PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes

Vamsi K Mootha[1,2,3,10], Cecilia M Lindgren[1,4,10], Karl-Fredrik Eriksson[4], Aravind Subramanian[1], Smita Sihag[1], Joseph Lehar[1], Pere Puigserver[5], Emma Carlsson[4], Martin Ridderstråle[4], Esa Laurila[4], Nicholas Houstis[1], Mark J Daly[1], Nick Patterson[1], Jill P Mesirov[1], Todd R Golub[1,5], Pablo Tamayo[1], Bruce Spiegelman[5], Eric S Lander[1,6], Joel N Hirschhorn[1,7,8], David Altshuler[1,2,7,9,11] & Leif C Groop[4,11]

DNA microarrays can be used to identify gene expression changes characteristic of human disease. This is challenging, however, when relevant differences are subtle at the level of individual genes. We introduce an analytical strategy, Gene Set Enrichment Analysis, designed to detect modest but coordinate changes in the expression of groups of functionally related genes. Using this approach, we identify a set of genes involved in oxidative phosphorylation whose expression is coordinately decreased in human diabetic muscle. Expression of these genes is high at sites of insulin-mediated glucose disposal, activated by PGC-1α and correlated with total-body aerobic capacity. Our results associate this gene set with clinically important variation in human metabolism and illustrate the value of pathway relationships in the analysis of genomic profiling experiments.

Type 2 diabetes mellitus (DM2) affects over 110 million people worldwide and is a principal contributor to atherosclerotic vascular disease, blindness, amputation and kidney failure[1]. Defects in insulin secretion are observed early in individuals with maturity-onset diabetes of the young, a monogenic form of type 2 diabetes[2]; insulin resistance at tissues including skeletal muscle is a cardinal feature of individuals with fully developed DM2. Many molecular pathways have been implicated in the disease process: β-cell development, insulin receptor signaling, carbohydrate production and utilization, mitochondrial metabolism, fatty acid oxidation, cytokine signaling, adipogenesis, adrenergic signaling and others. It is unclear, however, which of these or other pathways are disturbed in, and might be responsible for, DM2 in its common form.

Expression profiling using DNA microarrays enables researchers to survey the genome for transcripts whose levels are altered in tissue from individuals with disease. Microarray data can be used to classify individuals according to molecular characteristics and to generate hypotheses about disease mechanisms. This approach has been successful in the study of cancer[3], where large changes in the expression of individual genes have often been observed. When alterations in gene expression are more modest, however, the large number of genes tested, high variability between individuals and limited sample sizes typical of human studies make it difficult to distinguish true differences from noise.

One promising approach to increase power exploits the idea that alterations in gene expression might manifest at the level of biological pathways or coregulated gene sets, rather than individual genes. Subtle but coordinated changes in expression might be detected more readily by combining measurements across multiple members of each gene set. A straightforward strategy for identifying such differences is to examine top-ranking genes in a microarray experiment and then to create hypotheses about pathway membership. This is both subjective and *post hoc*, however, and thus prone to bias. A more objective set of approaches[4,5] tests for enrichment of pathway members among the top-ranking genes in a microarray study, comparing them to a null distribution in which genes are randomly distributed. Because functionally related genes are often coregulated, however, a positive result in such a test can be due solely to intrinsic correlation in gene expression rather than any relationship between expression of pathway members and the phenotype of interest.

We present an analytical technique designed to test *a priori* defined gene sets (for example, pathways) for association with disease phenotypes. We apply this method to gene expression profiles of human diabetic muscle, identifying a set of genes whose expression is correlated with insulin resistance and aerobic capacity. These results suggest hypotheses about pathways contributing to human metabolic disease and, more generally, show the value of incorporating information about functional relationships among genes in the analysis of microarray data.

[1]Whitehead Institute/MIT Center for Genome Research, Cambridge, Massachusetts, USA. [2]Department of Medicine, Harvard Medical School, Boston, Massachusetts, USA. [3]Department of Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA. [4]Department of Endocrinology, Wallenberg Laboratory, University Hospital MAS, Lund University, S-205 02 Malmo, Sweden. [5]Dana Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts, USA. [6]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. [7]Department of Genetics, Harvard Medical School, Boston, Massachusetts, USA. [8]Divisions of Pediatrics and Endocrinology, Children's Hospital, Boston, Massachusetts, USA. [9]Department of Molecular Biology and Diabetes Unit, Massachusetts General Hospital, Boston, Massachusetts 02114, USA. [10]These authors contributed equally to this work. [11]These two authors contributed equally to this work. Correspondence should be addressed to D.A. (altshuler@molbio.mgh.harvard.edu) or L.C.G. (leif.groop@endo.mas.lu.se).

RESULTS

We used DNA microarrays to profile expression of over 22,000 genes in skeletal muscle biopsy samples from 43 age-matched males (**Table 1**), 17 with normal glucose tolerance (NGT), 8 with impaired glucose tolerance (IGT) and 18 with DM2. We obtained samples at the time of diagnosis (before treatment with hypoglycemic medication) and under the controlled conditions of a hyperinsulinemic euglycemic clamp. When assessed with either of two different analytical techniques[3,6] that take into account the multiple comparisons implicit in microarray analysis, no single gene had a significant difference in expression between the diagnostic categories (data not shown). This result is consistent with smaller studies[7,8] that did not identify any individual gene whose expression difference was significant when corrected for the large number of hypotheses tested[9,10].

Gene Set Enrichment Analysis

To test for sets of related genes that might be systematically altered in diabetic muscle, we devised a simple approach called Gene Set Enrichment Analysis (GSEA), which we introduce here (**Fig. 1**) and describe in more detail elsewhere (A.S. *et al.*, manuscript in preparation). The method combines information from the members of previously defined sets of genes (for example, biological pathways) to increase signal relative to noise and improve statistical power.

