Genomics Data 4 (2015) 50–53

Contents lists available at ScienceDirect

Genomics Data

journal homepage: http://www.journals.elsevier.com/genomics-data/

Data in Brief

Gene expression profiling during intensive cardiovascular lifestyle modification: Relationships with vascular function and weight loss

Heather L. Blackburn [a], Seóna McErlean [b], Gera L. Jellema [b], Ryan van Laar [c], Marina N. Vernalis [d], Darrell L. Ellsworth [a,*]

[a] Integrative Cardiac Health Program, Windber Research Institute, Windber, PA 15963, USA
[b] Almac Diagnostics, Craigavon BT63 5QD, UK
[c] ChipDx, New York, NY 10128, USA
[d] Integrative Cardiac Health Program, Walter Reed National Military Medical Center, Bethesda, MD 20889, USA

ARTICLE INFO

Article history:
Received 24 February 2015
Received in revised form 3 March 2015
Accepted 3 March 2015
Available online 12 March 2015

Keywords:
Cardiovascular disease
Lifestyle modification
Microarray
Gene expression
Weight loss

ABSTRACT

Heart disease and related sequelae are a leading cause of death and healthcare expenditure throughout the world. Although many patients opt for surgical interventions, lifestyle modification programs focusing on nutrition and exercise have shown substantial health benefits and are becoming increasing popular. We conducted a year-long lifestyle modification program to mediate cardiovascular risk through traditional risk factors and to investigate how molecular changes, if present, may contribute to long-term risk reduction. Here we describe the lifestyle intervention, including clinical and molecular data collected, and provide details of the experimental methods and quality control parameters for the gene expression data generated from participants and non-intervention controls. Our findings suggest successful and sustained modulation of gene expression through healthy lifestyle changes may have beneficial effects on vascular health that cannot be discerned from traditional risk factor profiles. The data are deposited in the Gene Expression Omnibus, series GSE46097 and GSE66175.

Specifications	
Organism/cell line/tissue	*Homo sapiens*/whole blood
Sex	Male and female
Sequencer or array type	Affymetrix GeneChip HG-U133A 2.0 arrays
Data format	Raw data: CEL/TAR files, Normalized data: SOFT, MINiML, TXT
Experimental factors	Clinical: Standard demographic and clinical information, physiological and biochemical assessment; Molecular: RNA isolated from PAXgene™ tubes, globin reduction treatment of RNA, standard Affymetrix expression analysis, transcript validation by qRT-PCR
Experimental features	Intensive lifestyle modification to stabilize or reverse progression of heart disease over 1 year; participants and retrospectively matched controls with CAD or 2+ risk factors; group comparisons; risk factor correlations with gene expression; functional enrichment and pathways analysis; medication influence
Consent	All patients provided a written informed consent before participation. The study protocol (Pro00009375) was approved by the Chesapeake Institutional Review Board (https://www.chesapeakeirb.com/).
Sample source location	Windber, Pennsylvania, USA

* Corresponding author at: Windber Research Institute, 620 Seventh Street, Windber, PA 15963, USA. Tel.: +1 814 361 6911.
E-mail address: d.ellsworth@wriwindber.org (D.L. Ellsworth).

Direct link to deposited data

The study is registered as NCT01805492 at ClinicalTrials.gov. Expression data were deposited in the Gene Expression Omnibus (GEO) under series accession numbers GSE46097 and GSE66175 and are available here: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE46097 and http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE66175.

Experimental design, materials and methods

Objectives

The main objectives of this project were to 1) characterize longitudinal changes in gene expression in peripheral blood during an intensive cardiovascular lifestyle intervention, and 2) identify associations between gene expression profiles and changes in quantitative heart disease risk factors during the intervention. Our goal was to provide a global view of molecular changes associated with drastic lifestyle modification designed to stabilize or reverse heart disease and ascertain molecular pathways that are important in the development of coronary atherosclerosis.

http://dx.doi.org/10.1016/j.gdata.2015.03.001

Study participants

Inclusion criteria for participation included: 1) adult 21 + years of age, 2) mentally competent to provide informed consent and accurately report adherence, 3) physician diagnosis of coronary artery disease (CAD), which included stable angina, angioplasty, >50% luminal narrowing on coronary angiogram, acute myocardial infarction, bypass surgery, or stent placement, or 2 + CAD risk factors such as obesity (BMI >30), hypertension (systolic pressure >140 mm Hg or diastolic pressure >90 mm Hg), high total cholesterol (>200 mg/dL), diabetes, or family history of heart disease in parents or siblings, 4) approval from personal physician, 5) desire to pursue intensive lifestyle modification as an alternative to, or in conjunction with, standard therapy and motivation to follow the program guidelines for one year, and 6) successful abstinence from smoking for at least three months prior to and during enrollment.

Exclusion criteria were: 1) <21 years of age, 2) presence of unstable coronary syndromes, refractory congestive heart failure, uncontrolled arrhythmia, or high-grade uncorrected cardiac conduction abnormalities, 3) significant left main stenosis (>50%) and ejection fraction <35% in patients who did not have revascularization or were not candidates for revascularization, 4) hypotensive response to exercise, 5) known history of autoimmune disease or systemic/chronic disease requiring chemotherapy or long term treatment, 6) history of substance abuse (including alcohol) without self-certification of abstinence for at least three months, and 7) physical disabilities or medical conditions that would preclude program adherence.

