Evaluation of molecular profiling platforms in clinical pharmacology

Doorn, M.B.A. van

24 Lehrke M, Lazar MA. The many faces of PPARgamma. *Cell* 2005; **123**:993-999
25 Mootha VK, Lindgren CM, Eriksson KF *et al*. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat.Genet.* 2003; **34**:267-273
26 Patti ME, Butte AJ, Crunkhorn S *et al*. Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes: Potential role of PGC1 and NRF1. *Proc.Natl.Acad.Sci.U.S.A* 2003; **100**:8466-8471.

CHAPTER 7

Evaluation of global gene expression profiles as biomarkers for the pharmacological effects of fibrates in patients with type 2 Diabetes Mellitus and healthy volunteers

---

SUBMITTED

M. B. A. van Doorn[1], Q.S. Li[2], X. Liu[3], D. M. Ouwens[4], E. J. van Hoogdalem[5], A. Bittner[3], C. R. Jones[6], J. A. Romijn[7], J. Burggraaf[1] and A. F. Cohen[1]

1 *Centre for Human Drug Research, Leiden, The Netherlands*
2 *Johnson & Johnson Pharmaceutical Research & Development, Raritan, NJ, USA*
3 *Johnson & Johnson Pharmaceutical Research & Development, La Jolla, CA, USA*
4 *Leiden University Medical Centre; Department of Molecular Cell Biology, Leiden, The Netherlands*
5 *Johnson & Johnson Pharmaceutical Research & Development, Beerse, Belgium*
6 *Biopharma – PS, Basel, Switzerland;*
7 *Leiden University Medical Centre; Department of Internal Medicine, Leiden, The Netherlands.*

Abstract

We evaluated the usefulness of global gene expression profiling as a method to obtain a broad array of molecular biomarkers for the early clinical evaluation of novel peroxisome proliferator activated receptor alpha (PPARα) agonists in patients with type 2 Diabetes Mellitus (T2DM) and in healthy volunteers (HVs). Additionally, we explored differences in gene expression profile between T2DM patients and HVs at baseline. We used cDNA microarrays to evaluate the expression of 5954 genes in skeletal muscle, adipose tissue, and peripheral blood leukocytes (PBLs) of T2DM patients and HVs receiving ciprofibrate 100 mg or placebo once daily for 3 weeks. None of the genes in any of the three tissues showed a significant response to treatment with ciprofibrate *vs.* placebo. In contrast, baseline comparisons identified various genes in muscle and adipose tissue that were significantly differentially expressed between T2DM patients and HVs. We conclude that gene expression profiling using cDNA microarray technology is currently unable to provide robust molecular biomarkers for the action of PPARα agonists that are useful for early clinical 'proof of concept' studies. Nonetheless, the technology may be useful for disease exploration using pathway level analyses.

Introduction

Recent developments in molecular biology have provided important new tools for the assessment of global gene expression profiles. One of the most important developments is microarray technology which is capable of surveying the changes in expression of thousands of genes in a single measurement [1]. This technology has proven to be quite useful in clinical oncology, because several studies have shown that clinical outcome can be predicted on the basis of tumour gene expression profiles [2-5]. In addition, Gunther et al. demonstrated successful application of this technology in drug development by showing that clinical efficacy of psychoactive compounds could be predicted based on gene expression profiles of human neurons *in vitro* [6;7]. Moreover, recently a large gene expression database has been produced that characterizes the gene expression and physiological effects of hundreds of approved and withdrawn drugs, toxicants, and biochemical standards in various organs of live rats [8]. Nonetheless, relatively few publications have reported on the application of microarray technology in clinical pharmacology studies with the aim to evaluate drug effects at the transcriptional level in humans [9-11].

It is commonly assumed that microarray technology is capable of capturing the broad information about the *in vivo* effects of a drug into a specific gene expression 'fingerprint'. This assessment could be particularly useful for drugs - such as peroxisome proliferator activated receptor (PPAR) agonists - that regulate gene expression in their target tissues, and thus are expected to induce readily measurable changes in mRNA abundance. These changes in gene expression profiles could serve as important 'mechanism based' biomarkers for the pleiotropic effects of such drugs *in vivo*.

