Published in final edited form as:

Med Sci Sports Exerc. 2016 December ; 48(12): 2387–2397. doi:10.1249/MSS.0000000000001041.

PDPR Gene Expression Correlates with Exercise-Training Insulin Sensitivity Changes

Matthew D. Barberio¹, Kim M. Huffman^{3,4}, Mamta Giri¹, Eric P. Hoffman^{1,6}, William E. Kraus^{2,4}, and Monica J. Hubal, PhD^{1,5,6}

¹Research Center for Genetic Medicine, Children's National Medical Center, Washington, DC

²Division of Cardiology, Duke University School of Medicine, Durham, NC

³Division of Rheumatology, Department of Medicine, Duke University School of Medicine, Durham, NC

⁴Division of Duke Molecular Physiology Institute, Duke University School of Medicine, Durham, NC

⁵Department of Integrative Systems Biology, George Washington University, Washington, DC

⁶Department of Exercise and Nutrition Sciences, George Washington University, Washington, DC

Abstract

Purpose—Whole body insulin sensitivity (S_i) typically improves following aerobic exercise training; however, individual responses can be highly variable. The purpose of this study was to use global gene expression to identify skeletal muscle genes that correlate with exercise-induced S_i changes.

Methods—Longitudinal cohorts from the Studies of Targeted Risk Reduction Intervention through Defined Exercise (STRIDE) were utilized as Discovery (Affymetrix) and Confirmation (Illumina) of *vastus lateralis* gene expression profiles. Discovery ($n=39$; 21 men) and Confirmation ($n=42$; 19 men) cohorts were matched for age (52 ± 8 vs. 51 ± 10 yr), BMI (30.4 ± 2.8 vs. 29.7 ± 2.8 kg·m⁻²), and VO_{2max} (30.4 ± 2.8 vs. 29.7 ± 2.8 mL/kg/min). S_i was determined via intravenous glucose tolerance test pre- and post-training. Pearson product-moment correlation coefficients determined relationships between a) baseline and b) training-induced changes in gene expression and % S_i after training.

Results—Expression of 2454 (Discovery) and 1778 genes (Confirmation) at baseline were significantly ($P<0.05$) correlated to % S_i ; 112 genes overlapped. Pathway analyses identified Ca²⁺-signaling-related transcripts in this 112-gene list. Expression changes of 1384 (Discovery) and 1288 genes (Confirmation) following training were significantly ($P<0.05$) correlated to % S_i ; 33 genes overlapped, representing contractile apparatus of skeletal and smooth muscle genes.

Corresponding Author: Monica J Hubal, PhD, Assistant Professor, Dept. Integrative Systems Biology, Visiting Assistant Professor, Dept. Exercise Science and Nutrition, George Washington University, 950 New Hampshire Ave NW, Washington, DC 20052, mhubal@gwu.edu, Telephone: 202-994-1408.

Competing Interests: We report no competing or conflicts of interests. The results of the present study do not constitute endorsement by ACSM.

Pyruvate dehydrogenase phosphatase regulatory subunit (PDPR) expression at baseline ($p=0.01$, $r=0.41$) and post-training ($p=0.01$, $r=0.43$) were both correlated with % S_i .

Conclusion—Exercise-induced adaptations in skeletal muscle S_i are related to baseline levels of Ca^{+2} -regulating transcripts, which may prime the muscle for adaptation. Relationships between % S_i and PDPR, a regulatory subunit of the pyruvate dehydrogenase complex, indicate that the S_i response is strongly related to key steps in metabolic regulation.

Keywords

microarray analysis; metabolic flexibility; transcriptomics; insulin resistance; endurance exercise

Introduction

Insulin resistance (decreased insulin sensitivity) is a hallmark of pre-diabetes and Type 2 diabetes mellitus (T2DM) and can be a precursor to cardiovascular disease (8). Although routine physical activity ameliorates insulin resistance through enhanced skeletal muscle insulin sensitivity (11, 13, 16), the precise molecular mechanisms driving the adaptation remain unclear. Moreover, participants in randomized controlled exercise training interventions undergoing the same protocol exhibit a wide range of insulin sensitivity responses. For example, in a large family-based study after 20 weeks of aerobic exercise training, approximately 30% of participants did not improve insulin sensitivity (4). Exercise interventions commonly produce phenotypic responses with large heterogeneity; this heterogeneity has been attributed to participant differences in skeletal muscle adaptations (1, 27), aerobic capacity (35), and various cardiovascular risk factors (3). Thus, to individualize exercise prescriptions optimally, it is important to understand the mechanisms driving phenotypic response variation.

Insulin sensitivity is a complex trait that is influenced by both intrinsic (inherited) and extrinsic (environmental) factors (22, 39). Heritability estimates for insulin action on glucose homeostasis in individuals of European ancestry range between 25%-44% (31). Genetic variants account for small to modest amounts of trait variability (29). Examples of genetic variants associated with insulin sensitivity include: Cordon-Bleu WH2 Repeat Protein-Like 1 (COBLL1; rs7607980; T>C); Insulin Receptor Substrate 1 (IRS1; rs2943634; C>A); and Platelet Derived Growth Factor C (PDGFC; rs4691380; A>G) (22).

Environmental factors such as diet, physical activity patterns, and aging significantly influence insulin sensitivity (10). Numerous biochemical and physiologic mechanisms are implicated in exercise training-induced improvements in insulin sensitivity; these include increased glucose transport via glucose transporter type 4 (GLUT4) gene and protein expression (25), improved mitochondrial function (23), and improved skeletal muscle lipid handling (23).

While both intrinsic and extrinsic factors contribute to insulin resistance, the molecular interplay between the various contributing factors is not well understood. Here, we use an exercise training study cohort, in which both insulin sensitivity and gene expression microarray data are assessed longitudinally, to begin to probe the mechanisms at the gene

expression level that might be responsible for exercise-induced changes — and the variations in responses — in insulin action. In the Studies of Targeted Risk Reduction Intervention through Defined Exercise (STRRIDE) trials, three- to four-times weekly aerobic exercise training resulted in significant insulin sensitivity changes, regardless of exercise intensity or volume (16). STRRIDE participants completed eight months of structured endurance training, resulting in an average 56% improvement in whole body insulin sensitivity, similar in magnitude to responses seen in other studies (4, 9). Despite this robust average improvement in insulin sensitivity, a considerable amount of heterogeneity (range: -87% - 495%) in S_i changes was observed (17). Global skeletal muscle gene expression datasets from pre- and post-training time points afforded us the opportunity to dissect the S_i change trait in the STRRIDE cohort at the molecular level.

In this study, we utilized two longitudinal subcohorts from the STRRIDE clinical exercise trials to analyze relationships between skeletal muscle gene expression and changes in skeletal muscle insulin sensitivity with training. Furthermore, we utilized biological pathway analysis to connect significant exercise-induced changes in gene expression with biological function to better understand the molecular pathways involved improvements in skeletal muscle insulin sensitivity and the individual variation thereof.

