Diabetologia (2002) 45:1584–1593
DOI 10.1007/s00125-002-0905-7

Diabetologia

Microarray profiling of skeletal muscle tissues from equally obese, non-diabetic insulin-sensitive and insulin-resistant Pima Indians

X. Yang[2], **R. E. Pratley**[3], **S. Tokraks**[1], **C. Bogardus**[1], **P. A. Permana**[1]

[1] Clinical Diabetes and Nutrition Section, Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Phoenix, Arizona, USA
[2] Lundberg Laboratory for Diabetes Research, Sahlgrenska University Hospital, Göteborg, Sweden
[3] Department of Cardiovascular, Metabolic and Endocrine Clinical Research, Novartis Pharmaceutical Corp., East Hanover, New Jersey, USA

Abstract

Aims/hypothesis. We carried out global transcript profiling to identify differentially expressed skeletal muscle genes in insulin resistance, a major risk factor for Type II (non-insulin-dependent) diabetes mellitus. This approach also complemented the ongoing genomic linkage analyses to identify genes linked to insulin resistance and diabetes in Pima Indians.
Methods. We compared gene expression profiles of skeletal muscle tissues from 18 insulin-sensitive versus 17 insulin-resistant equally obese, non-diabetic Pima Indians using oligonucleotide arrays consisting of about 40,600 transcripts of known genes and expressed sequence tags, and analysed the results with the Wilcoxon rank sum test. We verified the mRNA expression of ten differentially (best-ranked) and ten similarly (worst-ranked) genes using quantitative Real Time PCR.
Results. There were 185 differentially expressed transcripts by the rank sum test. The differential expressions of two out of the ten best-ranked genes were confirmed and the similar expressions of all ten worst-ranked genes were reproduced.
Conclusion/interpretation. Of the 185 differentially expressed transcripts, 20 per cent were true positives and some could generate new hypotheses about the aetiology or pathophysiology of insulin resistance. Furthermore, differentially expressed genes in chromosomal regions with linkage to diabetes and insulin resistance serve as new diabetes susceptibility genes. [Diabetologia (2002) 45:1584–1593]

Keywords Genes, oligonucleotide array, RT-PCR, insulin resistance, diabetes.

Type II (non-insulin-dependent) diabetes mellitus is an increasingly common metabolic disease involving abnormal regulation of carbohydrate and lipid metabolism by insulin [1]. Insulin resistance, characterized as decreased insulin action on glucose uptake and metabolism, is a major predictor of Type II diabetes independent of obesity in Pima Indians of Arizona [2] who have a high prevalence of the disease [3]. Under physiological conditions, insulin-stimulated glucose metabolism occurs mainly in skeletal muscle (>80%) and adipose (~5–10%) tissues [4]. Similar to Type II diabetes, insulin resistance clusters in families [5], and is inherited as a non-Mendelian trait [6]. Since insulin resistance is a pre-diabetic phenotype, it is thought that genes influencing this metabolic abnormality could be fewer than those contributing to the complex diabetic syndrome. Chromosomal regions harbouring susceptibility genes for pre-diabetic phenotypes [7]

Received: 12 November 2001 / Revised: 26 April 2002
Published online: 23 August 2002

Corresponding author: Dr. P. A. Permana, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, 4212 N. 16th Street, Phoenix, AZ 85016, USA.
E-mail: ppermana@phx.niddk.nih.gov
Abbreviations: EST, Expressed sequence tag; IS, insulin sensitive; IR, insulin resistant; cRNA, complementary ribonucleic acid; Q-RT-PCR, Taqman Real Time PCR; SNP, single nucleotide polymorphism.

Table 1. Anthropometric and metabolic characteristics of the non-diabetic insulin-resistant (IR) and insulin-sensitive (IS) Pima Indian subjects

Characteristics	IS group	IR group	*p* value
Number of subjects (men/women)	11/7	12/5	
Age (year)	31±8	30±7	0.7
Height (cm)	167±7	170±10	0.4
Weight (kg)	108±25	106±23	0.9
Body fat (%)	36±6	33±6	0.2
Fasting glucose (mmol/l)	4.9±0.5	4.9±0.6	0.5
2-h glucose (mmol/l)	6.6±1.6	7.9±1.4	0.01
Fasting insulin (pmol/l)	230±50	310±120	0.01
2-h insulin (pmol/l)	745±305	1655±735	<0.0001
M-low ($mg \cdot min^{-1} \cdot kg^{-1}$ EMBS)	3.11±1.60	1.96±0.29	0.01
M-high ($mg \cdot min^{-1} \cdot kg^{-1}$ EMBS)	10.64±1.39	6.13±1.24	<0.0001

Data are expressed as means ± SD. *M-low* glucose disposal rate at physiological insulin concentration during a two-step hyperinsulinaemic-euglycaemic clamp, *EMBS* estimated metabolic body size, *M-high* glucose disposal rate at supra-physiological insulin concentration during a two-step hyperinsulinaemic-euglycaemic clamp

and diabetes [8] have been identified using linkage analyses in the Pima population.

The recent advances in genomic research include the utilization of microarrays to monitor the expression of thousands of genes in parallel. Global gene expression or transcriptional profiling has been used to identify molecular markers for various pathological states [9, 10, 11]. In this study we used global gene expression using oligonucleotide microarrays that included transcripts of known genes and Expressed Sequence Tags (ESTs) to identify potentially novel genes in pathways that are dysregulated at the transcript level in the skeletal muscle tissue of insulin-resistant non-diabetic Pima Indians. This approach provides data that can generate new hypotheses on the metabolic impairment that characterizes insulin resistance. Furthermore, this transcriptional profiling approach should complement the genomic linkage and positional cloning of diabetes susceptibility genes in the Pima Indian population by potentially identifying differentially expressed genes located on the chromosomal regions with suggestive linkage to diabetes and insulin resistance.

