 **4**: R7 [PMID: 12540299]
- [15] <http://www.genome.jp/kegg/>
- [16] Voit EO & Riley M, *Genome Biol.* 2003 **4**: 235 [PMID: 14611652]
- [17] Ge H *et al.* *Nat Genet.* 2001 **29**: 482 [PMID:11694880]
- [18] Jansen R & Gerstein M, *Nucleic Acids Res.* 2000 **28**: 1481 [PMID: 10684945]
- [19] Teichmann SA & Babu MM, *Trends Biotechnol.* 2002 **20**: 407 [PMID: 12220896]
- [20] Randle PJ, *Lancet.* 1963 **1**: 785 [PMID: 13990765]
- [21] Van Noort V *et al.* *Trends Genet.* 2003 **19**: 238 [PMID: 12711213]

Edited by P Kanguane

Citation: Sabouni *et al.* *Bioinformation* 9(17): 849-852 (2013)

**License statement:** This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited

## Supplementary material:

**Table 1:** Functional classification of differentially expressed genes: *A: Metabolism genes*

Gene name	Description
CD36	Binds long chain fatty acids and may function in the transport and/or as a regulator of fatty acid transport
Dgat1	Catalyzes the terminal and only committed step in triacylglycerol synthesis by using diacylglycerol and fatty acyl CoA as substrates
Adh4	Involved in the reduction of benzoquinones
Irs3	Insulin receptor substrate 3
Irs1	Mediate the control of various cellular processes by insulin.
IL2	Negative regulation of heart contraction
Fat4	Function in the regulation of planar cell polarity.
Pdhb	Complex catalyzes the overall conversion of pyruvate to acetyl-CoA and CO <sub>2</sub> .
Lpl	The apolipoprotein, APOC2, acts as a coactivator of LPL activity in the presence of lipids on the luminal surface of vascular endothelium
Ide	Plays a role in the cellular breakdown of insulin
Uqcrrh	This is a component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is part of the mitochondrial respiratory chain.
Uqcrc1	This is a component of the ubiquinol-cytochrome c reductase complex (complex III or cytochrome b-c1 complex), which is part of the mitochondrial respiratory chain
Gyk	Key enzyme in the regulation of glycerol uptake and metabolism
Ifng	Regulation of the force of heart contraction
Mb	Heart development and Serves as a reserve supply of oxygen and facilitates the movement of oxygen within muscles.

**Table 2:** Angiogenesis / Apoptosis genes

Gene name	Description
CD36	Receptor for thrombospondins, THBS1 AND THBS2, mediating their antiangiogenic effects
Thbs1	Adhesive glycoprotein that mediates cell-to-cell and cell-to-matrix interactions. Ligand for CD36 mediating antiangiogenic properties
Sdc4	thrombospondin receptor activity and Cell surface proteoglycan that bears heparan sulfate
CD9	Involved in platelet activation and aggregation.
Pdgfra	vascular endothelial growth factor receptor signaling pathway
Hand1	Transcription factor that plays an essential role in both trophoblast-giant cells differentiation and in cardiac morphogenesis.
Arnt2	negative regulation of apoptotic process and positive regulation of cell proliferation
Rag1	negative regulation of cysteine-type endopeptidase activity involved in apoptotic process
Api5	Antiapoptotic factor that have a role in protein assembly. Negatively regulates ACIN1. Also known to efficiently suppress E2F1-induced apoptosis
Map3K2	Component of a protein kinase signal transduction cascade, Plays a role in caveolae kiss-and-run dynamics
Mertk	Receptor tyrosine kinase that transduces signals from the extracellular matrix into the cytoplasm by binding to several ligands

**Table 3:** Cellular remodeling genes (cell structure).

Gene name	Description
Myl4	Myosin light chain
Tnnt2	atrial cardiac muscle tissue morphogenesis
Hand1	ventricular cardiac muscle tissue morphogenesis and cardiac septum morphogenesis,
Pdlim3	actin filament organization in heart development
Pdgfra	cardiac myofibril assembly