Materials and Methods

Study Design

The current study utilizes subjects and samples from the Studies of Targeted Risk Reduction Interventions through Defined Exercise (STRRIDE) trials (Figure 1), where all subjects completed eight months of supervised exercise training (19). We studied individuals from STRRIDE aerobic exercise groups with available pre- and post-exercise clinical data and intravenous glucose tolerance tests (N=319). From this group, two subcohorts (matched for race, age, and sex) with phenotype responses representative of the 319 subject cohort (Table 1) were selected for global mRNA analysis of skeletal muscle biopsies using transcriptomic microarrays from Affymetrix (n=39) and Illumina (n=42). The Affymetrix cohort served as the “Discovery Cohort”, while the Illumina cohort served as the “Confirmation Cohort”. There was significant overlap in the subjects between these subcohorts (n=28), for whom both platforms were run. For validation purposes, we analyzed the n=14 independent participants in our Confirmation cohort in a separate statistical analysis and provide statistical outcomes for this cohort when appropriate.

STRRIDE Subjects

In STRRIDE I (NCT00200993), sedentary (<1 day·wk⁻¹ exercise), overweight or obese (BMI 25-35 kg/m²) individuals with mild to moderate lipid abnormalities (LDL cholesterol between 130 – 190·mg·dL⁻¹ or HDL cholesterol < 45 mg·dL⁻¹ for woman and 40 mg·dL⁻¹ for men), and between the ages of 40 to 65 years, were randomized into one of three endurance exercise or control groups. In STRRIDE II (NCT00275145), a similar cohort of individuals, aged 18 to 70 years, was randomized into one of three endurance, resistance, or combined exercise or control groups. Subjects were excluded if they were currently using medications that could alter carbohydrate metabolism, blood pressure, or presented with

evidence of diabetes, heart disease, or orthopedic injury that impeded or prohibit unsupervised exercise.

In total, 380 (STRIDE I) and 249 (STRIDE II) subjects met inclusion criteria and were randomized into groups after providing written informed consent as approved by the Investigational Review Boards at Duke University and East Carolina University. Participants were counseled to maintain their normal dietary habits throughout the entirety of the study. To verify this, a three-day food record and 24-h dietary recall were assessed at baseline, mid-intervention, and post-intervention (Table 1).

Exercise Training

Subjects were randomized into exercise training groups which have been described in detail elsewhere (19). The exercise training programs were (volume, intensity): 1) low-volume/moderate intensity group = caloric equivalent of ~12 miles/wk, 1,200 kcal/wk at 40-55% $\text{VO}_{2\text{peak}}$ (MILD); 2) low-volume/high-intensity = caloric equivalent of ~20 miles/wk, 1,200 kcal/wk at 65-80% $\text{VO}_{2\text{peak}}$ (MOD); 3) high-volume/high-intensity group = caloric equivalent of ~20 miles/wk, 2,000 kcal/wk at 65-80% $\text{VO}_{2\text{peak}}$ or; 4) combined aerobic/resistance training group = caloric equivalent to 12 miles/wk at 75% $\text{VO}_{2\text{peak}}$ plus 3 days/wk, 3 sets/day, 8-12 repetitions/set, 8 total exercises (AT-RT). All exercise sessions were verified by direct supervision or the use of recorded heart rate monitoring (Polar Electro, Woodbury, NY). Caloric equivalents were determined by the approximate energy expenditure during walking or jogging for a 90 kg person, however multiple exercise modalities (cycle ergometers, treadmills, and elliptical trainers) were utilized for endurance training. A ramp period of 8 to 10 weeks (STRIDE I) and 12 weeks (STRIDE II) was utilized to familiarize individuals with the exercise regimen and minimize musculoskeletal injury followed by six months of the appropriate structured exercise protocol.

Insulin Sensitivity

Clinical insulin sensitivity measures were analyzed prior to and at completion of the structured exercise interventions via intravenous glucose tolerance test (IVGTT). A detailed description of the procedures is published elsewhere (19). Insulin sensitivity index (S_i ; $\text{mU}\cdot\text{l}^{-1}\cdot\text{min}^{-1}$) was determined by minimal model following a three-hour IVGTT. Although S_i represents whole body insulin sensitivity, in reality it is mostly reflective of peripheral skeletal muscle insulin sensitivity. Fasting samples were obtained and a 50% glucose solution was injected via an antecubital catheter at $0.3\text{ g}\cdot\text{kg}^{-1}$ body mass. At minute 20, a $0.025\text{ U}\cdot\text{kg}^{-1}$ body mass of insulin was injected. Blood samples were collected at 2, 3, 4, 5, 6, 8, 10, 12, 14, 16, 19, 22, 25, 30, 40, 50, 60, 70, 80, 90, 100, 120, 140, and 180 minutes. Insulin was determined via immunoassay (Access Immunoassay System, Bechman Coulter, Fullerton, CA) and glucose was analyzed with an oxidation reaction (YSI model 2300 Stat Plus, Yellow Springs Instruments, Yellow Springs, OH). A greater S_i indicates enhanced insulin sensitivity. Pre- and post-exercise intervention S_i values were used to determine percent change in S_i (% S_i ; $(\text{Post-exercise } S_i - \text{Pre-exercise } S_i) / \text{Pre-exercise } S_i \times 100$) for use in correlation analyses.

Skeletal Muscle Biopsies

Skeletal muscle biopsies were taken pre- and post-exercise training from the *vastus lateralis* muscle of the non-dominant leg via standard Bergstrom needle biopsy. Approximately 100 to 200 mg of tissue was obtained with a triple pass of the needle upon entry into the study (Pre) and after eight months of exercise training (Post; 24 h following the last bout of exercise). Biopsies were taken while the subject was in a fasted state. Tissues were flash frozen and approximately 30 to 50 mg was used for each expression profiling platform.

Gene Expression Analyses

Skeletal muscle biopsies from each of the two subcohorts were analyzed using global gene expression analyses. We used the two predominant platforms for global gene expression analysis currently available: Affymetrix Genechips and Illumina microarrays. These platforms use different approaches to generate global mRNA data; therefore, use of both platforms represented an opportunity for global confirmation of discovery microarray results. Description of gene expression analysis for each cohort is described below.

Gene Expression Analysis: Discovery (Affymetrix)—Skeletal muscle gene expression was analyzed from pre- and post-exercise training biopsies via global microarray analysis (Affymetrix Hu133 Plus 2.0 microarray; Affymetrix, Santa Clara, CA; Accession: GSE48278). Briefly, total RNA was extracted from 30 to 50 mg of starting material using Trizol (Invitrogen, Carlsbad, CA). Following two rounds of amplification, via Affymetrix kits, a total of 30µg of biotinylated cRNA from the second amplification was hybridized to the microarrays. CEL files were imported into Affymetrix Expression Console, where CHP files were generated using the PLIER (Probe Logarithmic Intensity Error) algorithm. PLIER is a model-based signal estimator beneficial to multi-array estimations. Standard quality control measures for adequate amplifications, thresholds for appropriate scaling factors, and RNA integrity (GAPDH 3'/5' and HSAC07 3'/5') were assessed (12). Samples not meeting quality control standards were reprocessed from original total RNA after RNA integrity was verified by agarose gel electrophoresis and imaging. CHP files generated from the PLIER algorithm were then imported into Partek Genomics Suite (Partek, Inc.; St. Louis, MO). Probe set intensities (PLIER) were log₂-transformed for gene expression data normalization.