For a given pairwise comparison (for example, highly expressed in individuals with NGT versus those with DM2), we rank all genes according to the difference in expression (using an appropriate metric, such as signal-to-noise ratio, SNR). The null hypothesis of GSEA is that the rank ordering of the genes in a given comparison is random with regard to the diagnostic categorization of the samples. The alternative hypothesis is that the rank ordering of the pathway members is associated with the specific diagnostic criteria used to categorize the groups of affected individuals.

We then measure the extent of association by a non-parametric, running sum statistic termed the enrichment score (ES) and record the maximum ES (MES) over all gene sets in the actual data from affected individuals (**Fig. 1**). To assess the statistical significance of the MES, we use permutation testing of the diagnostic labels of the individuals (for example, whether an individual is affected with NGT or DM2; **Fig. 1**). Specifically, we compare the MES achieved in the actual data to that seen in each of 1,000 permutations that shuffled the diagnostic labels among the samples. The significance of the MES score is calculated as the fraction of the 1,000 random permutations in which the top pathway gave a stronger result than that observed in the actual data. Because the permutation test involves randomization of the diagnostic labels, it is a test for the dependence on the actual diagnostic status of the affected individuals. Moreover, because the actual MES is compared to the distribution of maximal ES values over all pathways examined in each of the randomized data sets, it accounts for multiple pathways tested, and no further correction is required[9,10].

### Decreased expression of genes involved in oxidative phosphorylation

We applied GSEA to the microarray data described above, using 149 gene sets that we compiled (**Supplementary Table 1** online). Of these gene sets, 113 are grouped according to involvement in metabolic pathways (derived from public or local curation[11]) and 36 consist of gene clusters that are coregulated in a mouse expression atlas of 46 tissues[12]. The gene sets were selected without regard to the results of the microarray data from the affected individuals. The top gene set in

Table 1 Clinical and biochemical characteristics of male subjects with NGT, IGT and DM2

| | Class | | | | P value | |
	NGT	IGT	DM2	NGT versus IGT	IGT versus DM2	NGT versus DM2
n	17	8	18			
Age, y	66.1 (1.0)	66.4 (1.6)	65.5 (1.8)			
BMI, kg/m^2	23.6 (3.4)	27.1 (4.8)	27.3 (4.0)			5.70 10^{-3}
WHR	0.91 (0.09)	0.97 (0.04)	0.99 (0.03)	3.00 10^{-2}		3.83 10^{-3}
Triglycerides, mmol/l	1.03 (0.40)	1.83 (1.60)	2.04 (1.13)			2.63 10^{-3}
Cholesterol, mmol/l	5.39 (0.09)	4.60 (1.48)	5.77 (0.97)			
OGTT						
Glucose 0 min, mmol/l	4.67 (0.50)	5.05 (0.46)	7.83 (2.3)		9.22 10^{-5}	2.01 10^{-5}
Insulin 0 min, U/ml	5.41 (3.3)	13.38 (8.9)	12.0 (6.0)	4.05 10^{-2}		4.10 10^{-4}
Glucose 120 min, mmol/l	6.58 (0.94)	9.15 (0.8)	14.9 (4.0)	2.51 10^{-6}	8.91 10^{-6}	4.90 10^{-8}
Insulin 120 min, U/ml	33.5 (19.3)	125.1 (66.1)	43.5 (25.6)	5.47 10^{-3}	9.73 10^{-3}	
M value, mg/kg/min	8.74 (3.15)	6.32 (3.08)	4.22 (1.72)			2.30 10^{-5}
VO2max, ml O$_2$/kg/min	32.1 (5.46)	26.5 (4.6)	24.3 (5.6)	1.72 10^{-2}		3.09 10^{-4}
Glycogen, mmol/kg	371.1 (77.0)	326.5 (88.0)	350.6 (97.8)			
Type I fibers						
Number, %	37.2 (13.5)	33.5 (3.6)	36.4 (9.3)			
Area, %	39.1 (14.4)	32.7 (0.91)	40.1 (10.7)		2.35 10^{-2}	
Capillaries/fiber	3.91 (0.72)	4.05 (1.04)	4.14 (0.75)			
Type IIb fibers						
Number, %	73.8 (42.1)	60.2 (51.4)	72.2 (36.7)			
Area, %	31.3 (18.0)	24.7 (18.3)	36.2 (15.4)			
Capillaries/fiber	2.97 (0.71)	3.05 (0.87)	3.02 (0.65)			

Values are mean (s.d.). OGTT, oral glucose tolerance test. M value is the total body glucose uptake. VO2max is the total body aerobic capacity. Only values of $P < 0.05$ are shown for pairwise comparisons, using a two-sided t-test.