To assess the potential of (cDNA) microarray based gene expression profiling as a reliable tool for the evaluation of novel compounds, in this case the two clinically relevant isoforms (γ and α) of the PPAR agonist class, we performed two studies with so-called 'prototype' drugs. The first study investigated the effects of rosiglitazone, as prototype for the thiazolidinedione class, on global gene expression profiles in adipose tissue, skeletal muscle and peripheral blood leukocytes (PBLs) of type 2 Diabetes Mellitus (T2DM) patients and healthy volunteers (HVs). The results of that study were submitted for publication elsewhere. By analogy, the present study investigated the effects of the PPARα agonist ciprofibrate on global tissue gene expression profiles and was designed to address the following questions:

Is it possible to detect statistically robust changes in gene expression profiles ('treatment fingerprint') in adipose tissue and skeletal muscle samples of T2DM patients after 3 weeks of *in vivo* treatment with the PPARα agonist ciprofibrate *vs.* placebo? And if such a fingerprint is found, can it also be detected in more easily recruited HVs? Furthermore, can the anticipated change in gene expression profile in skeletal muscle adipose tissue also be observed in peripheral blood leukocytes (PBLs), which are easier to obtain than tissue samples? And finally, is it possible to identify 'enriched' (i.e. overlapping) T2DM-associated biological functions and pathways by comparing differences in gene expression profile between T2DM and HVs at baseline between the present and our previous microarray study?

Methods

Patients

Eight male and eight female patients with an established diagnosis of T2DM (i.e. increased fasting plasma glucose (FPG) concentrations >7.0 mm) uncontrolled by diet alone, aged between 18 and 75 years, and a fasting plasma triglyceride concentration > 1.5 mm (>133 mg/dL), were included. Patients were excluded if they had a significant medical history or current symptoms of clinically relevant conditions, or had used any non-steroidal anti-inflammatory drug, PPAR agonist (α,ß/∂ or γ) or lipid lowering medication within two weeks prior to the expected study start date.

In addition, eight male and eight female healthy subjects (as determined by medical history, physical examination and routine laboratory tests), aged between 18 and 45 years, were included.

Study design

The study subjects were studied in a four-week period consisting of six Visits (figure 1; p. 218). Within one week after medical screening all subjects started with a one-week single blind placebo run-in period. At the end of the run-in period (baseline) subjects were randomly assigned to a three-week treatment with 100 mg ciprofibrate or matching placebo once daily prior to breakfast. Treatment allocation took place according to randomly permuted blocks and was stratified by gender and subject type (T2DM patient or HV).

At baseline and after three weeks of treatment, skeletal muscle and abdominal subcutaneous adipose tissue samples were collected as previously described [12]. After sampling, the tissue samples were immediately snap-frozen in liquid nitrogen and stored at -80 °C until further analysis.

Blood samples for plasma efficacy parameters (lipid profile and inflammation markers), ciprofibrate trough levels and transcriptomic analyses in PBLs were collected on all Visits after approximately 30 min of supine rest (for details see Chapter 4). Four blood samples were collected at baseline (at 1, 3, 6 and 10 hrs post-dose) to investigate the immediate gene expression response in PBLs and to assess ciprofibrate's pharmacokinetic profile. PBLs were isolated from whole blood samples and RNA was extracted immediately and stored at -80 °C until analysis.

The protocol for this study was approved by the Medical Ethics Committee of the Leiden University Medical Centre and performed according to the principles of ICH-GCP, the Helsinki Declaration and Dutch law. All subjects gave their written informed consent.

Analytical methods for plasma markers

Concentrations of plasma lipid / (apo)lipoprotein and non-lipid plasma markers were determined as previously described in Chapter 4.

Compliance monitoring and pharmacokinetics

Compliance monitoring was performed using an electronic event monitoring system (Aardex® electronic drug exposure monitor (eDEM™). Pharmacokinetic assessments of ciprofibrate exposure were performed as previously described in Chapter 4.

RNA extraction and handling

RNA was extracted from the frozen tissue samples as previously described [12]. Purity and quantity of the RNA were determined by spectrophotometric analysis at 260 and 280 nm. Integrity of the RNA was verified by agarose gel electrophoresis.

Microarrays

Human cDNAs were spotted onto Corning ultraGAPS slides using an Amersham Biosciences Lucidea spotter. Each clone was spotted in duplicate on the array, with the exception of control clones that were spotted 4 to 12 times. The custom-made microarray contains 7777 replicated spots, representing 5954 unique human genes/expressed sequence tags. The Cy3-labeled cDNA probe preparation, hybridization and subsequent washes of the arrays were performed as previously described [12]. All arrays were scanned in an Agilent scanner model G2505B (Agilent Technologies, Inc., Wilmington, DE). We performed quantification using Imagene software (BioDiscovery, El Segundo, CA).