Gene Expression Analysis: Confirmation (Illumina)—Biotinylated total RNA was generated using the Illumina TotalPrep RNA amplification kit (Life Technologies, Grand Island, NY, USA); 200 ng of RNA were used for the kits. The quality of the RNA was determined using the Bioanalyzer RNA Nanochip assay (Agilent, Santa Clara, CA, USA). Quantification of the RNA was determined using the Quant-iT RiboGreen RNA Assay Kit. The Human HT-12v3 Expression BeadChip (Illumina, San Diego, CA; Accession: Pending) was used for quantitative whole genome RNA profiling. Biotinylated RNA (750 ng) was hybridized to the BeadChip and washed; Cy3-SA was then introduced to the hybridized samples and the BeadChips scanned on the Illumina iScan system according to manufacturer's protocol.

Quality control was performed using the Illumina GenomeStudio tools. Initial quality control was done using Illumina's recommended guidelines. Data were transformed and

normalized using bioconductor (www.bioconductor.org) lumi package in R statistical environment. Data was transformed using a combination of four different normalization methods – quantile normalization, robust spline normalization, rank invariant and simple scaling normalization, and two different transformation methods: $\log_2$ and variance stabilization transformation. Unsupervised clustering and principal component analysis were performed to choose the normalization method that produced the least background noise. Based on the biological and technical replicate grouping, technical outliers were removed and variance stabilized. The robust spline normalized dataset was chosen for downstream analyses.

Statistical Analyses

Normality of data was assessed with the Shapiro-Wilk test and visualization of the distribution. If data were non-normally distributed, the data were $\log_2$ -transformed and reassessed for normality. Differences in pre- and post-exercise values were analyzed with a repeated measures ANOVA (Cohort * Time). Differences in % S_i ((post-exercise S_i – pre-exercise S_i)/pre-exercise S_i) and exercise adherence (percentage) among cohorts was determined with a 1-Way ANOVA. To maintain independence of cohorts for statistical analysis, subjects represented in the Discovery and Confirmation sub-cohorts were removed from the STRRIDE I & II cohorts for analysis of outcomes. No subject was represented in multiple groups for demographic or clinical comparisons. Significance for phenotypic data was determined *a priori* as $p \leq 0.05$. When appropriate, pairwise comparisons were made using the Tukey test. Statistical analysis was performed using OriginLab Pro 9.1 (OriginLab Corp, Northampton, MA).

To analyze the relationship between pre-exercise gene expression and % S_i , pre-exercise, probe set intensities were correlated to % S_i via Pearson product-moment correlation analysis. To analyze the relationship between change in gene expression and % S_i , pre-exercise intervention probe set intensities were subtracted from post-exercise intervention probe set intensities using normalized data. The remaining gene expression value was used in correlation analysis to determine significant relationships to % S_i . For both analyses, Pearson product-moment correlation analysis was performed in Partek Genomics Suite. Significant ($p < 0.05$) genes represented in both the Discovery and Confirmation (Figure 1) datasets were then analyzed for biological interpretation, as described below. Unadjusted p-values and Pearson product-moment correlation coefficients presented in text are from the Discovery cohort; values for the Confirmation cohorts are provided in supplemental tables. For the determination of the overlapping significantly correlated genes, it is important to note that we chose to use a more lenient significance cutoff at this step because the use of global Confirmation and downstream pathway analyses both significantly reduce the chance of reporting false positive data.

Biological Pathway Analysis

Ingenuity Pathway Analysis (IPA; Qiagen; Germantown, MD) was utilized for probe set annotations and to query biological interpretation of gene sets related to S_i response to exercise training. IPA utilizes a curated literature-based database in which differentially

expressed genes from large data sets can be linked to biological processes, molecular networks, and biochemical function.

Results

Clinical Outcomes

Following eight months of aerobic exercise training, average % $\dot{V}_{O_2}$ increased by $44\% \pm 62$ (Discovery) and $61\% \pm 88$ (Confirmation) ($p < 0.01$); these responses were similar to the total STRRIDE cohort ($46\% \pm 84$). 1-Way ANOVA detected no differences ($p = 0.36$) among cohorts for % $\dot{V}_{O_2}$ and exercise adherence ($p = 0.20$; Table 1). RM ANOVA indicated that exercise significantly increased $\dot{V}_{O_2}$ ($p < 0.01$), increased $\dot{V}_{O_2}$ ($p < 0.01$), decreased fasting insulin ($p < 0.01$). The Discovery and Confirmation cohorts did not differ in cardiorespiratory fitness, body mass, or blood glucose homeostasis either at baseline or following exercise training. For baseline and post-training blood glucose homeostasis, body mass, and cardiorespiratory fitness, there were no significant relationships with % $\dot{V}_{O_2}$.

PDPR Expression at Baseline and Training-Induced Changes Correlate to % $\dot{V}_{O_2}$

The most consistent % $\dot{V}_{O_2}$ relationship was with pyruvate dehydrogenase phosphatase regulatory subunit (PDPR) gene expression. We first noted a strong relationship between PDPR baseline gene expression ($p=0.01$, $r=0.41$; Table 1) and % $\dot{V}_{O_2}$. Our analysis indicated one probe in the Discovery (Figure 2A) and two probes in Confirmation and Independent cohorts with similar significance and correlation coefficients (Table 2). We also found a significant relationship between the exercise-induced change in expression of PDPR ($p=0.01$, $r=0.43$) and % $\dot{V}_{O_2}$. We identified two probes in both Discovery (Figure 2B) and Confirmation (Figure 2C) cohorts where the magnitude of significance was similar and correlation coefficient direction was identical. These indicate a strong relationship for gene expression of this key metabolic regulatory protein. It should be noted that other genes were represented in both baseline and training-induced change in gene expression analysis, including BLVRA, BMP1, and ERC1 (Supplemental Table 3 and 6).

Baseline Gene Expression Correlated with % $\dot{V}_{O_2}$

Pearson product-moment correlation analyses identified 2454 (Discovery; Affymetrix) and 1778 (Confirmation; Illumina) genes for which baseline expression significantly ($P < 0.05$) related to % $\dot{V}_{O_2}$. Full lists of significant genes for Discovery (see Table, SDC 1, Baseline GEx In Discovery cohort) and Confirmation (see Table, SDC 2, Baseline GEx In Confirmation cohort) can be found in the supplemental material. 112 genes were represented in both the Discovery and Confirmation analyses, with 108 having the same relationship direction (see Table, SDC 3, 112 Confirmed Genes At Baseline). Table 2 lists the 21 genes where baseline expression most strongly related to % $\dot{V}_{O_2}$.