Subjects and methods

Subjects and clinical procedures. This study was approved by the Tribal Council of the Gila River Indian Community and the Institutional Review Board of the National Institutes of Diabetes and Digestive and Kidney Diseases (NIDDK). All subjects provided written informed consent prior to participation. The subjects in this study were non-diabetic Pima Indians or Tohono O'Odham Indians who were classified as insulin-sensitive (IS, *n*=18) or insulin-resistant (IR, *n*=17) based on a two-step hyperinsulinaemic-euglycaemic clamp and were matched for per cent body fat to minimize the differences due to obesity. The anthropometric and metabolic characteristics of the two groups at the time of the muscle biopsy are summarised in Table 1. All subjects were in good health as assessed by medical history, physical examination and routine blood and urine tests, and none were taking medication at the time of the study.

Volunteers were admitted to the Clinical Research Unit for 8 to 10 days. They were fed a weight maintaining diet (containing 50% of calories as carbohydrates, 30% as fat, and 20% as protein) for 2 to 3 days before metabolic testing. Body composition was measured by dual energy X-ray absorptiometry (DXA) using a total body scanner (DPX-L, Lunar Radiation, Madison, Wis., USA) [12]. Oral glucose tolerance tests using 75 g glucose were carried out after a 12-h overnight fast, and diabetes was defined according to the World Health Organization criteria [13]. Plasma glucose and insulin concentrations were measured in blood samples drawn before glucose ingestion and at 30, 60, 120 and 180 min thereafter. Insulin action was measured at physiologic and supraphysiologic insulin concentrations during a two-step hyperinsulinaemic-euglycaemic glucose clamp [2]. After an overnight fast, a primed continuous intravenous insulin infusion was administered for 100 min at a constant rate of 40 $mU \cdot m^{-2}$ body surface $area \cdot minute^{-1}$ (low dose), followed by a second insulin infusion for 100 min at 400 $mU \cdot m^{-2} \cdot minute^{-1}$ (high dose). These infusions achieved steady-state plasma insulin concentrations of 840±250 pmol/l and 13320±3480 pmol/l (Means ± SD), respectively. Plasma glucose concentrations were maintained at about 5.6 mmol/l with a variable infusion of a 20% glucose solution. Rates of insulin-stimulated glucose disposal at physiologic and maximally stimulating insulin concentrations were calculated for the last 40 min of each phase, and corrected for endogenous glucose output (EGO) [14]. During the low dose and baseline, EGO was calculated using a primed (1.11×10^6 Bq), continuous (1.11×10^4 $Bq \cdot min^{-1}$) 3-^{3}H-glucose infusion [2, 15]; during the high insulin dose, EGO was assumed to be 0. The glucose disposal rates during the clamp served as the selection criteria for the insulin-sensitive or resistant group. Subjects of a similar percentage of body fat with the highest or lowest glucose disposal rates were classified into the insulin-sensitive (IS) or insulin-resistant (IR) group, respectively. Indirect calorimetry using a ventilated hood system was used to calculate the rates of insulin-stimulated oxidative and non-oxidative glucose disposal [16]. All measurements derived from the glucose clamp were normalized to estimated metabolic body size (EMBS, which equals fat-free mass +17.7 kg) [17].

Percutaneous skeletal muscle biopsies of the vastus lateralis muscle were carried out in the morning after a 12-h overnight

fast using Bergstrom needles (Depuy) (Raynham, Mass., USA) under local anaesthesia with 1% lidocaine. The biopsy was cleaned of any visible fat, rinsed in sterile 0.9% NaCl solution, and immediately frozen in liquid nitrogen and stored at –70°C.

RNA preparation. Total RNA was isolated from the frozen tissues homogenized in Trizol Reagent (Life Technologies, Gaithersburg, Md., USA) and mRNA was subsequently isolated using oligo-dT latex beads (Qiagen, Santa Clarita, Calif., USA). The oligonucleotide microarray approach was chosen to allow comparison of multiple samples in IS and IR groups. Arrays that included about 40,600 transcripts of known genes and ESTs provided a comprehensive method to interrogate as many muscle transcripts as possible. Due to the high cost of the Affymetrix GeneChips, we pooled equal amounts mRNA of two to four subjects from either IS or IR group to make 1 ug mRNA for each GeneChip. This pooling strategy should also minimize individual variations. From the 17 IR and 18 IS subjects, we made 5 IS (IS1–5) and 5 IR (IR1–5) sample pools, respectively. Hybridization samples for subsequent GeneChip analysis were prepared as recommended by the manufacturer (Affymetrix, Santa Clara, Calif., USA). Briefly, double-stranded cDNA was synthesized from the mRNA samples using Superscript Choice system (Life Technologies, Gaithersburg, Md., USA). The cDNA served as a template for in vitro transcription reaction (Megascript kit from Ambion, Austin, Tex., USA) to generate biotinylated cRNA that included biotin-11-CTP and biotin-16-UTP (Enzo Biochemicals, Farmingdale, N.Y., USA). The cRNA was purified using RNeasy Mini kit (Qiagen, Santa Clarita, Calif., USA) and randomly fragmented by heat and alkaline treatment prior to hybridization to Affymetrix GeneChips.

For verification using TaqMan Real Time PCR (Q-RT-PCR), single-stranded oligo-dT primed cDNA were synthesized from either the available total RNA (pre-treated with DNA-free reagent from Ambion, Austin, Tex., USA) or mRNA of the majority of subjects using the Advantage RT-for-PCR kit (Clontech, Palo Alto, Calif., USA).