Figure 1 Schematic overview of GSEA. The goal of GSEA is to determine whether any *a priori* defined gene sets (step 1) are enriched at the top of a list of genes ordered on the basis of expression difference between two classes (for example, highly expressed in individuals with NGT versus those with DM2). Genes $R_1,…R_N$ are ordered on the basis of expression difference (step 2) using an appropriate difference measure (for example, SNR). To determine whether the members of a gene set *S* are enriched at the top of this list (step 3), a Kolmogorov-Smirnov (K-S) running sum statistic is computed: beginning with the top-ranking gene, the running sum increases when a gene annotated to be a member of gene set *S* is encountered and decreases otherwise. The ES for a single gene set is defined as the greatest positive deviation of the running sum across all *N* genes. When many members of *S* appear at the top of the list, ES is high. The ES is computed for every gene set using actual data, and the MES achieved is recorded (step 4). To determine whether one or more of the gene sets are enriched in one diagnostic class relative to the other (step 5), the entire procedure (steps 2–4) is repeated 1,000 times, using permuted diagnostic assignments and building a histogram of the maximum ES achieved by any pathway in a given permutation. The MES achieved using the actual data is then compared to this histogram (step 6, red arrow), providing us with a global *P* value for assessing whether any gene set is associated with the diagnostic categorization.
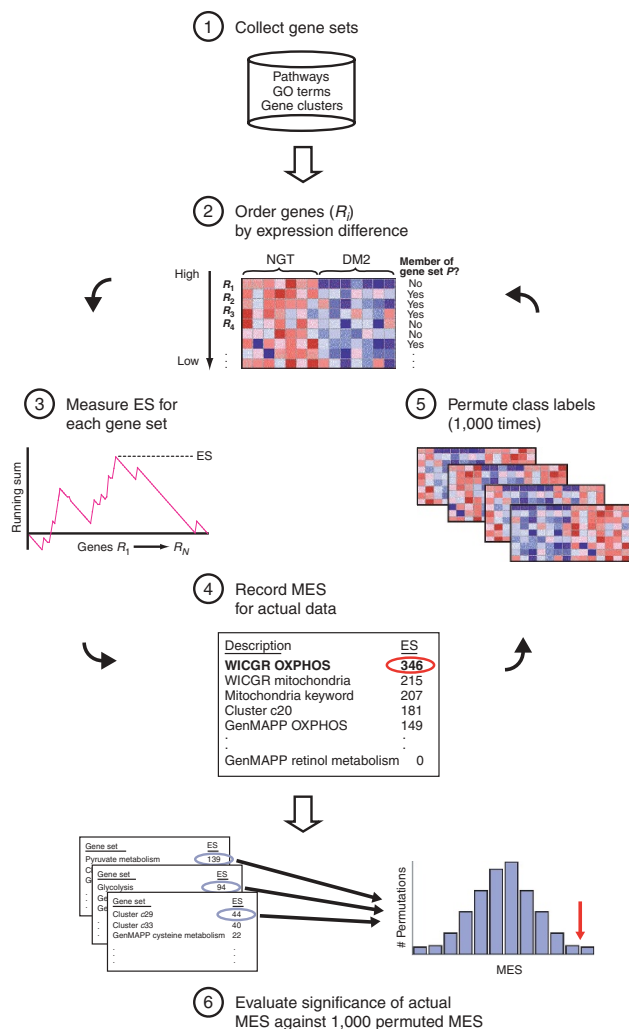
GSEA analysis yielded an MES score (MES = 346) that was significant at $P = 0.029$ over the 1,000 permutations of the 149 pathways. That is, in only 29 of 1,000 permutations did the top pathway (of the 149) exceed the score achieved by the top pathway achieved using the actual diagnostic labels.

The maximal ES score was obtained for an internally curated set consisting of genes involved in oxidative phosphorylation (we refer to this gene set as OXPHOS). Notably, the four gene sets with the next highest ES scores overlap with this OXPHOS gene set, and their enrichment is almost entirely explained by the overlap: a locally curated set of genes involved in mitochondrial function, a set of genes identified with the keyword 'mitochondria,' a cluster (referred to here as c20) of coregulated genes derived from the comparison of publicly available mouse data and a set of genes related to oxidative phosphorylation defined at the Affymetrix website[11].

Examination of the individual expression values for the 106 OXPHOS genes identifies the source of this signal (**Fig. 2**). Although the typical decrease in expression for individual OXPHOS genes is very modest (~20%), the decrease is consistent across the set: 89% (94 of 106) of the genes showing lower expression in individuals with DM2 relative to those with NGT (**Fig. 2**). As controls, we confirmed that the result is independent of specific aspects of data processing (such as scaling, thresholding, filtering) or of selection of difference metrics (data not shown). Moreover, the result identified by GSEA is supported by previous observations: others have suggested that oxidative capacities are altered in insulin resistant muscle[13,14], and recent microarray analyses of human diabetic muscle have identified genes in oxidative phosphorylation among their top-ranked genes (ref. 7 and M.E. Patti *et al.*, manuscript submitted).

OXPHOS-CR: a coregulated subset of OXPHOS genes
One of the overlapping gene sets identified by GSEA is cluster c20, defined as a set of genes that are tightly coregulated across many tissues. The partial overlap of OXPHOS with the coregulated cluster led us to ask whether all OXPHOS genes are coordinately regulated. We examined transcriptional coregulation of mouse homologs of OXPHOS genes across a mouse tissue expression atlas[12]. This identified a previously unrecognized subset of the OXPHOS biochemical pathway, corresponding to about two-thirds of the OXPHOS genes, that are strongly correlated across mouse tissues ($r = 0.61$; **Fig. 3a**).

We term this subset OXPHOS-CR (oxidative phosphorylation co-regulated). The remaining OXPHOS genes show little co-regulation with OXHPOS-CR genes or with each other (**Fig. 3a**). The OXPHOS-CR subset was strongly expressed in 3 of 46 tissues: skeletal muscle, heart and brown fat. We note that these are the principal sites of insulin-mediated glucose disposal in mice.