Pathway analysis mapped 89 of the 108 genes significantly correlated with % $\dot{V}_{O_2}$. Using IPA's functional categorization tool (which groups genes of interest together based on their major functional role within the cell), we identified Cellular Function and Maintenance as the top molecular classification. Included in this list were the skeletal muscle contraction and Ca^{2+} -regulating genes: calcium channel; voltage-dependent, L type, alpha 1S subunit

(CACNA1S; $p=0.006$, $r=0.42$); calcium/calmodulin-dependent protein kinase II delta (CaMK2D; $p=0.03$, $r=0.33$), sodium-channel, voltage gated, type IV beta subunit (SCN4B; $p=0.008$, $r=0.41$); Annexin A7 (ANXA7; $p=0.007$, $r=0.42$); and myosin light chain kinase 2 (MYLK2; $p=0.003$, $r=0.47$). Also, basal levels of the transcriptional factors hepatocyte nuclear factor 4; alpha (HNF4 α ; $p=0.03$, $r=0.4$); and GATA binding protein 4 (GATA4; $p=0.009$, $r=-0.41$) had significant relationships to % S_i .

We also used the IPA network analysis tool (which explores the relationship between the genes of interest and previously established genes in the literature) to identify related groups of genes. We found gene networks related to contractile activity, transcription regulation, and molecular transport represented among the 108 S_i -correlated genes. The top IPA network (score = 45; 22/35 genes; Figure 3) included the contraction-related genes (CaMK2D and CACNA1S); and transcription regulation (GATA4).

Training-Induced Gene Expression Changes Correlated with % S_i

Expression changes in 1384 (Discovery) and 1288 (Confirmation) genes were significantly ($p < 0.05$) related to % S_i . Full lists of significant genes for Discovery (see Table, SDC 4, Change In GEx In Discovery cohort) and Confirmation (see Table, SDC 5, Change In GEx In Confirmation Cohort) can be found in the supplemental material. Thirty-three genes were found in common with the Discovery and Confirmation analyses; of these, 27 genes had correlation coefficients in the same direction in both analyses (see Table, SDC 6, 27 Confirmed Genes Following Exercise Intervention). Table 3 shows the genes whereby training-induced expression changes were most positively and negatively related to S_i change.

For those 27 genes, biological pathway analysis mapped 23. The top functional category was Cellular Compromise, containing the genes MYH11 ($p=0.03$, $r=0.34$), MYLK ($p=0.03$, $r=0.34$), and DNA excision repair protein (ERCC1; $p=0.04$, $r=0.33$). The next highest ranked function, Cellular Function and Maintenance, was largely made up of the genes listed under Cellular Compromise as well as Huntingtin-associate protein-interaction protein (KALRN; $p=0.01$, $r=-0.40$); Fibronectin Type III Domain Containing 3B (FNDC3B; $p=0.03$, $r=-0.35$); synaptogamin 7 (SYT7; $p=0.04$, $r=-0.32$), and developmental pluripotency-associated 4 (DPPA4; $p=0.05$, $r=0.33$).

Discussion

Aerobic exercise training is essential to mitigating Type 2 diabetes mellitus (28) and cardiovascular disease risk (38). Understanding the molecular drivers of beneficial training-induced skeletal muscle adaptations may serve to identify new targets for therapy and provide novel insight into the interaction of various cellular processes prior to and during adaptation. This is the largest study to assess skeletal muscle gene expression profiles prior to and following aerobic exercise training interventions to identify genes and gene networks specifically related to changes in insulin sensitivity. A major strength of our approach was the use of dual global gene expression profiling platforms in STRRIDE subcohorts; serving as global Discovery and Confirmation groups. Using this clinically well-phenotyped cohort,

we leveraged the wide variability in clinical outcomes of large exercise training cohort studies (3), combined with unbiased global gene expression profiling.

PDPR's Relationship to % S_i

PDPR gene expression at baseline (Figure 2A) and after training (Figure 2B-C) were both related to % S_i. This gene encodes the regulatory subunit of pyruvate dehydrogenase phosphatase, which serves a critical role in the activation of the pyruvate dehydrogenase (PDH) complex. Though central to glucose oxidation and fatty acid synthesis in response to insulin stimulation, little is known about the role of PDH complex, and its regulation, in the pathobiology of skeletal muscle insulin resistance. There is increased insulin-stimulated activation of the PDH complex following exercise training (24). Insulin resistance is associated metabolic inflexibility (18), where the shift from predominately lipid oxidation and fatty acid uptake to insulin-stimulated glucose oxidation is blunted. Thus, finding of a strong relationship between % S_i and PDPR expression at baseline and following exercise training, which may result in a higher activation of the PDH complex, provides a potential mechanistic role for how the PDH complex influences S_i. Further exploration of PDPR's role in exercise-induced adaptation will be of significant interest to understanding the role of metabolic flexibility in skeletal muscle health.

Ca²⁺-Signaling Related Baseline Transcripts Related to % S_i

Skeletal muscle gene expression profiles prior to exercise training (baseline) allows for the identification of genes that potentially prime skeletal muscle for exercise-induced adaptations. Our results suggest that exercise training-induced changes in S_i are strongly related to baseline expression of genes involved in skeletal muscle contraction and Ca²⁺ handling. The CaMK2D and MYLK2 genes code for Ca²⁺-dependent enzymes activated upon excitation-contraction coupling. MYLK2 is a sarcomeric protein that co-localizes with the phosphorylated regulatory myosin light chain. Ablation of MYLK2 in mice alters twitch responses to electrical stimulation (40), while polymorphisms in this gene are associated with exercise-induced muscle damage and loss of strength after eccentric exercise (5). ANXA7 is another Ca²⁺-regulating gene; located in the t-tubule and plasma membrane of striated muscle and believed to play a role the excitation-coupling mechanism of skeletal muscle contraction (2). CACNA1S encodes for the alpha 1S subunit of the L type voltage-dependent calcium channel found in the ryanodine receptor RyR1. The role of Ca²⁺ in cellular energy metabolism and the contractile process in skeletal muscle is well understood. Thus, skeletal muscle more adept at Ca²⁺ handling may in fact prime the muscle for training-induced adaptations that can occur earlier and/or to a greater extent. Given the variability of the exercise response, these genes may serve as a baseline predictor of exercise responsiveness and should be further explored for this potential.

The relationship between baseline gene expression for solute carrier family 25 (carnitine/acylcarnitine translocase), member 20 (SLC25A20) and % S_i further supports the role of baseline Ca²⁺ signaling gene expression and change in insulin sensitivity. Over-expression of the calcium-regulated serine/threonine protein phosphatase, calcineurin results in significant up-regulation of lipid-metabolism regulating genes, including SLC25A2, and increased skeletal muscle lipid metabolism (21). Additionally, more SLC25A20 expression

implies more metabolic potential for lipids, the primary energy source for aerobic exercise. Reduced SLC25A20 mRNA and protein content correlated with reduced acylcarnitine mitochondrial flux in insulin resistant muscle (26). Last, increased acylcarnitine flux – as a result of higher SLC25A20 expression – could prevent cytosolic accumulation of lipids, which is suspected to play a crucial role in metabolic inflexibility (18).