Data normalization

VARIANCE-STABILIZING TRANSFORMATION (VST)

Each RNA sample was hybridized to two microarray slides run in parallel, yielding quadruplicate data points for each sample. For analysis,

the intensities of the quadruplicate data set in each sample were scaled to the 75th percentile (the 75th percentile value of each data set was set to 100), followed by log-linear hybrid transformation, a VST previously described by Rocke and Durbin [13]. The average of the quadruplicate transformed intensities of each spot was designated as the transformed expression intensity in the RNA sample.

NORMALIZATION AMONG SAMPLES WITHIN EACH TISSUE

A sorted nonlinear smooth spline normalization procedure [14] was applied on all samples from the same tissue. Concordance correlation coefficients were calculated for all samples from the same tissue as previously described [15].

Statistical Analysis

INCLUDED DATA

All subjects that dropped out were replaced with newly recruited subjects receiving the same treatment. Data of subjects dropping out after Visit 2 (baseline) were included in the statistical analysis even when a subject was replaced. Consequently, data of subjects dropping out during the run-in period were not used in the analysis.

This resulted in 229 blood RNA samples, 64 muscle RNA samples and 60 adipose tissue RNA samples that were analyzed by microarray. The hybridization of the blood RNA samples were carried out in five hybridization batches on chips from eight print lots; while the hybridization of the muscle and adipose tissue RNA samples was carried out in one hybridization batch from two and three print lots, respectively. Two samples from adipose tissue tissue were identified as outliers and were excluded from the statistical analysis. The total number of high quality tissue samples is presented in table 1.

MIXED MODEL ANALYSIS

Differentially expressed gene-lists were constructed using mixed model analysis as implemented in Partek Pro™ (Partek Inc. St. Louis, Missouri). The results obtained with Partek Pro™ were consistent with results using the SAS System (SAS Institute Inc. Cary, NC) in a small subset of genes tested. Specifically, the log-linear hybrid VST data were analyzed in a mixed model with both fixed factors, i.e. disease status, treatment regimen, time course, sex, and random factors, i.e. subject, print lot, hybridization batch. Restricted maximal likelihood method was used in variance estimation. As a first pass, a general model considering main effects of fixed factors and interaction among fixed factors were fitted for each of the 9514 clones in the microarray, the p-values for both main effects and interactions were used to rank the genes to identify the differentially expressed genes (α=0.05).

GENE NETWORK AND FUNCTIONAL ANALYSES

Both gene network and functional analyses were performed via the use of Ingenuity Pathways Analysis (IPA; Ingenuity® Systems, www.ingenuity.com).

The Ingenuity Pathways Knowledge Base (IPKB) collects millions of gene to gene interactions, both direct and indirect, as a result of manual curation from the literature. Given a list of significantly differentially expressed genes that were selected from a data set by a particular p-value cut-off, networks consisted of 35 genes, based on the interactions collected in the (IPKB), are selected to uncover the maximum number of input genes. A statistic score, equivalent to minus log p-value, is given for each selected network to reflect the probability of selecting such a network by random chance. This score is dependent both on the number of input genes and on the number of genes (named as focus genes) in the selected network.

For a given list of genes, either directly selected form an entire data set, or from focus genes in a selected network, the functional analysis identifies the biological functions and/or diseases that were most significant to the selected genes. Fischer's exact test is used to calculate a p-value determining the probability that each biological function and/or disease assigned to the selected genes is due to chance alone.

TABLE 1 **Summary of tissue samples with high quality gene expression data per Visit and treatment group**. Number of samples with high quality gene expression data used in the analysis. Missing data was modelled using Partek mixed model analysis. DF: Diabetic Female; HF: Healthy Female; DM: Diabetic Male; HM: Healthy Male; SK: Skeletal Muscle; AT: Adipose Tissue.