We next asked whether the downregulation of OXPHOS observed in DM2 was a general property of all OXPHOS genes or was specific to OXPHOS-CR genes. Notably, the bulk of the statistical signal we observe in GSEA is accounted for by OXPHOS-CR (**Supplementary Fig. 1** online). Namely, the OXPHOS-CR subset showed a stronger mean deviation than the remainder of the OXPHOS gene set (mean SNR of 0.235 versus 0.128; $P = 0.04$) and was itself significant in the GSEA analysis (nominal P value = 0.001, as compared with nominal $P = 0.226$ for the remainder of the OXPHOS set). To see if these changes were secondary to hyperglycemia *per se* or preceded the onset of frank diabetes, we compared expression of OXPHOS-CR in individuals with NGT to that in individuals with IGT, the pre-diabetic state. We found that expression of OXPHOS-CR was also downregulated in individuals with IGT (nominal $P < 10^{-4}$). This suggests that downregulation of OXPHOS-CR precedes onset of hyperglycemia. Thus, GSEA allowed us to detect a subset of OXPHOS genes, called OXPHOS-CR, with three key properties: (i) they are members of the oxidative phosphorylation pathway (ii) they are tightly coregulated
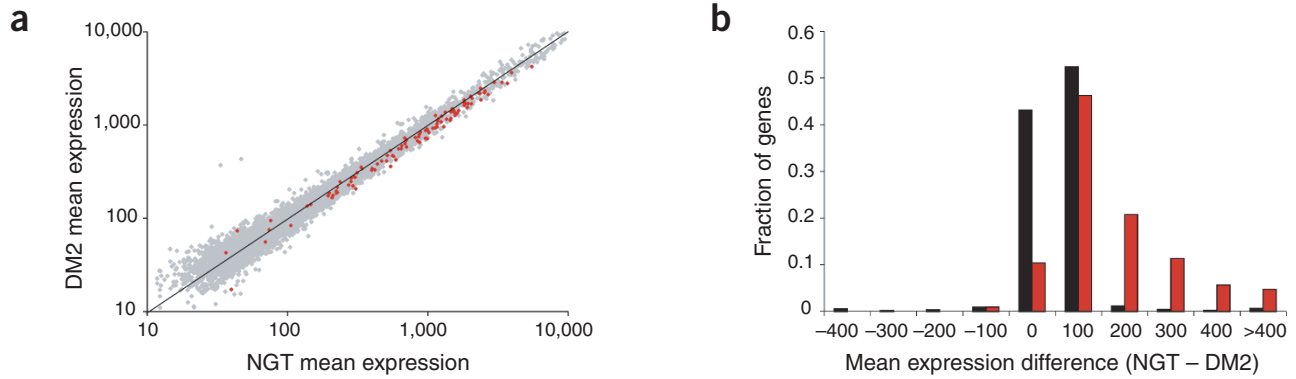
Figure 2 OXPHOS gene expression is reduced in diabetic muscle. (**a**) The mean expression of all genes (gray) and of OXPHOS genes (red) is plotted for individuals with DM2 versus those with NGT. (**b**) Histogram of mean gene expression level differences between individuals with NGT and DM2, using the data from **a**, for all genes (black) and for OXPHOS genes (red).

across many tissues and are highly expressed in the principal sites of insulin-mediated glucose disposal and (iii) their expression is subtly but consistently lower in muscle from individuals with both the pre-diabetic state IGT and DM2.

PGC-1α induces expression of OXPHOS-CR

The strong correlation in expression of the OXPHOS-CR genes and their coordinated downregulation in diabetic muscle led us to explore mechanisms that might mediate this tight control. We reasoned that peroxisome proliferator-activated receptor γ coactivator 1α (PGC-1α, encoded by *PPARGC1*), a cold-inducible regulator of mitochondrial biogenesis, thermogenesis and skeletal muscle fiber–type switching[15–17], was a prime candidate for mediating these effects. Consistent with this hypothesis, we observed that mean levels of *PPARGC1* transcript were similarly lower (by ~20%) in the diabetic muscle and noted that the promoters of several of the OXPHOS-CR genes have been reported to contain binding sites for nuclear respiratory factor 1, a transcription factor coactivated by PGC-1α[18].

To test directly whether OXPHOS-CR genes might be transcriptional targets of PGC-1α, we expressed PGC-1α in a mouse skeletal

muscle cell line using an adenoviral expression vector[17] and used DNA microarrays to profile expression of the OXPHOS genes over a 3-d period. We found that a subset of OXPHOS genes was strongly upregulated in a time-dependent manner in response to PGC-1α and that this subset corresponded almost precisely to OXPHOS-CR (**Fig. 3b**). These *in vitro* results support the hypothesis that PGC-1α has a role in the regulation of OXPHOS-CR, both across the mouse tissue compendium and in the observed downregulation in diabetes.

Expression of OXPHOS-CR and measures of whole-body physiology

Metabolic control theory suggests that small adjustments in many sequential steps of a metabolic pathway can lead to a substantial change in the total flux through the pathway, whereas large changes in a single enzyme might have no measurable effects[19]. To test the hypothesis that differences in OXPHOS-CR gene expression in diabetic individuals might be related to changes in total body metabolism, we examined the relationships between diabetes status, expression of OXPHOS-CR genes and maximal oxygen uptake (VO2max) as measured in affected individuals (**Fig. 4**). Consistent
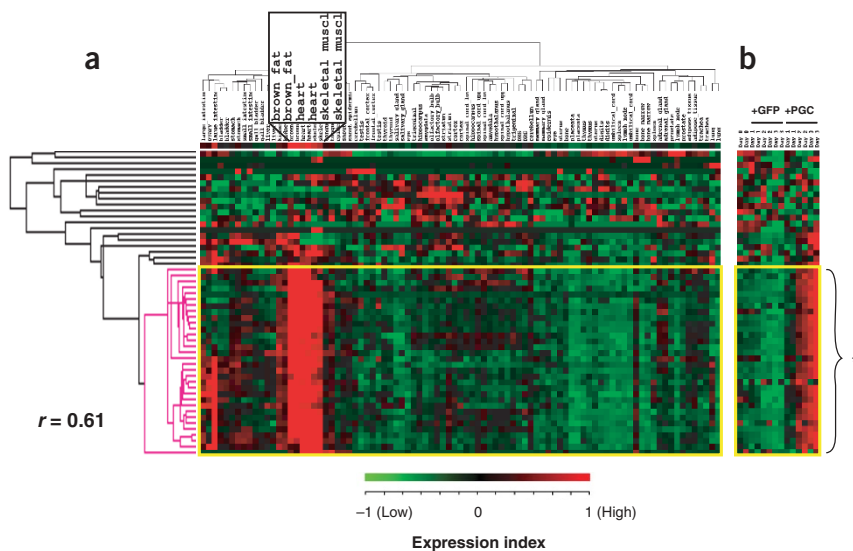