Transcriptional Regulation Transcripts at Baseline Related to % S_i

Transcripts involved in transcriptional regulation were also found to be related at baseline to % S_i following exercise training. HNF4 α is a nuclear transcription factor that interacts with the regulatory units of genes involved in glucose metabolism, lipid metabolism, and mitochondrial function (20, 34). Recent global analysis demonstrates differential methylation of the HNF4 α promoter region among diabetic and non-diabetic twins (30). GATA4 is GATA family zinc finger transcription factor that works in conjunction with calcineurin-dependent nuclear factor of activated T cells (NFAT) (32). Originally believed to be a cardiac-specific transcription factor, GATA4 is expressed in skeletal muscle and critical to early skeletal muscle myogenesis (6). Network analysis (Figure 3) showed GATA4 downstream of Ca²⁺-dependent enzyme CaMK2D, and classic signaling pathways NF κ B and Vascular Endothelial Growth Factor (VEGF). Continued expression of these genes in adult skeletal muscle may be a mechanism whereby skeletal muscle responds rapidly to changes in contractile activity (6).

Training-Induced Changes in Gene Expression of Contractile Apparatus Is Correlated with % S_i

Our analysis identified changes in 27 genes following exercise training with significant relationship to % S_i. These genes are involved in a variety of molecular functions that promote adaptation in skeletal muscle in response to various stimuli, including aerobic exercise training (7). MYLK (also referred to as MLCK) is a ubiquitous Ca²⁺/calmodulin-regulated enzyme responsible for the initiation of contraction through phosphorylation of the regulatory light chain (14). Biological network analysis (not shown) places MYLK as a central node that is upstream of the classic NF κ B inflammatory signaling pathway. The role of NF κ B signaling in peripheral S_i and cellular energy metabolism is well studied (33). Our analysis indicates this role may also include regulation by the contractile functions of skeletal muscle, potentially linking repeated bouts of contraction (exercise) to inflammation and S_i.

We also found a significant correlational relationship of % S_i with MYH11 expression, which is a smooth muscle myosin belong to the myosin heavy chain family also responsible for the initiation of contraction. Skeletal muscle is heterogeneous tissue including many cell types (skeletal, vascular, inflammatory, etc.) and we cannot rule out that presence of these cell and tissue types in our samples. Our findings of a significant relationship between the changes in expression of both MYLK and MYH11 and % S_i may be indicative of changes in the contractile apparatus of both skeletal and smooth muscle within the skeletal muscle architecture.

% S_i Correlated-Genes Associated with CVD Risk

Analysis of baseline and training-induced changes in gene expression resulted in significant relationships between exercise-induced changes in S_i and genes previously linked to CVD and CVD risk factors. The intronic variant rs9289231 in the KALRN (Table 3) gene has previously been linked (36) and validated (15) to the early-onset of CAD. Significant relationship between the genetic variant rs16834817 in MYLK and coronary artery disease was also established in this cohort. The rs745975 C>T HNF4 α (identified in baseline gene expression analysis; Supplemental Table 3) SNP is associated with S_i and glucose tolerance; both associations are modified by physical activity (31). This common variant has also previously been linked to serum lipid levels (37). While the role of these genes in skeletal muscle and improvements in skeletal muscle S_i require further investigation, the previously established link between CVD and genetic variants in these genes along with our findings are interesting because insulin resistance is one of the strongest predictors of CVD developments. Perhaps this is indicative of a role for these genes in the well-established cardiovascular benefits of regular exercise training.

Limitations

While these results are intriguing, we acknowledge some limitations in our analyses. First, we did not address the differential effects of the exercise programs on changes in the genes of interest. In the original STRRIDE investigations, exercise amount was a significant factor in promoting greater changes in S_i (all exercise groups significantly improved) (16). Moving forward, it will be important to understand how exercise regimens may differentially alter not only specific cardiometabolic risk factors, but also the molecular processes that drive these changes at the cellular level. Second, our Confirmation cohort combines 28 subjects that completely overlap with the Discovery cohort; further, there are 14 Independent subjects, for a total of 42 subjects. Here, we posit that comparing the results from the Discovery cohort to the full 42-subject Confirmation cohort is the most powerful approach for the cross-platform confirmation of S_i-related genes in our study. We believe is less important to have a completely overlapping Confirmation dataset than a more robust combined cohort; there is typically high fidelity across platforms. We have broken out the 14 Independent subjects in several Tables within the manuscript to show how each gene from the Discovery cohort performs in an independent validation set.

Conclusion

Chronic aerobic exercise results in significant improvements in cardiometabolic health and significantly decreases risk for chronic diseases such as CVD and T2DM. However, large clinical exercise trials have resulted in a wide variability in improvements in cardiometabolic risk phenotypes (VO_{2max}, blood lipid profiles, and S_i). Using a well-defined clinical exercise-training cohort and an unbiased global gene expression profiling of skeletal muscle, we identified relationships between: 1) baseline gene expression and exercise-induced change in whole body insulin sensitivity; and 2) exercise-induced changes in gene expression and exercise-induced changes in whole body insulin sensitivity. Analysis of both baseline and training-induced changes in gene expression indicates a significant relationship between genes involved in Ca²⁺ regulation and changes in insulin sensitivity following

exercise training. Last, we confirmed three probes at baseline and four probes for training-induced changes for the PDPR gene. The role of this gene in the regulation of the PDH, a flux-limiting step in glucose metabolism, provides an intriguing regulatory gene for further exploration for its role in glucose homeostasis responses to exercise training, particularly insulin sensitivity changes mediated through skeletal muscle insulin action.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Information: This project was supported by the following grants from the National Institutes of Health: 1R01HL57354 (Kraus), F32AR052596 (Hubal), K23 AR0549404 (Huffman) and T32AR056993 (Barberio).

References

1. Bamman MM, Petrella JK, Kim JS, Mayhew DL, Cross JM. Cluster analysis tests the importance of myogenic gene expression during myofiber hypertrophy in humans. *J Appl Physiol*. 2007; 102(6): 2232–9. [PubMed: 17395765]
2. Berchtold MW, Brinkmeier H, Muntener M. Calcium ion in skeletal muscle: its crucial role for muscle function, plasticity, and disease. *Physiol Rev*. 2000; 80(3):1215–65. [PubMed: 10893434]
3. Bouchard C, Blair SN, Church TS, et al. Adverse metabolic response to regular exercise: is it a rare or common occurrence? *PLoS One*. 2012; 7(5):e37887. [PubMed: 22666405]
4. Boule NG, Weisnagel SJ, Lakka TA, et al. Effects of exercise training on glucose homeostasis: the HERITAGE Family Study. *Diabetes Care*. 2005; 28(1):108–14. [PubMed: 15616242]
5. Clarkson PM, Hoffman EP, Zambraski E, et al. ACTN3 and MLCK genotype associations with exertional muscle damage. *J Appl Physiol*. 2005; 99(2):564–9. [PubMed: 15817725]
6. Clause KC, Tchao J, Powell MC, et al. Developing cardiac and skeletal muscle share fast-skeletal myosin heavy chain and cardiac troponin-I expression. *PloS One*. 2012; 7(7):e40725. [PubMed: 22808244]
7. Coffey VG, Hawley JA. The molecular bases of training adaptation. *Sports Med*. 2007; 37(9):737–63. [PubMed: 17722947]
8. DeFronzo RA, Barzilai N, Simonson DC. Mechanism of metformin action in obese and lean noninsulin-dependent diabetic subjects. *J Clin Endocrinol Metab*. 1991; 73(6):1294–301. [PubMed: 1955512]
9. DeFronzo RA, Sherwin RS, Kraemer N. Effect of physical training on insulin action in obesity. *Diabetes*. 1987; 36(12):1379–85. [PubMed: 3315786]
10. Ford ES, Giles WH, Dietz WH. Prevalence of the metabolic syndrome among US adults: findings from the third National Health and Nutrition Examination Survey. *JAMA*. 2002; 287(3):356–9. [PubMed: 11790215]
11. Goodyear LJ, Kahn BB. Exercise, glucose transport, and insulin sensitivity. *Annu Rev Med*. 1998; 49:235–61. [PubMed: 9509261]
12. Group TABPW. Expression profiling--best practices for data generation and interpretation in clinical trials. *Nat Rev Genet*. 2004; 5(3):229. [PubMed: 14970825]
13. Hawley JA, Lessard SJ. Exercise training-induced improvements in insulin action. *Acta Physiol (Oxf)*. 2008; 192(1):127–35. [PubMed: 18171435]
14. Herring BP, Dixon S, Gallagher PJ. Smooth muscle myosin light chain kinase expression in cardiac and skeletal muscle. *Am J Physiol Cell Physiol*. 2000; 279(5):C1656–64. [PubMed: 11029314]
15. Horne BD, Hauser ER, Wang L, et al. Confirmation study of genetic associations with coronary artery disease on chromosome 3q13-21 and potential effect modification by smoking. *Ann Hum Genet*. 2009; 73(Pt 6):551–8. [PubMed: 19706030]