Hybridization, staining, scanning and analysis of image. The hybridization solution consisted of the 0.05 µg/µl fragmented cRNA sample, 50 pmol/l of a control biotinylated oligonucleotide for image alignment, biotin-labelled bacterial and phage cRNAs for hybridization control (1.5 pmol/l bioB, 5 pmol/l bioC, 25 pmol/l bioD, 100 pmol/l Cre), and 0.1 mg/ml degraded herring sperm DNA in hybridization buffer. The hybridization mixture was heated to 99°C for 5 min and equilibrated at 45°C for 5 min before hybridization in the oligonucleotide array chamber at 45°C for 16 to 17 h. Each hybridization sample was hybridized to the Affymetrix GeneChip Hu6800 Array (Santa Clara, Calif., USA) and Human 35 K set consisting of about 5600 unambiguous full-length cDNAs (after masking for the ambiguous probe set designs using the class AB mask as per the manufacturer's instruction that filtered out probe sets containing less than 10 unambiguous probe pairs in the Hu6800 array) and about 35,000 clustered human EST transcripts, respectively. After hybridization, the solution was removed and the probe arrays were washed and stained using the GeneChip Fluidics station protocol EukGE-WS2, as described previously [18]. The protocol consisted of non-stringent and stringent washes after hybridization, followed by a staining procedure using streptavidin-phycoerythrin solution (SAPE), and a post-stain wash. Signal amplification was achieved using antibody against streptavidin, after a final wash. The probe arrays were then scanned twice using a scanning confocal microscope (GeneChip scanner, Molecular Dynamics and Affymetrix, Santa Clara, Calif., USA). The stored images were aligned and analysed using the GeneChip software MAS 4.0 (Affymetrix) as per manufacturer's instruction. Signal intensity for each cDNA or EST entry (represented as "value" in Table 4) was calculated as the difference of filtered Perfect Match (PM) probes minus Mismatch (MM) probes. The signal intensities were normalized to the mean intensity of all the genes represented on the array, and then scaled to the image intensity of a randomly chosen sample pool (IS1) for that particular array. The comprehensive list on signal intensity and present/absent calling for each cDNA/EST in every sample pool can be found on http://www.ncbi.nlm.nih.gov/geo/; accession numbers GSE64–73.

GeneChip data analysis. We only considered cDNA or EST entries that were called "present" by the GeneChip software in at least three sample pools of one group (either IS or IR) versus none in the other group. For example, we would exclude an entry that was present in only two IS sample pools and two IR pools, but we would include an entry that was present in three IS pools versus no IR pool. Excluded entries were designated "below detection threshold." From the 40600 cDNAs/ESTs represented on the arrays, there were only 10831 cDNAs/ESTs that met the inclusion criteria. We then applied the non-parametric statistical analysis, Wilcoxon Rank Sum test to the average difference values of the cDNAs and ESTs in all sample pools, and listed those with a p value of less than or equal to 0.05 (hereafter called the best-ranked list).

Quantitative Q-RT-PCR. Verification of transcript quantity in several selected cDNAs/ESTs was carried out using TaqMan Real Time PCR (Q-RT-PCR) on cDNA from individual samples that had enough remaining total RNA or mRNA. The primer pairs and probe for each cDNA/EST were designed using Primer Express software (Applied Biosystems, Foster City, Calif., USA), and are available upon request. The quantification was carried out using the standard protocol of ABI PRISM 7700 (Applied Biosystems). For each primer and probe set, a standard curve was generated by a serial dilution of a cDNA sample synthesized from muscle RNA of a healthy subject that was done in triplicate. Each sample was run in duplicate and the mean value of the duplicate was used to calculate the mRNA expression. The transcript quantity of a particular cDNA/EST in each cDNA sample was normalized to that of cyclophilin using the TaqMan Pre-Developed Assay Reagent for human endogenous controls (Applied Biosystems).

Q-RT-PCR data analysis. Differential gene expression between the IR and the IS groups as measured using Q-RT-PCR was analysed using one-tailed Student's t test, since we hypothesized that the relative abundance of each transcript in the IS and the IR groups measured by the Q-RT-PCR method would be the same as that measured by the oligonucleotide array analysis.

Results

Oligonucleotide microarray analysis. There were 195 differentially expressed transcripts in IS versus IR groups as assessed by the rank sum test. These transcripts included nine ESTs that contained repetitive elements and one EST that was no longer documented in the GenBank database (http://www.ncbi.nih.gov/Genbank/). The remaining 185 transcripts were categorized based on their (putative) functions as well as their known chromosomal locations to complement

Table 2. The best-ranked functionally-known genes with higher expression in the insulin sensitive (IS) group

Gene name (Gene Symbol[a])	Chromosomal location
Insulin signalling	
Insulin receptor substrate 1 (IRS1)	2q36
Protein phosphatase 1, regulatory (inhibitor) subunit 11 (PPP1R11)	6p21.3
Protein phosphatase 1, regulatory (inhibitor) subunit 2 (PPP1R2)	3q29
Signal transduction	
FK506-binding protein 5 (FKBP5)	6
Disabled (Drosophila) homolog 1 (DAB1)	1p32-p31
Nuclear autoantigen (GS2NA)	14q13-q21
Tumour rejection antigen (gp96) 1 (TRA1)	12q24.2-q24.3
KIAA0382 protein; leukaemia-associated rho guanine nucleotide exchange factor (ARHGEF12)	11q23.3
Cell growth	
Insulin-like growth factor binding protein 5 (IGFBP5)	2q33–36
LIM domain only 4 (LMO4)	1p22
Ion transport	
ATP-binding cassette, sub-family B, member 10 (ABCB10)	1q42
ATPase, Ca++ transporting, type 2C, member 1 (ATP2C1)	3q21-q24
Potassium inwardly-rectifying channel, subfamily J, member 2 (KCNJ2)	17q23.1-q24.2
Energy metabolism	
NADH dehydrogenase (ubiquinone) 1; subcomplex unknown 1 (NDUFC1)	4q28-q31
ATP synthase, H+ transporting, mitochondrial F1 complex, gamma polypeptide 1 (ATP5C1)	10q22-q23
Lipid metabolism	
Glyceronephos-phate O-acyl-transferase (GNPAT)	1q42.11–42.3
Vacuolar sorting protein 4 (VPS4)	16
Transcription regulation	
General transcription factor IIIA (GTF3A)	13q12.3-q13.1
TATA box binding protein (TBP)-associated factor (TAF2D)	10q24-q25.2
Nuclear transcription factor Y, beta (NFYB)	12q22-q23
Zinc finger protein 161 (ZNF161)	3q26.2
FOXJ2 forkhead factor (LOC55810)	12pter-p13.31
Neural polypyrimidine tract binding protein (PTB)	1
Heat shock transcription factor 4 (HSF4)	16q21
MAX-interacting protein 1 (MXI1)	10q24-q25
Hairy/enhancer-of-split related with YRPW motif 1 (HEY1)	8q21
Cold shock domain protein A (CSDA)	12p13.1
Protein synthesis	
Signal recognition particle 9kD (SRP9)	1q41
Ribosomal protein L7 (RPL7)	8q
Nucleolin (NCL)	2q12-qter
Protein degradation	
Cathepsin F (CTSF)	11q13
Protective protein for beta-galactosidase (PPGB)	20q13.1
Proteasome activator subunit 3; PA28 gamma; Ki (PSME3)	17q12-q21
Ubiquitin specific protease 14 (tRNA-guanine transglycosylase) (USP14)	18
Ubiquitin C-terminal hydrolase (UCH37)	1q32
Homo sapiens ubiquitin protein ligase (UBE3B) mRNA, partial cds	12
Ariadne (Drosophila) homolog, ubiquitin-conjugating enzyme E2-binding protein, 1 (ARIH1)	15q24
Huntingtin interacting protein 2 (HIP2)	4p14
Cell adhesion	
Integrin α chain, alpha 6 (ITGA6)	2
Pinin, desmosome associated protein (PNN)	14
DNA replication	
Topoisomerase II binding protein 1 (TOPBP1)	3p13-q26.1