Treatment	Group	Pre-1day		Post-6wk	
		SM	AT	SM	AT
Placebo	DF	4	4	4	3
Placebo	NF	4	4	4	4
Placebo	DM	3	4	4	4
Placebo	NM	4	3	4	4
Ciprofibrate	DF	3	3	4	2
Ciprofibrate	NF	6	6	4	4
Ciprofibrate	DM	4	4	4	3
Ciprofibrate	NM	4	4	4	4
Sub-total		32	32	32	28

IDENTIFICATION OF 'ENRICHED' FUNCTIONS AND PATHWAYS

The IPA results from the baseline comparisons of the present and our previous study were compared to identify 'enriched' putative T2DM-associated biological functions and pathways.

Analogous to the present study, our previous microarray study included 18 T2DM patients (9 male and 9 female), mean age 56.8 ± 10.53 and 54.0 ± 8.56 years, and body mass index (BMI) of 28.8 ± 2.39 and 32.7 ± 4.61, respectively and 16 HVs (8 male and 8 female), mean age 22.1 ± 4.79 and 24.4 ± 7.25 years and BMI of 24.0 ± 4.07 and 24.6 ± 5.07, respectively. The patients were withdrawn from their regular oral antidiabetic medication and baseline comparisons of global tissue gene expression profiles between T2DM and HVs were performed after a two-week single-blind placebo run-in period. Virtually identical methods and statistical criteria for differential gene expression were used in the present and previous study.

A biological function or pathway was considered 'enriched' if it was significantly differentially expressed in the respective tissue datasets of both studies, and thus showed overlap. Enriched biological functions in the adipose tissue and skeletal muscle datasets were identified by comparing the significant biological functions within the top 2 gene networks (containing 70 focus genes) of each data set across both studies. Enriched canonical pathways in the adipose tissue dataset were identified by comparing the significant pathways within the top two gene networks from both studies. Enriched canonical pathways in skeletal muscle were identified by comparing the significant canonical pathways derived from the top 200 focus genes from each study.

Results

Subjects

Detailed demographics of the study populations were submitted for publication elsewhere. Briefly, the T2DM patients group consisted of 16 subjects (8 male and 8 female), mean age 59 ± 7.8 years and BMI of 30.4 ± 5.2, with a mean disease duration of 6 ± 4.2 years. The T2DM patients were allowed to continue their regular oral hypoglycaemic medication. The HVs group consisted of 16 subjects (8 male and 8 female), mean age 27 ± 7.8 years and BMI of 24.0 ± 2.6.

Two subjects (healthy female subjects in the ciprofibrate group) withdrew consent for personal reasons and were replaced with newly recruited volunteers receiving the same treatment.

Mean overall compliance was good in both patient (99.7% of prescribed dose) and HVs group (97% of prescribed dose) and both groups had equal exposure to ciprofibrate.

Effects on plasma markers

After 3 weeks of treatment with ciprofibrate *vs.* placebo, significant decreases in total cholesterol (TC), triglycerides (TG), apolipoprotein B (apo-B), and apo-lipoprotein CIII (apo-CIII) concentrations were observed in the T2DM group. We found no significant change in high density lipoprotein cholesterol (HDL-c), low density lipoprotein cholesterol (LDL-c), apo-lipoprotein CII (apo-CII), Lipoprotein (a) (Lp(a)), free fatty acids (FFA), glucose, monocyte chemoattractant protein 1 (MCP-1) and tumour necrosis factor alpha (TNFα) plasma concentrations in this group. In the HVs group, we observed a significant decrease in mean TC, HDL-c, FFA, apo-B and apo-CIII and no significant change in TG, LDL-c, apo-CII, Lp(a), FPG, MCP-1 and TNFα plasma concentrations. A more detailed description of the effects on the plasma markers measured in this study was submitted for publication elsewhere (see Chapter 4).

Gene expression profiling

When assessed by ANOVA mixed model, our analyses indicated that none of the genes in any of the three tissues, showed a significant treatment response at p-value cut-off of 0.05 after correction for multiple comparisons using the Bonferroni adjustment (table 2).

In contrast, analysis at baseline revealed several genes in adipose tissue, skeletal muscle and PBLs that were significantly differentially expressed between T2DM patients and HVs after Bonferroni adjustment and p-value cut-off 0.05 (table 2). Using these criteria, there was no significant disease*treatment interaction in any of the tissues, indicating that there were no significant effects when the treatment response was assessed in one of the groups separately (i.e. in the T2DM or HVs group alone). In addition, a number of genes, mainly in adipose tissue, were significant for the disease*gender interaction, indicating some gender specific changes in disease state (table 2).