16. Houmard JA, Tanner CJ, Slentz CA, Duscha BD, McCartney JS, Kraus WE. Effect of the volume and intensity of exercise training on insulin sensitivity. *J Appl Physiol.* 2004; 96(1):101–6. [PubMed: 12972442]
17. Huffman KM, Koves TR, Hubal MJ, et al. Metabolite signatures of exercise training in human skeletal muscle relate to mitochondrial remodelling and cardiometabolic fitness. *Diabetologia.* 2014; 57(11):2282–95. [PubMed: 25091629]
18. Kelley DE, Mandarino LJ. Fuel selection in human skeletal muscle in insulin resistance: a reexamination. *Diabetes.* 2000; 49(5):677–83. [PubMed: 10905472]
19. Kraus WE, Torgan CE, Duscha BD, et al. Studies of a targeted risk reduction intervention through defined exercise (STRRIDE). *Med Sci Sports Exerc.* 2001; 33(10):1774–84. [PubMed: 11581566]
20. Ladas JA, Hadzopoulou-Cladaras M, Kardassis D, et al. Transcriptional regulation of human apolipoprotein genes ApoB, ApoCIII, and ApoAII by members of the steroid hormone receptor superfamily HNF-4, ARP-1, EAR-2, and EAR-3. *J Biol Chem.* 1992; 267(22):15849–60. [PubMed: 1639815]
21. Long YC, Glund S, Garcia-Roves PM, Zierath JR. Calcineurin regulates skeletal muscle metabolism via coordinated changes in gene expression. *J Biol Chem.* 2007; 282(3):1607–14. [PubMed: 17107952]
22. Manning AK, Hivert MF, Scott RA, et al. A genome-wide approach accounting for body mass index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat Genet.* 2012; 44(6):659–69. [PubMed: 22581228]
23. Meex RC, Schrauwen-Hinderling VB, Moonen-Kornips E, et al. Restoration of muscle mitochondrial function and metabolic flexibility in type 2 diabetes by exercise training is paralleled by increased myocellular fat storage and improved insulin sensitivity. *Diabetes.* 2010; 59(3):572–9. [PubMed: 20028948]
24. Nakai N, Sato Y, Oshida Y, Fujitsuka N, Yoshimura A, Shimomura Y. Insulin activation of pyruvate dehydrogenase complex is enhanced by exercise training. *Metabolism.* 1999; 48(7):865–9. [PubMed: 10421227]
25. O'Gorman DJ, Karlsson HK, McQuaid S, et al. Exercise training increases insulin-stimulated glucose disposal and GLUT4 (SLC2A4) protein content in patients with type 2 diabetes. *Diabetologia.* 2006; 49(12):2983–92. [PubMed: 17019595]
26. Peluso G, Petillo O, Margarucci S, et al. Decreased mitochondrial carnitine translocase in skeletal muscles impairs utilization of fatty acids in insulin-resistant patients. *Front Biosci.* 2002; 7:a109–16. [PubMed: 11991847]
27. Petrella JK, Kim JS, Cross JM, Kosek DJ, Bamman MM. Efficacy of myonuclear addition may explain differential myofiber growth among resistance-trained young and older men and women. *Am J Physiol Endocrinol Metab.* 2006; 291(5):E937–46.
28. Praet SF, van Loon LJ. Exercise: the brittle cornerstone of type 2 diabetes treatment. *Diabetologia.* 2008; 51(3):398–401. [PubMed: 18183362]
29. Prudente S, Morini E, Trischitta V. Insulin signaling regulating genes: effect on T2DM and cardiovascular risk. *Nat Rev Endocrinol.* 2009; 5(12):682–93. [PubMed: 19924153]
30. Ribel-Madsen R, Fraga MF, Jacobsen S, et al. Genome-wide analysis of DNA methylation differences in muscle and fat from monozygotic twins discordant for type 2 diabetes. *PloS One.* 2012; 7(12):e51302. [PubMed: 23251491]
31. Ruchat SM, Weisnagel JS, Rankinen T, Bouchard C, Vohl MC, Perusse L. Interaction between HNF4A polymorphisms and physical activity in relation to type 2 diabetes-related traits: results from the Quebec Family Study. *Diabetes Res Clin Pract.* 2009; 84(3):211–8. [PubMed: 19406499]
32. Schulz RA, Yutzey KE. Calcineurin signaling and NFAT activation in cardiovascular and skeletal muscle development. *Dev Bio.* 2004; 266(1):1–16. [PubMed: 14729474]
33. Sriwijitkamol A, Christ-Roberts C, Berria R, et al. Reduced skeletal muscle inhibitor of kappaB beta content is associated with insulin resistance in subjects with type 2 diabetes: reversal by exercise training. *Diabetes.* 2006; 55(3):760–7. [PubMed: 16505240]
34. Stoffel M, Duncan SA. The maturity-onset diabetes of the young (MODY1) transcription factor HNF4alpha regulates expression of genes required for glucose transport and metabolism. *Proc Natl Acad Sci.* 1997; 94(24):13209–14. [PubMed: 9371825]

35. Timmons JA, Knudsen S, Rankinen T, et al. Using molecular classification to predict gains in maximal aerobic capacity following endurance exercise training in humans. *J Apply Physiol*. 2010; 108(6):1487–96.
36. Wang L, Hauser ER, Shah SH, et al. Peakwide mapping on chromosome 3q13 identifies the kalirin gene as a novel candidate gene for coronary artery disease. *Am J Hum Genet*. 2007; 80(4):650–63. [PubMed: 17357071]
37. Weissglas-Volkov D, Huertas-Vazquez A, Suviolahti E, et al. Common hepatic nuclear factor-4alpha variants are associated with high serum lipid levels and the metabolic syndrome. *Diabetes*. 2006; 55(7):1970–7. [PubMed: 16804065]
38. Wisloff U, Najjar SM, Ellingsen O, et al. Cardiovascular risk factors emerge after artificial selection for low aerobic capacity. *Science*. 2005; 307(5708):418–20. [PubMed: 15662013]
39. Zabaneh D, Chambers JC, Elliott P, Scott J, Balding DJ, Kooner JS. Heritability and genetic correlations of insulin resistance and component phenotypes in Asian Indian families using a multivariate analysis. *Diabetologia*. 2009; 52(12):2585–9. [PubMed: 19763535]
40. Zhi G, Ryder JW, Huang J, et al. Myosin light chain kinase and myosin phosphorylation effect frequency-dependent potentiation of skeletal muscle contraction. *Proc Natl Acad Sci*. 2005; 102(48):17519–24. [PubMed: 16299103]

Figure 1.