[a] Whenever possible, the HUGO (Human Genome Organization) nomenclature for each gene product is provided (62)

Table 3. The best-ranked functionally-known genes with higher mRNA expression in the insulin resistant (IR) group

Gene name (Gene Symbol)	Chromosomal location
Signal transduction	
A kinase (PRKA) anchor protein 11 (AKAP11)	13q12.2–13q14.3
A kinase (PRKA) anchor protein 2 (AKAP2)	9q31-q33
Platelet-derived growth factor receptor, beta polypeptide (PDGFRB)	5q31-q32
Endothelin receptor type B (EDNRB)	13q22
Carbohydrate metabolism	
Malic enzyme 1, NADP(+)-dependent, cytosolic (ME1)	6q12
Lipid metabolism	
Peroxisomal biogenesis factor 11B (PEX11B)	1p36.13-q24.1
Amino acid metabolism	
3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase (HMGCL)	1p36.1-p35
Iron metabolism	
Ferritin, light polypeptide (FTL)	19q13.3-q13.4
Transcription regulation	
PAI-1 mRNA-binding protein (PAI-RBP1)	1p31-p22
Transcription factor AP-2 alpha; activating enhancer-binding protein 2 alpha (TFAP2A)	6p24
Signal transducer and activator of transcription 5B (STAT5B)	17q11.2
Survival of motor neuron protein interacting protein 1 (SIP1)	14q13
Protein synthesis	
Mitochondrial ribosome recycling factor (MRRF)	9q32-q34.1
Diptheria toxin resistance protein (Saccharomyces)-like 2 (DPH2L2)	1p34
Protein degradation	
Ubiquitin-conjugating enzyme E2E3 (UBE2E3)	2q32.1
Prenylcysteine lyase (PCL1)	9q34.3
Palmitoyl-protein thioesterase 1 (PPT1)	1p32
Cytoskeletal function	
Myosin phosphatase, target subunit 1 (MYPT1)	12q15-q21
Fer-1 (C. elegans)-like 3 (FER1L3)	10q24
Epithelial protein lost in neoplasm beta (EPLIN)	12q13
Actin related protein 2/3 complex, subunit 5 (16 kD) (ARPC5)	1
Cell adhesion	
Integrin Beta-3; platelet glyco-protein IIIa, antigen CD61 (ITGB3)	17q21.32
Cell growth	
Ornithine decarboxylase antizyme 2 (OAZ2)	15q11.2
Apoptosis	
Death associated protein 3 (DAP3)	1q21

the positional cloning strategy for diabetes susceptibility gene(s). Subsets of functionally known genes with higher expression in either the IS or the IR group are listed (Table 2, 3).

The expression of some candidate genes previously suggested to be transcriptionally dysregulated in diabetic and/or insulin resistant people are shown in Table 4.

Q-RT-PCR verification. For verification of differential mRNA expression using a different method (Q-RT-PCR), we selected ten out of the best-ranked 185 differentially expressed transcripts (Table 5). The ten transcripts were called "present" in all ten sample pools, and contained sequence regions that allowed the design of specific primer-probe sets for this assay. Nine of the ten transcripts were derived from known cDNAs and the remaining transcript was an EST representing a gene with an unknown function. As a comparison, we also quantified ten transcripts that were similarly called "present" in all ten sample pools, but

Table 4. Comparison of several previously suggested differentially expressed human muscle genes at basal level with the microarray data

Gene (Gene Symbol)	Expression data in the literature		Signal intensity (value) in the microarray data		
	In diabetic subjects	In IR subjects	Means ± SD in IS group	Means ± SD in IR group	Rank sum *p* value
Phosphoprotein enriched in diabetes or phosphoprotein enriched in astrocytes 15 (PED/PEA15)	Increased (45)		335±59	487±146	0.1
Glucose transporter 1 (GLUT1)	No difference (47)	Increased (48)	202±76	352±162	0.1
Hexokinase II (HK2)	Decreased (49)		258±143	110±201	0.3
Glycogen synthase (GYS1)	Decreased (50)	No difference (51)	1311±336	1934±812	0.3
Ras associated with diabetes (RRAD)	Increased (52)	No difference (53)	508 ± 444	338 ± 529	0.7
Uncoupling protein 3 (UCP3)	Increased (54)	No difference (54)	541±259	483±159	0.7
calpain10 (CAPN10)		Decreased (55)	Below detection threshold		
Glucose transporter 3 (GLUT3)		Decreased (48;53)	Below detection threshold		
Uncoupling protein 2 (UCP2)	Increased (53;54)	No difference (53;54)	Below detection threshold		
Tumour necrosis factor (TNF)	Increased (56)	Increased (56)	Below detection threshold		
Peroxisome proliferator activated receptor γ (PPARγ)	Increased (57)		Below detection threshold		
Insulin receptor (INSR)	Increased isoform B (58)		Masked probe sets[a]		
Glycogen-associated regulatory subunit of protein phosphatase 1 (PPP1R3)	No difference (59)	Decreased (60)	Masked probe sets[a]		
Glycoprotein plasma cell 1 or PC-1 (ENPP1)		Increased (61)	Not represented		

[a] Ambiguous probe set design; taken out from analysis according to manufacturer's instructions

were not differentially expressed as assessed by the rank sum test (i.e. worst-ranked) in the same individual cDNA samples (Table 5). The ten worst-ranked transcripts were chosen to be in a similar range of expression in the GeneChip analysis as the ten best-ranked ones. We measured the expression of these transcripts in individual samples (13 IS and 10–12 IR) that had enough remaining RNA available. The average values of each transcript in the IS and the IR groups obtained by the GeneChip analysis and the Q-RT-PCR method were compared.