Figure 3 OXPHOS-CR represents a coregulated subset of OXPHOS genes responsive to the transcriptional coactivator PGC-1α. (**a**) Normalized expression profile of 52 mouse homologs of the human OXPHOS genes across the mouse expression atlas[12]. These 52 genes were hierarchically clustered[32]. The pink tree on the left corresponds to a subcluster with a correlation coefficient of 0.65. We call the human homologs of these mouse genes the OXPHOS-CR set. The human homologs of this tightly coregulated cluster, marked with an asterisk and delimited with a yellow box, are *ATP5J*, *ATP5L*, *ATP5O*, *COX5B*, *COX6A2*, *COX7A1*, *COX7B*, *COX7C*, *CYC1*, *CYCS*, *GRIM19*, *HSPC051*, *NDUFA2*, *NDUFA5*, *NDUFA7*, *NDUFA8*, *NDUFB3*, *NDUFB5*, *NDUFB6*, *NDUFC1*, *NDUFS2*, *NDUFS3*, *NDUFS5*, *SDHA*, *SDHB*, *UQCRB* and *UQCRC1*. (**b**) Normalized expression profile of OXPHOS mouse homologs in a mouse skeletal muscle cell line during a 3-d time course in response to PGC-1α. The expression profile includes infection with control vectors (expressing GFP) or with vectors expressing PGC-1α before infection (d 0) and 1, 2 and 3 d after adenoviral infection, all done in duplicate.

with previous reports[20], diabetes and VO2max were correlated in affected individuals ($R^2_{adj} = 0.28$, $P = 0.0005$). Notably, we found that the expression of OXPHOS-CR genes in muscle was strongly correlated with VO2max ($R^2_{adj} = 0.22$, $P = 0.0012$; **Fig. 4**), a measure of total-body physiology. The top ranking OXPHOS-CR gene, ubiquinol cytochrome c reductase binding protein (*UQCRB*), was even a stronger predictor ($R^2_{adj} = 0.31$, $P < 0.0001$). Expression of OXPHOS-CR genes is not merely a proxy for diabetes status, however, because a two-variable regression of VO2max on diabetes status and OXPHOS-CR expression level shows that both variables contribute significantly to the correlation ($P = 0.05$ for the model with both variables as compared to the model with only diabetes status).

These results do not seem to be secondary to other known predictors of oxidative capacity. We found no relationship between body mass index or waist-to-hip ratio and OXPHOS-CR gene expression ($R^2_{adj} < 0.01$ in both cases). In addition, there was no significant relationship between quantitative measures of fiber types and OXPHOS-CR expression (data not shown). Thus, decreased in expression of OXPHOS-CR genes in muscle seems to be associated with changes in total-body aerobic capacity, even beyond their correlation to diabetes status, body habitus or muscle-fiber type.

DISCUSSION

Our results indicate that decreases in expression of OXPHOS-CR genes accompany, and might possibly contribute to, DM2. The relationship between OXPHOS and DM2 is richly supported by clinical investigation, exercise physiology, pharmacology and genetics. For example, the mitochondria of diabetic individuals show ultrastructural changes as well as decreases in oxidative phosphorylation activity[13,21]. Whole-body V02max (which we have shown to be correlated with OXPHOS-CR expression) predicts future development of DM2 (ref. 20). Exercise and caffeine consumption both increase oxidative phosphorylation capacity and can delay or prevent onset of diabetes[17,20,22,23]. Inherited mutations in mitochondrial DNA, which encodes 13 subunits of the electron transport chain, and whose copy number is under the control of PGC-1α[16], cause rare, inherited forms of diabetes[24]. Missense variants in PGC1-α have been reported to be associated with DM2 (refs. 25,26), although it is not yet clear if this association is reproducible[27]. Moreover, of the handful of genes in which variants have been clearly shown to influence risk of human diabetes, two are transcriptional partners of PGC1-α: HNF4-α (mutations of which cause early-onset diabetes) and PPARG, in which the Pro12Ala polymorphism is associated with risk of DM2 (reviewed in ref. 24). Further investigation will be required to test the hypothesis that the PGC-1α-regulated, OXPHOS-CR genes might represent a common link to these varied phenomena. If this hypothesis is valid, it would suggest that modulation of OXPHOS-CR activity might represent a target for the prevention and treatment of DM2.

More generally, methods like GSEA may be valuable in efforts to relate genomic variation to disease and measures of total-body physiology. Single-gene methods are powerful only when the individual gene effect is marked and the variance is small across individuals, which may not be the case in many disease states. Methods like GSEA are complementary to single-gene approaches and provide a framework with which to examine changes operating at a higher level of biological organization. This may be needed if common, complex disorders typically result from modest variation in the expression or activity of multiple members of a pathway. As gene sets are systematically assembled using functional and genomic approaches, methods such as GSEA will be valuable in detecting coordinated variation in gene function that contributes to common human diseases.

a

	Predictor(s)	R^2_{adj}	*P* value
	Diabetes status	0.28	0.0006
1	OXPHOS-CR	0.22	0.0012
	Diabetes status, OXPHOS-CR	0.33	0.0004
2	*UQCRB*	0.31	<0.0001
	Diabetes status, *UQCRB*	0.38	0.0001

[1] Addition of OXPHOS-CR improves the model with $P = 0.05$.

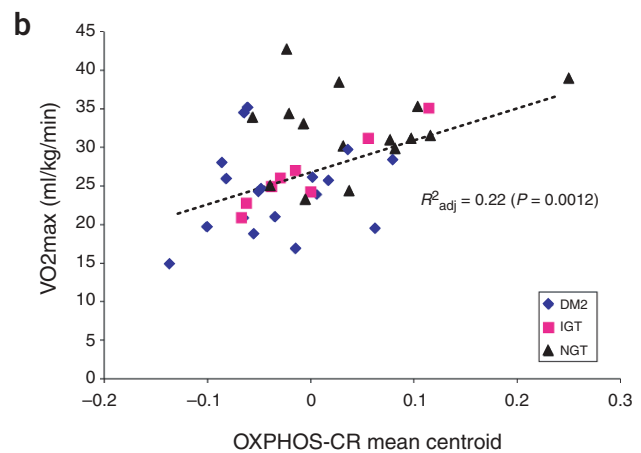[2] Addition of *UQCRB* improves the model with $P = 0.03$.

b

Figure 4 OXPHOS-CR predicts total-body aerobic capacity (VO2max). (**a**) Linear regression was used to model VO2max with diabetes status, the mean centroid of OXPHOS-CR gene expression, expression of *UQCRB* or in combination as explanatory (predictor) variables. The explanatory power and significance of the model are shown in the table. (**b**) Linear regression of VO2max against the mean centroid of OXPHOS-CR gene expression.