Analysis work flow. Participants completing 9-months of aerobic exercise training were considered for analysis. Two representative cohorts were selected for skeletal muscle global gene expression analysis for Discovery (n=29; Affymetrix HG U133 Plus 2.0 Array) and Confirmation (n=42; Illumina HumanHT-12 v4 Expression) prior to and following exercise training. An independent cohort (n=14) from the Confirmation cohort not represented in Discovery was analyzed separately. Skeletal muscle insulin sensitivity (S_i) was determined prior to and following exercise via an intravenous glucose tolerance test. Pearson product-moment correlation analysis was performed for basal gene expression and exercise induced change in gene expression (post gene expression – pre gene expression) relationship to % S_i (Post-exercise S_i – Pre-exercise S_i) / Pre-exercise $S_i \times 100$). Genes found to have a significant relationship to % S_i in both the Discovery and Confirmation cohorts were assessed for biological function using Ingenuity Pathway Analysis software.

Figure 2.

Baseline (A) and Change in Expression (B and C) of pyruvate dehydrogenase phosphatase regulatory subunit (PDPR). A: Baseline probe(s) from Discovery (224902_at; raw intensities) and Confirmation (110364 and 194016; intensities log transformed) cohorts in relationship to $\% \Delta S_i$. B and C: Change in Expression of probes following 8 month of aerobic exercise from Discovery (224902_at and 220236_at; raw intensities) and Confirmation (110364 and 1940167; intensities log transformed) in relationship to $\% \Delta S_i$.

Figure 3. Ingenuity Pathway Analysis modified Network 1 (Score 45; Clustered 21) of baseline gene expression positively (dark gray) and negatively (light gray shaded) correlated to % S_i. Gene names are presented in Table 2 and Supplemental Table 3.

Table 1

Effects Of Exercise On Insulin Action, Body Mass, And Cardiorespiratory Fitness.

	STRIDE I & II		Discovery (Affymetrix)		Confirmation (Illumina)		Independent (Illumina)	
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
N (m, f)	319 (171, 148)		39 (21, 18)		42 (19, 23)		14 (7, 7)	
Age, yr	50 ± 9		52 ± 8		51 ± 10		46 ± 12	
Body Mass								
Height, cm	171 ± 9		169 ± 10		168 ± 9.8		170 ± 7.5	
Mass, kg	88.0 ± 13.2	87.0 ± 13.3	86.6 ± 13.4	85.0 ± 13.4	84.2 ± 10.9	83.8 ± 11.1	85.6 ± 7.2	85.2 ± 9.0
BMI, kg/m ²	30.0 ± 3.0	29.5 ± 3.5	30.4 ± 2.8	29.8 ± 2.7	29.7 ± 2.8	29.5 ± 2.8	29.3 ± 2.7	29.2 ± 3.3
Insulin Action								
Change in S _i , %	46 ± 84		44 ± 62		61 ± 88		89 ± 115	
S _i , mU·l ⁻¹ ·min ⁻¹	3.9 ± 3.4	4.8 ± 4.1 *	3.8 ± 2.6	4.9 ± 3.2 *	4.5 ± 3.0	6.7 ± 5.7 *	5.6 ± 3.4	9.6 ± 7.9 *
AIRG, mU·mM ⁻¹	483 ± 384	433 ± 360	465 ± 466	423 ± 431	456 ± 391	389 ± 338	533 ± 318	416 ± 338
S _G , ·min ⁻¹	1623 ± 1566	1848 ± 1823	1413 ± 1256	1702 ± 1505	1923 ± 2003	2338 ± 2288	2927 ± 2683	3524 ± 2895
Insulin, mU·ml ⁻¹	9.3 ± 5.8	8.2 ± 4.8 *	8.8 ± 6.1	7.5 ± 4.2	8.6 ± 6.7	7.5 ± 4.2	6.1 ± 6.7	5.5 ± 1.7
Glucose, mg·dl ⁻¹	94.3 ± 10.3	94.25 ± 9.5	95.0 ± 10.4	95.8 ± 9.7	92.6 ± 11.6	94.2 ± 9.3	86.2 ± 13.0	91.3 ± 8.7
Cardiorespiratory Fitness								
VO ₂ , ml·kg ⁻¹ ·min ⁻¹	28.0 ± 5.8	30.8 ± 6.9 *	27.9 ± 6.1	30.9 ± 6.9 *	27.5 ± 6.3	31.0 ± 7.4 *	29.4 ± 7.2	33.5 ± 8.4 *
Exercise Adherence, %	89 ± 14		88 ± 12		85 ± 15		80 ± 8.4	

Data are means ± SD.

* p < 0.05 as compared to Pre

Table 2
Genes At Baseline Negatively And Positively Correlated To Percent Change In Insulin Sensitivity.