Out of the ten best-ranked transcripts that were differentially expressed in the IS and IR groups by GeneChip analysis, two transcripts, Insulin Receptor Substrate 1 (IRS1) and Insulin-like Growth Factor Binding Protein 5 (IGFBP5), were confirmed to be differentially expressed by Q-RT-PCR (Table 5). None of the ten worst-ranked transcripts showed a difference between the IS and IR groups by Q-RT-PCR (Table 5).

Discussion

We have used global microarray analysis to identify potential candidate genes for insulin resistance in skeletal muscle tissues of Pima Indians. We analysed the data above detection threshold using the Rank Sum Test to select the transcripts that were differentially expressed (the best-ranked transcripts). The resulting number of 195 transcripts was fewer than the expected 541 transcripts (out of 10,831) that should have achieved a *p* value of less than or equal to 0.05. This could be due, at least partly, to redundancies in the cDNA/EST representation on the probe arrays resulting in less than 10,831 unique transcripts. Knowledge of new genes that are differentially regulated in insulin resistance can generate new hypotheses on molecular mechanisms of the syndrome. Furthermore, differentially expressed genes in chromosomal regions with suggestive linkage to diabetes and insulin action in Pima and other populations would serve as candidate susceptibility genes for insulin resistance and diabetes.

We also checked the expression of some genes previously suggested to be transcriptionally dysregulated in diabetic and insulin-resistant people, which could be either primary causes or secondary effects of the insulin resistance state, e.g. modulated by chronically higher plasma insulin or glucose concentrations.

Table 5. Expression of selected ten best-ranked and ten worst-ranked genes as assessed by the oligonucleotide array and Q-RT-PCR

Probe set	Gene description	Rank sum *p* value	*t* test *p* value
S62539	Insulin receptor substrate 1 (IRS1)	0.02	0.04
L27559	Insulin-like growth factor binding protein 5 (IGFBP5)	0.01	0.06
AA460511	Ubiquitin-conjugating enzyme E2E3 (UBE2E3)	0.03	0.1
M60858	Nucleolin (NCL)	0.01	0.3
L07648	MAX-interacting protein 1 (MXI1)	0.01	0.3
X57959	Ribosomal protein L7 (RPL7)	0.03	0.3
AA435791	Speckle-type POZ protein (SPOP)	0.02	0.4
D57916	Clone 24775 mRNA	0.01	0.5
U20998	Signal recognition particle 9kD (SRP9)	0.03	0.5
L07033	3-hydroxymethyl-3-methylglutaryl-Coenzyme A lyase (HMGCL)	0.01	0.5
D21235	RAD23 (S. cerevisiae) homolog A (RAD23A)	1.0	0.2
U36764	Eukaryotic translation initiation factor 3, subunit 2 (beta, 36kD) (EIF3S2)	1.0	0.3
L49054	Myeloid leukaemia factor 1 (MLF1)	1.0	0.3
U58089	Cullin 3 (CUL3)	1.0	0.3
V00599	Tubulin, beta 5 (Tubb5)	1.0	0.4
AA417126	Translocase of inner mitochondrial membrane 10 (yeast) homolog (TIMM10)	1.0	0.4
D13900	Enoyl Coenzyme A hydratase, short chain, 1, mitochondrial (ECHS1)	1.0	0.4
AA448347	Prefoldin 2 (PFDN2)	1.0	0.5
D21853	KIAA0111 gene product (KIAA0111)	1.0	0.5
U28963	G protein pathway suppressor 2 (GPS2)	1.0	0.5

The first ten probe sets or gene entries (in bold font) are in the best-ranked list; the last ten entries are in the worst-ranked list. For GeneChip analysis, n=17 IR and 18 IS. For Q-RT-PCR analysis, n=10–12 IR and 13 IS

There were no statistically significant differences in the expression of these genes in the microarray data, which could be due to the small number of samples in each group and lower sensitivity of microarray technology compared to RT-PCR [19]. The expression of some genes are below detection threshold, and determination of differential expression for such probe sets is deemed unreliable [20].

The expensive cost of the GeneChips did not allow experimental replication that would be necessary to assess the signal intensity variance for each probe set and the traditional confidence level. As there were thousands of genes represented on the GeneChips, the statistical analysis involved multiple testing issues. A traditional way to test for statistical significance in the face of such multiple testing is to apply the Bonferroni correction or a related step-up or step-down procedure [21], which would result in very small alpha levels. For example, an experiment-wise alpha level of 0.05 with 10,000 genes would require a p-value of 5×10^{-6}; this might result in absurdly low power for realistic sample sizes and would disregard many biologically significant changes [22]. Thus, we estimated the number of genes with true expression differences between the IS and IR groups by empirical testing using a different method (Q-RT-PCR).

We carried out Q-RT-PCR on the ten best-ranked and ten worst-ranked transcripts in the majority of subjects individually. This empirical testing confirmed about 20% of the differentially expressed genes and 100% of the similarly expressed genes. Despite the confirmation of true negative results in the data set above detection threshold, the 20% enrichment in true positive result was low and indicated that the differential expression of the other best-ranked genes has to be confirmed with a different method or a larger number of samples. The high false positive rate could be due to either type I error in the microarray analysis or type II error in the RT-PCR analysis. Nevertheless, the 185 potentially differentially expressed genes can be used to generate new hypotheses on molecular features of insulin resistance and to complement the positional cloning effort on diabetes susceptibility genes.