METHODS

Human subjects and clinical measurements. We selected 54 men of similar age but with varying degree of glucose tolerance who had been participating in The Malmö Prevention Study in southern Sweden for more than 12 years[20]. The investigation was approved by the Ethics Committee at Lund University, and informed consent was obtained from each of the volunteers. All subjects were Northern Europeans, and their glucose tolerance status was assessed using standardized 75-gram oral glucose tolerance test (OGTT) and by applying WHO85 criteria[20]. At the initial OGTT done 10 years earlier, none of the men had DM2 (ref. 20). An OGTT done at the time of the biopsy showed that 20 of the subjects had developed DM2, 8 fulfilled the criteria for IGT and 26 had NGT. As diabetes was diagnosed at the time of the repeat OGTT, none of the subjects were on medication for hyperglycemia or diabetes-related conditions.

Anthropometric and insulin sensitivity measures were done as previously described[28]. We measured height, weight, waist-to-hip ratio and fat-free mass on the day of the euglycemic clamp. We measured VO2max using an incremental work-conducted upright exercise test with a bicycle ergometer (Monark Varberg) combined with continuous analysis of expiratory gases and minute ventilation. Exercise was started at a workload varying between 30W and 100W depending on the previous history of endurance training or exercise habits and then increased by 20–50W every 3 min until a perceived exhaustion or a respiratory quotient of 1.0 was reached. Maximal aerobic capacity was defined as the VO2 during the last 30 s of exercise and is expressed per lean body mass. We determined insulin sensitivity with a standard 2-h euglycemic hyperinsulinemic clamp combined with infusion of tritiated glucose to estimate endogenous glucose production and indirect calorimetry (Deltatrac, Datex Instrumentarium) to estimate substrate oxidation[28]. We calculated the rate of glucose uptake (also referred to as the M value) from the infusion rate of glucose and the residual rate of endogenous glucose production measured by the tritiated glucose tracer during the clamp.

We took percutaneous muscle biopsy samples (20–50 mg) from the vastus lateralis muscle under local anesthesia (1% lidocaine) after the 2-h euglycemic hyperinsulinemic clamp using a Bergström needle[29]. We determined fiber-type composition and glycogen concentration as previously described[30]. We quantified and calculated the fibers using the COMFAS image analysis system (Scan Beam).

Cell culture and adenoviral infection. We cultured mouse myoblasts (C2C12 cells) and differentiated them into myotubes as previously described[16]. After 3 d of differentiation, we infected them with an adenovirus expressing either green fluorescent protein (GFP) or PGC-1α as previously described[17].

mRNA isolation, target preparation and hybridization. We prepared targets from human biopsy or mouse cell lines as previously described[3] and hybridized them to the Affymetrix HG-U133A or MG-U74Av2 chip, respectively. We selected only those scans with 10% Present calls and a *GAPD 3′/GAPD 5′* expression ratio <1.33. We obtained gene expression data for 54 human samples, but only 43 met these selection criteria; the analysis in this paper is limited to these 43 individuals.

Data scaling and filtering. We subjected human microarray data to global scaling to correct for intensity-related biases. For each scan, we binned all genes according to their expression intensity and recorded the median intensity of each to serve as a calibration curve for that scan. We then scaled the expression to the calibration curve of the scan from one individual with NGT (individual mm12), which we visually inspected and deemed high-quality, using a linear interpolation between the calibration points. We then filtered the 22,283 genes on the HG-U133A chip to eliminate genes that had extremely low expression. A previous study suggested that an Affymetrix average difference level of 100 corresponds to an extremely low level ('not expressed'; ref. 12). Therefore, we only considered genes for which there was at least a single measure (average difference) greater than 100. Of the 22,283 genes on the HG-U133A chip, 10,983 genes met this filtering criterion.

Single gene microarray analysis. We carried out microarray analysis to identify individual genes that are significantly different between diagnostic classes using two software packages. First, we carried out marker analysis as previously described using GeneCluster[3]. Significance of individual genes was tested by permutation of class labels (5,000 iterations). We used both the *t*-test and SNR difference metrics in these analyses, both yielding comparable results. Second, we used the software package SAM, using a Δ = 0.5, to search for gene expression values significantly different between classes[6].

Compilation of gene sets. We analyzed 149 gene sets consisting of manually curated pathways and clusters defined by public expression compendia (**Supplementary Table 1** online). First, we used two different sets of metabolic pathway annotations. We manually curated genes belonging to the following pathways: free fatty-acid metabolism, gluconeogenesis, glycolysis, glycogen metabolism, insulin signaling, ketogenesis, pyruvate metabolism, reactive oxygen species homeostasis, Krebs cycle, oxidative phosphorylation (OXPHOS) and mitochondria, using standard textbooks, literature reviews and LocusLink. We also downloaded NetAFFX[11] annotations (October 2002) corresponding to GenMAPP metabolic pathways. To identify sets of coregulated genes, we used self-organizing maps to group the GNF mouse expression atlas into 36 clusters[12,31]. Genes in these 36 groups were converted to Affymetrix HG-U133A probe sets using the ortholog tables available at the NetAFFX website (October 2002).