Gene Symbol	Gene Title	Discovery Cohort (n = 39)				Confirmation Cohort (n = 42)				Independent Cohort* (n = 14)			
		Probeset ID	r	r ²	p-value	Probeset ID	r	r ²	p-value	r	r ²	p-value	(correlation)
MAP4K1	mitogen-activated protein kinase kinase kinase 1	214339_s_at	-0.43	0.18	0.01	6660368	-0.31	0.10	0.04	-0.35	0.12	0.22	
SAP30	Sin3A-associated protein, 30kDa	213963_s_at	-0.42	0.18	0.01	1780673	-0.38	0.14	0.01	-0.49	0.24	0.07	
GATA4	GATA binding protein 4	553131_a_at	-0.41	0.17	0.01	3440239	-0.34	0.12	0.03	-0.29	0.08	0.32	
KLHDC7B	kelch domain containing 7B	1552639_at	-0.41	0.17	0.01	2350452	-0.35	0.12	0.02	-0.51	0.26	0.06	
STAT8	stathmin	206835_at	-0.40	0.16	0.01	20113	0.36	0.13	0.02	0.57	0.33	0.03	
NLRP8	NLR family, pyrin domain containing 8	1553526_at	-0.39	0.15	0.01	6770326	-0.31	0.09	0.05	-0.62	0.39	0.02	
GDPD1	glycerophosphodiester phosphodiesterase domain containing 1	555606_a_at	-0.39	0.15	0.02	4780403	-0.33	0.11	0.03	-0.62	0.38	0.02	
ERC1	ELKS/RAB6-interacting/CAST family member 1	563947_a_at	-0.39	0.15	0.02	7040193	-0.43	0.19	<0.01	-0.63	0.40	0.02	
SEZ6L2	seizure related 6 homolog (mouse)-like 2	238406_x_at	-0.38	0.14	0.02	5490692	-0.31	0.10	0.04	-0.47	0.22	0.09	
XIST	X (inactive)-specific transcript (non-protein coding)	243712_at	-0.37	0.14	0.02	1690440	-0.34	0.11	0.03	-0.59	0.35	0.03	
PDP1	pyruvate dehydrogenase phosphatase regulatory subunit	224902_at	0.41	0.17	0.01	110364	0.50	0.25	<0.01	0.65	0.43	0.01	
DENND5A	DENN/MADD domain containing 5A	212561_at	0.43	0.18	0.01	1940167	0.46	0.21	<0.01	0.56	0.31	0.04	
C17orf90	chromosome 17 open reading frame 90	50374_at	0.44	0.19	0.01	2640408	0.40	0.16	0.01	0.57	0.32	0.03	
SEC13	SEC13 homolog (S. cerevisiae)	221797_at	0.38	0.14	0.02	2070373	0.31	0.09	0.05	0.35	0.13	0.21	
LMNA	lamin A/C	207707_s_at	0.45	0.20	<0.01	940300	0.34	0.12	0.03	0.46	0.21	0.10	
		203411_s_at	0.46	0.21	<0.01	1580017	0.31	0.10	0.04	0.49	0.24	0.07	
		212086_x_at	0.36	0.13	0.02	2630768	0.31	0.10	0.05	0.37	0.14	0.19	
		554600_s_at	0.37	0.14	0.02								
MYLK2	myosin light chain kinase 2	231792_at	0.47	0.22	<0.01	4290168	0.34	0.12	0.03	0.50	0.25	0.07	
MUL1	mitochondrial E3 ubiquitin protein ligase 1	218246_at	0.47	0.22	<0.01	4120487	0.31	0.10	0.04	0.65	0.42	0.01	
SERINC3	serine incorporator 3	221471_at	0.48	0.23	<0.01	840259	0.35	0.12	0.02	0.41	0.17	0.15	

Gene Symbol	Gene Title	Discovery Cohort (n = 39)				Confirmation Cohort (n = 42)				Independent Cohort* (n = 14)			
		Probeset ID	r	r ²	p-value (correlation)	Probeset ID	r	r ²	p-value (correlation)	r	r ²	p-value (correlation)	
SLC25A20	solute carrier family 25 (carnitine/acylcarnitine translocase), member 20	221473_x_at	0.33	0.11	0.04								
		203658_at	0.49	0.24	< 0.01	6220161	0.33	0.11	0.03	0.24	0.06		0.40
TIMM22	translocase of inner mitochondrial membrane 22 homolog (yeast)	229100_s_at	0.50	0.25	< 0.01	3390373	0.42	0.18	0.01	0.38	0.14		0.19
		219184_x_at	0.42	0.18	0.01								
MLXIP	MLX interacting protein	225157_at	0.50	0.25	< 0.01	4280458	0.36	0.13	0.02	0.57	0.32		0.04

Table 3

Top Changes In Genes Following Exercise Intervention Negatively And Positively Correlated To Percent Change In Insulin Sensitivity.

Gene Symbol	Gene Title	Discovery Cohort (n=39)				Confirmation Cohort (n=42)				Independent Cohort* (n = 14)			
		Probeset ID	r	r ²	p-value (correlation)	Probeset ID	r	r ²	p-value (correlation)	r	r ²	p-value (correlation)	
KALRN	kalirin, RhoGEF kinase	236651_at	-0.40	0.16	0.01	6660403	-0.33	0.11	0.03	-0.30	0.09	0.30	
EFCAB6	EF-hand calcium binding domain 6	233345_at	-0.39	0.15	0.01	3840639	-0.32	0.10	0.04	-0.42	0.17	0.14	
GPR125	G protein-coupled receptor 125	1555122_at	-0.35	0.12	0.03	2060707	-0.32	0.10	0.04	-0.31	0.10	0.27	
BLVRA	biliverdin reductase A	203773_x_at	-0.35	0.12	0.03	2940022	-0.35	0.12	0.02	-0.30	0.09	0.29	
FNDC3B	fibronectin type III domain containing 3B	1569490_at	-0.35	0.12	0.03	4250736	-0.35	0.12	0.02	-0.42	0.17	0.14	
ERC1	ELKS/RAB6-interacting/CAST family member 1	1563947_a_at	-0.33	0.11	0.04	3460343	-0.33	0.11	0.03	-0.32	0.10	0.26	
TTC39C	tetratricopeptide repeat domain 39C	1552663_a_at	-0.32	0.10	0.05	3310142	-0.33	0.11	0.04	-0.29	0.08	0.32	
SYT7	synaptotagmin VII	226167_at	-0.32	0.10	0.05	2640575	-0.46	0.22	<0.01	-0.53	0.29	0.05	
SCARA3	scavenger receptor class A, member 3	219416_at	0.34	0.11	0.04	2190278	0.31	0.09	0.05	0.78	0.60	<0.01	
IFT3	interferon-induced protein with tetratricopeptide repeats 3	229450_at	0.34	0.11	0.03	1500280	0.32	0.10	0.04	0.18	0.03	0.53	
DIRAS2	DIRAS family, GTP-binding RAS-like 2	240122_at	0.34	0.12	0.03	5340605	0.32	0.10	0.04	0.47	0.22	0.09	
MYLK	myosin light chain kinase	224823_at	0.34	0.12	0.03	7160343	0.31	0.10	0.05	-0.03	0.00	0.91	
MYH11	myosin, heavy chain 11, smooth muscle	201497_x_at	0.35	0.12	0.03	6280133	0.36	0.13	0.02	0.26	0.07	0.37	
RNF39	ring finger protein 39	207961_x_at	0.33	0.11	0.04	3190156	0.35	0.12	0.02	0.33	0.11	0.25	
TIMM22	translocase of inner mitochondrial membrane 22 homolog (yeast)	219916_s_at	0.35	0.12	0.03	1980255	0.39	0.15	0.01	0.66	0.43	0.01	
		219184_x_at	0.35	0.12	0.03	3390373	0.35	0.12	0.02	0.34	0.12	0.23	
TOMM34	translocase of outer mitochondrial membrane 34	201870_at	0.35	0.12	0.03	1190445	0.37	0.13	0.02	0.59	0.35	0.03	
OPA3	optic atrophy 3 (autosomal recessive, with chorea and spastic paraplegia)	229624_at	0.35	0.12	0.03	1450615	0.38	0.14	0.01	0.58	0.34	0.03	

Gene Symbol	Gene Title	Discovery Cohort (n=39)			Confirmation Cohort (n=42)			Independent Cohort* (n = 14)		
		Probeset ID	r	r ²	p-value (correlation)	Probeset ID	r	r ²	p-value (correlation)	p-value (correlation)
PDPR	pyruvate dehydrogenase	220236_at	0.38	0.14	0.02	1940167	0.42	0.18	0.01	0.06
	phosphatase regulatory subunit	224902 at	0.43	0.18	0.01	110364	0.44	0.19	<0.01	0.08