Exploratory pathway analysis

To perform exploratory pathway analyses of differentially expressed genes using IPA and allow comparisons with our previous results,

we initially applied the same p-value cut-off of 0.001 as was used in our prior microarray study. At this cut-off, ciprofibrate treatment significantly affected the expression of 2 genes in skeletal muscle tissue, 0 genes in adipose tissue, and 3 genes in PBLs (table 2). Evidently, this small number of differentially expressed genes was insufficient for meaningful pathway analyses.

When comparing T2DM and HVs at this p-value cut-off at baseline, a large number of differentially expressed genes were identified in both adipose, skeletal muscle tissue and PBLs, allowing subsequent gene network and canonical pathway analyses (table 2).

Subsequently, we varied the selection criterion (p-value cut-off for the number of genes in the input list) to attain the highest score for the top ranking network(s) in the adipose tissue, skeletal muscle and PBLs dataset. A top score of 59, corresponding with 200 focus genes, was reached for two gene networks in the adipose tissue dataset. Similarly, a top score of 59 corresponding with 200 focus genes was reached for a single gene network in the skeletal muscle dataset (table 3). The analysis failed to identify significant gene networks in PBLs.

The top ranking network(s) and a selection of potentially important canonical pathways within each network for adipose tissue are presented in figure 2a (gene network 1) (p. 218) and figure 2b (gene network 2) (p. 219), and for skeletal muscle in figure 3 (p. 220). In addition, tables 4 and 5 provide an overview of the different functions covered by the genes within the two top ranking networks for each tissue.

Subsequently, a global pathway analysis using the top 200 focus genes of each tissue identified several putative disease related canonical pathways. The top 15 putative disease related canonical pathways for adipose tissue and skeletal muscle are presented in figure 4a (p. 221) and 4b (p. 221), respectively.

In adipose tissue, the functional analysis identified several genes implicated in lipid and carbohydrate metabolism as well as several previously reported T2DM-associated genes (table 6a). In addition, the two top rating gene networks and canonical pathways identified in the adipose tissue dataset harbour genes that are either components of the insulin signalling system (i.e. Insulin receptor substrate 1 (IRS1), cAMP-dependent protein kinase, catalytic subunit alpha (PRKACa), cAMP-dependent protein kinase, regulatory subunit beta 2 (PRKAR2B); figures 2a and 3; p. 218- 220), or have been characterized as negative modulators of molecular insulin action (p38 mitogen-activated protein kinase (P38MAPK); figure 4a; p. 221) [16]. In addition, the glycolysis/gluconeogenesis canonical pathway was identified as most significant differentially expressed

TABLE 2 **Significant genes in ANOVA mixed model analysis.** Number of genes in each tissue that were significant for each of the factors assessed in the ANOVA mixed model after adjustment for multiple comparisons (Bonferroni) and the number of genes significant when using a less stringent raw p-value cut-off of p=0.001 is indicated. Only genes differentially expressed between T2DM patients and HVs reached statistical significance after adjustment for multiple comparisons. PBLs: peripheral blood leukocytes.

	Skeletal muscle		**Adipose tissue**		**PBLs**	
Contrast or interaction term	Bonferroni; p=0.05	p=0.001	Bonferroni; p=0.05	p=0.001	Bonferroni; p=0.05	p=0.001
Treatment	0	2	0	0	0	3
Disease	29	584	268	1349	83	605
Disease * Treatment	0	1	0	2	0	6
Disease * Gender	1	81	23	425	1	16

TABLE 3 **Variation of selection criterion to attain the highest score for the top ranking gene network(s).** The number of input genes derived from the baseline comparison between the T2DM and HVs group at baseline, number of genes from this list that is covered in Ingenuity Pathway Analysis ('focus genes'), top score of gene network and the number of focus genes in each gene network are given for several p-value cut-offs in the adipose tissue and skeletal muscle dataset. A top score of 59, corresponding with 200 focus genes, was reached for two gene networks in the adipose tissue dataset (bold) and for one gene network in the skeletal muscle dataset (bold). The analysis failed to identify significant gene networks in the PBLs dataset (data not shown). NA: not available; IPA: Ingenuity Pathway Analysis.