Non-intervention controls were recruited prospectively and matched to program participants based on age (within ±5 years), gender, and disease status (presence of CAD or diabetes mellitus) [1]. Control subjects received only standard care from their primary care physician, did not receive any advice, counseling, or information regarding healthy lifestyle behaviors, and did not participate in any component of the lifestyle intervention.

Intervention

A prospective nonrandomized trial based on the Multicenter Lifestyle Demonstration Project was designed to stabilize or reverse progression of heart disease through comprehensive changes in lifestyle [2]. Participants were recruited by referral from physicians and through advertisements in the media. The lifestyle intervention consisted of a low-fat vegetarian diet (<10% of calories from fat), 180 min/week of moderate aerobic exercise, 1 h of stress management each day, and weekly group support sessions. The year-long program was divided into 2 stages, consisting of an intensive 3-month intervention during which participants were taught to adopt and strictly adhere to the program guidelines followed by a 9-month primarily self-directed maintenance phase.

Clinical information was collected by review of medical records, standard questionnaires, and physical examinations at the baseline, 3-month, and 1-year time points. Demographic and lifestyle factors included the following: age, gender, ethnicity, family history of disease, medication use, various psychometric parameters, and daily caloric intake. Clinical information encompassed: height and weight, systolic and diastolic blood pressure, general endurance, standard lipid panel, lipoprotein profiles, and plasma biomarkers including C-reactive protein, ultra-sensitive insulin, and leptin.

This research was conducted in accordance with the Code of Ethics of the World Medical Association. Participants and controls volunteered to participate in the research study and provided a written informed consent. All research activities were governed by the United States Army Medical Research and Materiel Command (MRMC)/Telemedicine and Advanced Technology Research Center (TATRC) and the Henry M. Jackson Foundation for the Advancement of Military Medicine. Our data reporting followed recommendations of the Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) group [3].

Blood collection, RNA isolation, and microarray analysis

Peripheral blood was collected from participants and controls at each time point using the PreAnalytiX PAXgene™ Blood RNA System (Qiagen, Valencia, CA). Blood was placed at room temperature for 4–24 h and frozen at −80 °C. PAXgene™ tubes were thawed overnight at room temperature and RNA isolation was performed using the PAXgene™ blood RNA Kit. Globin mRNA transcripts were depleted from a portion of the total RNA using the GLOBINclear™-Human Kit (Life Technologies, Carlsbad, CA). RNA quality was assessed with a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA), and quantity was measured with the Nanodrop ND-1000 spectrophotometer (Thermo Scientific, Waltham, MA). One microgram of globin-depleted RNA was then amplified using the MessageAmp™ II aRNA Amplification System (Life Technologies). Resulting double-stranded cDNA was purified, amplified, and labeled with biotin-11-UTP. Labeled aRNA (15 μg) was subsequently fragmented and hybridized to GeneChip® Human Genome U133A 2.0 arrays (Affymetrix, Santa Clara, CA) and scanned on a GeneChip® Scanner 3000 using standard Affymetrix protocols. All 3 time points for each participant/control were processed together to minimize technical artifact. Further details of RNA isolation and gene expression analysis are available in the Data Supplement of Ellsworth et al. [4].

Quality control analysis

All CEL files (n = 480) were subjected to pre-processing using the Robust Multichip Algorithm (RMA). Probe set intensities were obtained by RMA background correction, quantile normalization, median polish summarization, and $\log_2$ transformation. To assess data integrity, evaluate assay performance, and ensure suitability for analysis, the processed intensity data was subjected to standard GeneChip® quality control parameters: background intensity, raw noise (Q) values, percent present calls, scaling factors, and GAPDH 3′/5′ ratio, and Actin 3′/5′ ratio. In addition, the following QC assessments were conducted: array image analysis to identify artifacts on the array surface, distribution analysis to assess the spread of the data relative to the full probe set, and principal component analysis to summarize overall variance.

Arrays included in the final dataset passed the recommended GeneChip® quality control assessments. The RMA normalized $\log_2$ intensity plot showed consistency of individual arrays relative to the entire dataset (Fig. 1). Principal Component Analysis identified limited variability attributable to laboratory procedures across all arrays (Fig. 2). Comparable percent present values (median = 59.2%, range 48.1–64.8%), assessed using the mean absolute deviation, were observed for all samples (Fig. 3).

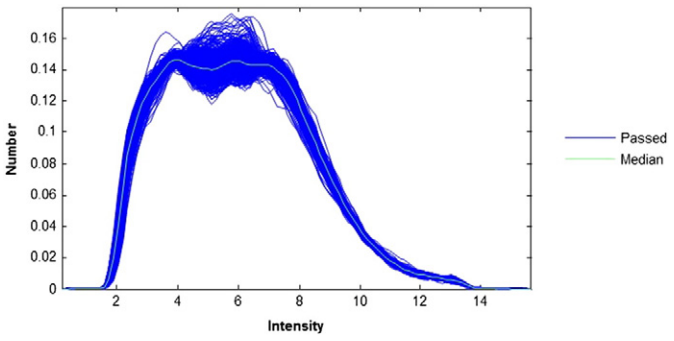

Fig. 1. Intensity graph showing the RMA normalized $\log_2$ intensity for each array. The median intensity curve is highlighted in green.