Insulin receptor substrate 1 (IRS1) was one of the best-ranked transcripts confirmed to be differentially expressed between the IR and the IS groups. Considering that IRS1 mRNA was reduced by a higher plasma insulin concentration during a euglycaemic-hyperinsulinaemic clamp in healthy subjects [23], our result can be explained by normal insulin regulation of IRS1 expression in the IR subjects, i.e. the decreased muscle IRS1 expression is secondary to hyperinsulinaemia induced by insulin resistance. This explanation would support the hypothesis that muscle IRS1 regulation is not affected by insulin resistance in diabetic patients [24]. Alternatively, the reduced IRS1 expression in the IR subjects could play a role in the development of insulin resistance in non-diabetic subjects and could be due to genetically determined abnormal regulation of the *IRS1* gene or altered function of transcription factors regulating *IRS1* transcription. This explanation, together with the gene location on 2q36 in the vicinity

of a suggestive linkage to insulin action in Pima Indians [7], lead to the possibility of a polymorphism(s) in the *IRS1* gene that regulates its expression. The frequency of the Gly972Arg polymorphism in the *IRS1* gene is higher in Type II diabetic patients of some, but not all, studied ethnic groups [25, 26, 27, 28]. This polymorphism is absent, or at least very rare, in Pima Indians [29]; thus, the differentially regulated IRS1 expression in the subjects of our study was not associated with Gly972Arg polymorphism. The *IRS1* gene has been screened for single nucleotide polymorphisms (SNPs), and preliminary data indicate that SNPs in the gene are associated with diabetes in Pima Indians [30].

IRS1 was one of the transcripts in insulin signaling pathway with higher expression in the IS group. The other transcripts encode PPP1R11 and PPP1R2, regulatory subunits of protein phosphatase-1 (PP1), a key enzyme that regulates the activity of glycogen synthase. PPP1R2 has been proposed to act as a molecular "chaperone" that aids the folding of newly synthesized PP1 into a biologically active conformation [31, 32] and it has been shown to translocate to the nucleus during the S-phase of cell cycle [33]. There were no polymorphisms in the exons and exon-intron splice junctions of the *PPP1R2* gene in selected Pima Indian subjects [34] but potential polymorphisms in the regulatory region of the gene might account for the differential gene expression. PPP1R11 is a heat-stable inhibitor of PP1 encoded by a gene on chromosome 6p21 that was previously known as the Hemochromatosis Candidate Gene V [35]. Of interest, an apparent pseudogene of PPP1R2 is also located on chromosome 6p21 [36, 37]. Both PPP1R11 and PPP1R2 serve as candidate genes for insulin resistance.

Apart from IRS1, the other gene in the ten selected best-ranked list with confirmed lower mRNA expressions in the IR subjects was IGFBP5 (on chromosome 2q33–36), a modulator of insulin growth factor 1 in inducing muscle differentiation [38, 39, 40]. The serum protein concentration of IGFBP5 was lower in diabetic patients compared to the control subjects [41], and diabetic rat kidney contains reduced IGFBP5 mRNA concentrations [42]. Nevertheless, the expression of IGFBP5 in skeletal muscle is not altered acutely by nutrients and insulin [43]. Note that a transcriptional regulator of myogenesis, LIM domain Only 4 (LMO4), was assessed to be lower in the IR group by the GeneChip analysis. These findings lead to a hypothesis that insulin resistance is associated with dysregulation of myogenic development.

One of the objectives of this study is to complement the positional cloning effort currently ongoing in the Pima population. The genes encoding many of these transcripts had been localized to particular chromosomal regions, and a few are in regions with suggestive linkage to diabetes and insulin action in Pima Indians. One such gene on chromosome 11q23 encodes a rho guanine nucleotide exchange factor (ARHGEF12). Preliminary findings indicate an association between several SNPs in the gene with insulin sensitivity in Pima Indians [44]. This result provides evidence for the utility of transcriptional profiling to identify new candidate genes as a complementary approach to positional cloning.

Another chromosomal region with suggestive linkage to diabetes in the Pima population is 1q21 that harbours many genes, including phosphoprotein enriched in diabetes or phosphoprotein enriched in astrocytes 15 (PED/PEA15). The mRNA concentration of PED/PEA15 is increased in fibroblasts, skeletal muscle and adipose tissue of Type II diabetic subjects [45]. In line with this finding, our GeneChip analysis indicated that the PED/PEA15 was expressed marginally higher in the IR versus IS groups. Despite the recent finding that several SNPs in the non-coding region of the *PEA15* gene were not associated with diabetes in 50 affected and 50 control Pimas [46], this could still be a good candidate gene for insulin resistance.

In summary, we have carried out a global transcriptional profiling of insulin-resistant and insulin-sensitive skeletal muscle tissue. This approach, coupled with the current metabolic knowledge and the positional cloning efforts, provides several novel candidate genes that can generate new hypotheses on the pathophysiology of insulin resistance and the development of Type II diabetes.

Acknowledgements. We would like to thank the members and leaders of the Gila River Indian Community, without whose cooperation these studies would have been impossible. We gratefully acknowledge M. Milner, PA, C. Massengill, RN, the nurses of the Clinical Research Ward, and Dr. A. Salbe and the metabolic kitchen staff for the care of the research volunteers, and Drs. R. Hanson and D. Allison for helpful discussions.