Rationale for grouped gene analysis. Consider a microarray data set with samples in two categories, *A* and *B*. For the sake of simplicity, let the size of *A* and *B* each be *n*. Consider a gene set *S* for which the expression levels differ between samples of *A* and *B*. Model the data set so that the entry D_{ij} for gene *i* and sample *j* is normally distributed with mean μ_{ij} and standard deviation σ, where

$$\mu_{ij} = \begin{cases} 0, & i \notin S \\ +\alpha, & i \in S, j \in A \\ -\alpha, & i \in S, j \in B \end{cases} .$$

Then the SNR for an individual gene in *S* is proportional to

$$\frac{\alpha\sqrt{n}}{\sigma} .$$

Suppose, on the other hand, that we know *S* and add the expression levels for all genes in *S*. Then the SNR is proportional to

$$\frac{\alpha\sqrt{nM}}{\sigma} ,$$

where *M* is the number of genes in *S*. This increases the mean of our statistic (which is standard normal for the null hypothesis of no gene set association) by a factor of $\sqrt{M}$. If the noise is in fact correlated for genes of *S*, this reduces the benefit, but we can still expect a large gain. In practice we will not be able to select a gene set containing fully concordant expression levels, but as long as an appreciable fraction of our gene set has this property, we can expect a benefit from the grouped gene approach.

Gene Set Enrichment Analysis (GSEA). GSEA determines if the members of a given gene set are enriched among the most differentially expressed genes between two classes. First, the genes are ordered on the basis of a difference metric. The results presented in the current manuscript use the SNR difference metric, which is simply the difference in means of the two classes divided by the sum of the standard deviations of the two diagnostic classes. In general, other difference metrics can also be used.

For each gene set, we then make an enrichment measure called the ES, which is a normalized Kolmogorov-Smirnov statistic. Consider the genes $R_1,.., R_N$ that are ordered on the basis of the difference metric between the two classes and a gene set *S* containing *G* members. We define

$$X_i = -\sqrt{\frac{G}{N-G}}$$

if R_i is not a member of *S*, or

$$X_i = \sqrt{\frac{N-G}{G}}$$

if R_i is a member of *S*.

We then compute a running sum across all *N* genes. The ES is defined as

$$\max_{1 \le j \le N} \sum_{i=1}^{j} X_i$$

or the maximum observed positive deviation of the running sum. ES is measured for every gene set considered. To determine whether any of the given gene sets shows association with the class phenotype distinction, we permute the class labels 1,000 times, each time recording the maximum ES over all gene sets. In this regard, we are testing a single hypothesis. The null hypothesis is that no gene set is associated with the class distinction.

In this manuscript, after identifying OXPHOS-CR as a subset of co-regulated OXPHOS genes, we tested it (a single gene set) for association with clinical status using GSEA. Because OXPHOS-CR is not independent of the OXPHOS set interrogated in the initial analysis, this cannot be viewed as an independent hypothesis. For this reason, these *P* values are explicitly marked as nominal *P* values.

GSEA has been implemented as a software tool for use with microarray data and will be presented in fuller detail, including a discussion of different varieties of multiple hypothesis testing and applications to other biomedical problems, in a companion paper (A.S. *et al.*, manuscript in preparation).

Evaluating OXPHOS coregulation in mouse expression data sets. We used the NetAFFX website to identify probe sets on the mouse expression chips corresponding to human OXPHOS probe sets. We identified a total of 114 (106 of which passed our filtering criterion) probe sets corresponding to the human genes related to oxidative phosphorylation. Using the October 2002

ortholog tables at NetAFFX, we identified 61 mouse orthologs on the Affymetrix MG-U74Av2 chip. Of these 61 probe sets, 52 were represented in the GNF mouse expression atlas[12]. These expression data were normalized to a mean of 0 and a variance of 1. Data were hierarchically clustered and visualized using the Cluster and TreeView software packages[32]. We parsed these 52 genes into 32 coregulated probe sets and 20 probe sets that are not coregulated, based on the dendrogram in **Figure 3**. Forty distinct HG-U133A probe sets mapped to the 32 coregulated mouse probe sets, and 19 distinct HG-U133A probe sets mapped to the 20 mouse probe sets that are not coregulated. Five HG-U133A probe sets are shared between these two groups, representing ambiguous cases (human probe sets that map to two mouse probe sets, one of which is coregulated and the other of which is not). We omitted these five ambiguous human probe sets from our analysis. This left a total of 35 HG-U133A probe sets, which we call OXPHOS-CR genes, and a total of 14 HG-U133A probe sets, which we call OXPHOS not CR. Thirty-four and 13 of these genes, respectively, passed our filtering criteria, and these were used in **Supplementary Figure 1** online as well as in the OXPHOS-CR analysis described in the paper.

Linear regression analysis. We generated linear regression models using SAS (SAS Institute). We used clinical variables as dependent variables and OXPHOS-CR gene expression levels or other clinical/biochemical measures as the independent (explanatory or predictor) variables. To compute the mean centroid of OXPHOS-CR, we normalized the gene expression levels of the 34 OXPHOS-CR genes to a mean of 0 and a variance of 1 across all 43 individuals. The OXPHOS-CR mean centroid vector is simply the mean of these 34 expression vectors. In some regression analyses, we introduced dummy variables to represent diabetes status. For the regressions we carried out, we report the adjusted squared correlation coefficient (R^2_{adj}), which corrects for the degrees of freedom.

URLs. Further details on microarray data sets and analysis are available at http://www-genome.wi.mit.edu/mpg/oxphos/. Further data on microarrays are available at http://www-genome.wi.mit.edu/cancer/, http://www-stat.stanford.edu/~tibs/SAM/ and http://www.affymetrix.com/. The gene expression atlas is available at http://expression.gnf.org/.

Note: Supplementary information is available on the Nature Genetics website.

COMPETING INTERESTS STATEMENT
The authors declare that they have no competing financial interests.

1. Zimmet, P. Globalization, coca-colonization and the chronic disease epidemic: can the Doomsday scenario be averted? *J. Intern. Med.* **247**, 301–310 (2000).
2. Fajans, S.S., Bell, G.I. & Polonsky, K.S. Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. *N. Engl. J. Med.* **345**, 971–980 (2001).
3. Golub, T.R. *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537 (1999).
4. Doniger, S.W. *et al.* MAPPFinder: using Gene Ontology and GenMAPP to create a global gene-expression profile from microarray data. *Genome Biol.* **4**, R7 (2003).
5. Draghici, S., Khatri, P., Martins, R.P., Ostermeier, G.C. & Krawetz, S.A. Global functional profiling of gene expression. *Genomics* **81**, 98–104 (2003).
6. Tusher, V.G., Tibshirani, R. & Chu, G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. USA* **98**, 5116–5121 (2001).
7. Sreekumar, R., Halvatsiotis, P., Schimke, J.C. & Nair, K.S. Gene expression profile in skeletal muscle of type 2 diabetes and the effect of insulin treatment. *Diabetes* **51**, 1913–1920 (2002).
8. Yang, X., Pratley, R.E., Tokraks, S., Bogardus, C. & Permana, P.A. Microarray profiling of skeletal muscle tissues from equally obese, non-diabetic insulin-sensitive and insulin-resistant Pima Indians. *Diabetologia* **45**, 1584–1593 (2002).
9. Kropf, S. & Lauter, J. Multiple tests for different sets of variables using a data-driven ordering of hypotheses, with an application to gene expression data. *Biometrical J.* **44**, 789–800 (2002).
10. Storey, J.D. A direct approach to false discovery rates. *J. R. Statist. Soc. B* **64**, 479–498 (2002).
11. Liu, G. *et al.* NetAffx: Affymetrix probesets and annotations. *Nucleic Acids Res.* **31**, 82–86 (2003).
12. Su, A.I. *et al.* Large-scale analysis of the human and mouse transcriptomes. *Proc. Natl. Acad. Sci. USA* **99**, 4465–4470 (2002).
13. Bjorntorp, P., Schersten, T. & Fagerberg, S.E. Respiration and phosphorylation of mitochondria isolated from the skeletal muscle of diabetic and normal subjects. *Diabetologia* **3**, 346–352 (1967).
14. Simoneau, J.A., Colberg, S.R., Thaete, F.L. & Kelley, D.E. Skeletal muscle glycolytic and oxidative enzyme capacities are determinants of insulin sensitivity and muscle composition in obese women. *FASEB J.* **9**, 273–278 (1995).
15. Puigserver, P. *et al.* A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* **92**, 829–839 (1998).
16. Wu, Z. *et al.* Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1. *Cell* **98**, 115–124 (1999).
17. Lin, J. *et al.* Transcriptional co-activator PGC-1α drives the formation of slow-twitch muscle fibres. *Nature* **418**, 797–801 (2002).
18. Scarpulla, R.C. Nuclear activators and coactivators in mammalian mitochondrial biogenesis. *Biochim. Biophys. Acta* **1576**, 1–14 (2002).
19. Brown, G.C. Control of respiration and ATP synthesis in mammalian mitochondria and cells. *Biochem. J.* **284**, 1–13 (1992).
20. Eriksson, K.F. & Lindgarde, F. Impaired glucose tolerance in a middle-aged male urban population: a new approach for identifying high-risk cases. *Diabetologia* **33**, 526–531 (1990).
21. Kelley, D.E., He, J., Menshikova, E.V. & Ritov, V.B. Dysfunction of mitochondria in human skeletal muscle in type 2 diabetes. *Diabetes* **51**, 2944–2950 (2002).
22. van Dam, R.M. & Feskens, E.J. Coffee consumption and risk of type 2 diabetes mellitus. *Lancet* **360**, 1477–1478 (2002).
23. Ojuka, E.O., Jones, T.E., Han, D.H., Chen, M. & Holloszy, J.O. Raising Ca^{2+} in L6 myotubes mimics effects of exercise on mitochondrial biogenesis in muscle. *FASEB J.* **17**, 675–681 (2003).
24. Florez, J.C., Hirschhorn, J.N. & Altshuler, D. The inherited basis of diabetes mellitus: general lessons for the genetic analysis of complex traits. *Annu. Rev. Genomics Hum. Genet.* (in the press).
25. Ek, J. *et al.* Mutation analysis of peroxisome proliferator-activated receptor-gamma coactivator-1 (PGC-1) and relationships of identified amino acid polymorphisms to Type II diabetes mellitus. *Diabetologia* **44**, 2220–2226 (2001).
26. Hara, K. *et al.* A genetic variation in the *PGC-1* gene and insulin resistance. *Diabetologia* **45**, 740–743 (2002).
27. Lacquemant, C., Chikri, M., Boutin, P., Samson, C. & Froguel, P. No association between the G482S polymorphism of the proliferator-activated receptor-γ coactivator-1 (*PGC-1*) gene and Type II diabetes in French Caucasias. *Diabetologia* **45**, 602–603; author reply 604 (2002).
28. Groop, L. *et al.* Metabolic consequences of a family history of NIDDM (the Botnia study): evidence for sex-specific parental effects. *Diabetes* **45**, 1585–1593 (1996).
29. Eriksson, K.F., Saltin, B. & Lindgarde, F. Increased skeletal muscle capillary density precedes diabetes development in men with impaired glucose tolerance. A 15-year follow-up. *Diabetes* **43**, 805–808 (1994).
30. Schalin-Jantti, C., Laurila, E., Lofman, M. & Groop, L.C. Determinants of insulin-stimulated skeletal muscle glycogen metabolism in man. *Eur. J. Clin. Invest.* **25**, 693–698 (1995).
31. Tamayo, P. *et al.* Interpreting patterns of gene expression with self-organizing maps: methods and application to hematopoietic differentiation. *Proc. Natl. Acad. Sci. USA* **96**, 2907–2912 (1999).
32. Eisen, M.B., Spellman, P.T., Brown, P.O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**, 14863–14868 (1998).