References

1. Saltiel AR (2001) New perspectives into the molecular pathogenesis and treatment of type 2 diabetes. Cell 104: 517–529
2. Lillioja S, Mott DM, Spraul M et al. (1993) Insulin resistance and insulin secretory dysfunction as precursors of non-insulin-dependent diabetes mellitus. Prospective studies of Pima Indians. N Engl J Med 329:1988–1992
3. Knowler WC, Bennett PH, Hamman RF, Miller M (1978) Diabetes incidence and prevalence in Pima Indians: a 19-fold greater incidence than in Rochester, Minnesota. Am J Epidemiol 108:497–505
4. Reaven GM (1988) Banting lecture 1988. Role of insulin resistance in human disease. Diabetes 37:1595–1607
5. Lillioja S, Bogardus C (1988) Obesity and insulin resistance: lessons learned from the Pima Indians. Diabetes Metab Rev 4:517–540
6. Kahn CR, Vicent D, Doria A (1996) Genetics of non-insulin-dependent (type-II) diabetes mellitus. Annu Rev Med 47:509–531

7. Pratley RE, Thompson DB, Prochazka M et al. (1998) An autosomal genomic scan for loci linked to prediabetic phenotypes in Pima Indians. J Clin Invest 101:1757–1764
8. Hanson RL, Ehm MG, Pettitt DJ et al. (1998) An autosomal genomic scan for loci linked to type II diabetes mellitus and body-mass index in Pima Indians. Am J Hum Genet 63:1130–1138
9. Alon U, Barkai N, Notterman DA et al. (1999) Broad patterns of gene expression revealed by clustering analysis of tumor and normal colon tissues probed by oligonucleotide arrays. Proc Natl Acad Sci USA 96:6745–6750
10. Thykjaer T, Workman C, Kruhoffer M et al. (2001) Identification of gene expression patterns in superficial and invasive human bladder cancer. Cancer Res 61:2492–2499
11. Wilson SB, Kent SC, Horton HF et al. (2000) Multiple differences in gene expression in regulatory Valpha 24Jalpha Q T cells from identical twins discordant for type I diabetes. Proc Natl Acad Sci USA 97:7411–7416
12. Tataranni PA, Ravussin E (1995) Use of dual-energy X-ray absorptiometry in obese individuals. Am J Clin Nutr 62:730–734
13. Alberti KG, Zimmet PZ (1998) Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. Diabet Med 15:539–553
14. Weyer C, Bogardus C, Mott DM, Pratley RE (1999) The natural history of insulin secretory dysfunction and insulin resistance in the pathogenesis of type 2 diabetes mellitus. J Clin Invest 104:787–794
15. Bogardus C, Lillioja S, Howard BV, Reaven G, Mott D (1984) Relationships between insulin secretion, insulin action, and fasting plasma glucose concentration in nondiabetic and noninsulin-dependent diabetic subjects. J Clin Invest 74:1238–1246
16. Bogardus C, Lillioja S, Ravussin E et al. (1986) Familial dependence of the resting metabolic rate. N Engl J Med 315:96–100
17. Gerich JE (1998) The genetic basis of type 2 diabetes mellitus: impaired insulin secretion versus impaired insulin sensitivity. Endocr Rev 19:491–503
18. Warrington JA, Nair A, Mahadevappa M, Tsyganskaya M (2000) Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. Physiol Genomics 2:143–147
19. Valerius MT, Patterson LT, Witte DP, Potter SS (2002) Microarray analysis of novel cell lines representing two stages of metanephric mesenchyme differentiation. Mech Dev 112:219–232
20. Schadt EE, Li C, Su C, Wong WH (2000) Analyzing high-density oligonucleotide gene expression array data. J Cell Biochem 80:192–202
21. Hochberg Y, Tamhane AC (1987) Multiple comparison procedures. John Wiley & Sons, New York
22. Claverie JM (1999) Computational methods for the identification of differential and coordinated gene expression. Hum Mol Genet 8:1821–1832
23. Ducluzeau PH, Perretti N, Laville M et al. (2001) Regulation by insulin of gene expression in human skeletal muscle and adipose tissue. Evidence for specific defects in type 2 diabetes. Diabetes 50:1134–1142
24. Huang X, Vaag A, Hansson M, Groop L (2002) Down-regulation of insulin receptor substrates (IRS)-1 and IRS-2 and Src homologous and collagen-like protein Shc gene expression by insulin in skeletal muscle is not associated with insulin resistance or type 2 diabetes. J Clin Endocrinol Metab 87:255–259
25. Almind K, Bjorbaek C, Vestergaard H, Hansen T, Echwald S, Pedersen O (1993) Aminoacid polymorphisms of insulin receptor substrate-1 in non-insulin-dependent diabetes mellitus. Lancet 342:828–832
26. Hitman GA, Hawrami K, McCarthy MI et al. (1995) Insulin receptor substrate-1 gene mutations in NIDDM; implications for the study of polygenic disease. Diabetologia 38:481–486
27. Shimokawa K, Kadowaki H, Sakura H et al. (1994) Molecular scanning of the glycogen synthase and insulin receptor substrate-1 genes in Japanese subjects with non-insulin-dependent diabetes mellitus. Biochem Biophys Res Commun 202:463–469
28. Ito K, Katsuki A, Furuta M et al. (1999) Insulin sensitivity is not affected by mutation of codon 972 of the human IRS-1 gene. Horm Res 52:230–234
29. Celi FS, Silver K, Walston J, Knowler WC, Bogardus C, Shuldiner AR (1995) Lack of IRS-1 codon 513 and 972 polymorphism in Pima Indians. J Clin Endocrinol Metab 80:2827–2829
30. Kovacs P, Yang X, Permana PA, Kobes S, Bogardus C, Baier LJ (2002) Variants in the insulin receptor substrate-1 gene are associated with type 2 diabetes in Pima Indians (abstract). Diabetes 51 (Suppl 2):A266
31. Alessi DR, Street AJ, Cohen P, Cohen PT (1993) Inhibitor-2 functions like a chaperone to fold three expressed isoforms of mammalian protein phosphatase-1 into a conformation with the specificity and regulatory properties of the native enzyme. Eur J Biochem 213:1055–1066
32. MacKintosh C, Garton AJ, McDonnell A et al. (1996) Further evidence that inhibitor-2 acts like a chaperone to fold PP1 into its native conformation. FEBS Lett 397:235–238
33. Kakinoki Y, Somers J, Brautigan DL (1997) Multisite phosphorylation and the nuclear localization of phosphatase inhibitor 2-green fluorescent protein fusion protein during S phase of the cell growth cycle. J Biol Chem 272:32308–32314
34. Permana PA, Mott DM (1997) Genetic analysis of human type 1 protein phosphatase inhibitor 2 in insulin-resistant Pima Indians. Genomics 41:110–114
35. Zhang J, Zhang L, Zhao S, Lee EY (1998) Identification and characterization of the human HCG V gene product as a novel inhibitor of protein phosphatase-1. Biochemistry 37:16728–16734
36. Helps NR, Street AJ, Elledge SJ, Cohen PT (1994) Cloning of the complete coding region for human protein phosphatase inhibitor 2 using the two hybrid system and expression of inhibitor 2 in E. coli. FEBS Lett 340:93–98
37. Sanseau P, Jackson A, Alderton RP et al. (1994) Cloning and characterization of human phosphatase inhibitor-2 (IPP-2) sequences. Mamm Genome 5:490–496
38. James PL, Stewart CE, Rotwein P (1996) Insulin-like growth factor binding protein-5 modulates muscle differentiation through an insulin-like growth factor-dependent mechanism. J Cell Biol 133:683–693
39. Rotwein P, James PL, Kou K (1995) Rapid activation of insulin-like growth factor binding protein-5 gene transcription during myoblast differentiation. Mol Endocrinol 9:913–923
40. Imai Y, Moralez A, Andag U, Clarke JB, Busby WH Jr, Clemmons DR (2000) Substitutions for hydrophobic amino acids in the N-terminal domains of IGFBP-3 and -5 markedly reduce IGF-I binding and alter their biologic actions. J Biol Chem 275:18188–18194
41. Jehle PM, Jehle DR, Mohan S, Bohm BO (1998) Serum levels of insulin-like growth factor system components and relationship to bone metabolism in Type 1 and Type 2 diabetes mellitus patients. J Endocrinol 159:297–306

42. Bach LA, Dean R, Youssef S, Cooper ME (2000) Aminoguanidine ameliorates changes in the IGF system in experimental diabetic nephropathy. Nephrol Dial Transplant 15:347–354
43. Coenen Schimke JM, Ljungqvist OH, Sarkar G, Conover CA, Sreekumaran NK (1999) A quantitative PCR measurement of messenger RNA expression of IGF-I, IGF-II and IGFBP-5 in human skeletal muscle. Growth Horm IGF Res 9:179–186
44. Kovaks P, Yang X, Permana PA, Bogardus C, Baier LJ (2001) A Rho-associated guanine exchange factor is associated with insulin resistance in non-diabetic Pima Indians. Diabetes 50:A245
45. Condorelli G, Vigliotta G, Iavarone C et al. (1998) PED/PEA-15 gene controls glucose transport and is overexpressed in type 2 diabetes mellitus. EMBO J 17:3858–3866
46. Wolford JK, Bogardus C, Ossowski V, Prochazka M (2000) Molecular characterization of the human PEA15 gene on 1q21-q22 and association with type 2 diabetes mellitus in Pima Indians. Gene 241:143–148
47. Pedersen O, Bak JF, Andersen PH et al. (1990) Evidence against altered expression of GLUT1 or GLUT4 in skeletal muscle of patients with obesity or NIDDM. Diabetes 39: 865–870
48. Stuart CA, Wen G, Williamson ME et al. (2001) Altered GLUT1 and GLUT3 gene expression and subcellular redistribution of GLUT4: protein in muscle from patients with acanthosis nigricans and severe insulin resistance. Metabolism 50:771–777
49. Vestergaard H, Bjorbaek C, Hansen T, Larsen FS, Granner DK, Pedersen O (1995) Impaired activity and gene expression of hexokinase II in muscle from non-insulin-dependent diabetes mellitus patients. J Clin Invest 96:2639–2645
50. Vestergaard H, Bjorbaek C, Andersen PH, Bak JF, Pedersen O (1991) Impaired expression of glycogen synthase mRNA in skeletal muscle of NIDDM patients. Diabetes 40:1740–1745
51. Majer M, Mott DM, Mochizuki H et al. (1996) Association of the glycogen synthase locus on 19q13 with NIDDM in Pima Indians. Diabetologia 39:314–321
52. Reynet C, Kahn CR (1993) Rad: a member of the Ras family overexpressed in muscle of type II diabetic humans. Science 262:1441–1444
53. Paulik MA, Hamacher LL, Yarnall DP et al. (1997) Identification of Rad's effector-binding domain, intracellular localization, and analysis of expression in Pima Indians. J Cell Biochem 65:527–541
54. Bao S, Kennedy A, Wojciechowski B, Wallace P, Ganaway E, Garvey WT (1998) Expression of mRNAs encoding uncoupling proteins in human skeletal muscle: effects of obesity and diabetes. Diabetes 47:1935–1940
55. Baier LJ, Permana PA, Yang X et al. (2000) A calpain-10 gene polymorphism is associated with reduced muscle mRNA levels and insulin resistance. J Clin Invest 106:R69–R73
56. Saghizadeh M, Ong JM, Garvey WT, Henry RR, Kern PA (1996) The expression of TNF alpha by human muscle. Relationship to insulin resistance. J Clin Invest 97:1111–1116
57. Park KS, Ciaraldi TP, Abrams-Carter L, Mudaliar S, Nikoulina SE, Henry RR (1997) PPAR-gamma gene expression is elevated in skeletal muscle of obese and type II diabetic subjects. Diabetes 46:1230–1234
58. Kellerer M, Sesti G, Seffer E et al. (1993) Altered pattern of insulin receptor isotypes in skeletal muscle membranes of Type II (non-insulin-dependent) diabetic subjects. Diabetologia 36:628–632
59. Chen YH, Hansen L, Chen MX et al. (1994) Sequence of the human glycogen-associated regulatory subunit of type 1 protein phosphatase and analysis of its coding region and mRNA level in muscle from patients with NIDDM. Diabetes 43:1234–1241
60. Xia J, Scherer SW, Cohen PT et al. (1998) A common variant in PPP1R3 associated with insulin resistance and type 2 diabetes. Diabetes 47:1519–1524
61. Frittitta L, Sbraccia P, Costanzo BV et al. (2000) High insulin levels do not influence PC-1 gene expression and protein content in human muscle tissue and hepatoma cells. Diabetes Metab Res Rev 16:26–32
62. White JA, McAlpine PJ, Antonarakis S et al. (1997) Guidelines for human gene nomenclature (1997). HUGO Nomenclature Committee. Genomics 45:468–471