Selection Criterion	Number of genes	Number of genes covered in IPA	Top score of gene network	Number of focus genes in gene network
Adipose tissue				
p<0.0001	693	411	47	35
p<0.0000107	**321**	**200**	**59**	**35**
p<0.00001	311	194	57	30
Muscle				
p<0.001	584	315	52	35
p<0.00039	**387**	**200**	**59**	**35**
p<0.0001	214	110	21	16

pathway in the adipose tissue dataset (figure 4a; p. 221). Moreover, the chemokine Interleukin 1 beta (IL1ß), which is a component of the PPAR signalling pathway and believed to be an important negative modulator of insulin signalling, was identified as one of the central nodes in gene network 1 (figure 2a; p. 218). Also, PPARγ, the target for thiazolidinedione class drugs, which activation is associated with improvements in insulin sensitivity [17], was identified in gene network 2 (figure 2b; p. 219).

TABLE 4 **Functional analysis of top 200 focus genes within the adipose tissue data set.** The functional analysis of the top 200 focus genes in the adipose tissue data set identified several focus genes that were significantly implicated in lipid and carbohydrate metabolism of which several were identified as T2DM-associated genes. Fischer's exact test was used to calculate a p-value determining the probability that each biological function and/or disease category within the input list is due to chance alone. If the association between the input list and biological function has a p-value less than 0.05 the function and/or disease category was considered significant.

Function	Gene symbol	p-value	number of genes
Lipid Metabolism			
quantity of lipid	EDN1, FCGR2B, GCK, GNAS, IL1B, LIF, LPL, PPARG, PTEN, TGFB1, UCP3, XBP1	6.0E-7	12
quantity of progesterone	EDN1, LHCGR, PTGFR, TIMP1	1.3E-5	4
quantity of fatty acid	GCK, LPL, PPARG, TGFB1, UCP3	1.4E-4	5
quantity of phosphatidylethanolamine	IL1B, XBP1	1.8E-4	2
accumulation of lipid	BCL2, IL1B, LPL, PPARG, PTEN, TIMP1	5.1E-5	6
production of lipid	EDN1, GCK, IL1B, PLAA, SYK, TGFB1	1.5E-4	6
production of eicosanoid	EDN1, IL1B, PLAA, SYK, TGFB1	1.7E-4	5
Carbohydrate Metabolism			
uptake of 2-deoxyglucose	AGTR1, EDN1, GCK, GNAS, IL1B, IRS1, PPARG, PTEN, UCP3	2.0E-11	9
uptake of monosaccharide	AGTR1, EDN1, GCK, GNAS, IL1B, IRS1, LPL, PPARG, PTEN, UCP3	2.8E-9	10
uptake of d-glucose	IL1B, IRS1, LPL, PPARG, UCP3	1.9E-4	5
production of carbohydrate	AGTR1, EDN1, GCK, HLA-A, IL1B, LPL, PPARG	1.2E-6	7
production of monosaccharide	AGTR1, EDN1, GCK, HLA-A, LPL, PPARG	1.0E-5	6
oxidation of d-glucose	GCK, IL1B, LPL, UCP3	6.0E-6	4
quantity of carbohydrate	EDN1, GCK, IL1B, IRS1, LPL, PPARG, TGFB1, UCP3	1.2E-5	8
quantity of monosaccharide	EDN1, GCK, IL1B, IRS1, LPL, PPARG, UCP3	1.8E-5	7
quantity of d-glucose	EDN1, IL1B, IRS1, LPL, PPARG, UCP3	2.7E-5	6
modification of carbohydrate	GCK, IL1b, LPL, TGFB1, UCP3	2.0E-5	5
transport of d-glucose	EDN1, GCK, IL1B, IRS1, PTEN	4.6E-5	5
synthesis of polysaccharide	GCK, IL1B, IRS1, TGFB1, TGFB2	1.4E-4	5
glycolysis	BCL2, GCK, UCP3	2.4E-4	3
Metabolic Disease			
diabetes associated	AGTR1, B2M, GCK, IRS1, LPL, PPARG, PTEN, TGFB1, UCP3	2.0E-7	9

TABLE 5 **Functional analysis of top 200 focus genes within the skeletal muscle data set.** The functional analysis of the top 200 focus genes in the skeletal muscle data set identified several focus genes that were significantly implicated in lipid and carbohydrate metabolism. Fischer's exact test was used to calculate a p-value determining the probability that each biological function and/or disease category within the input list is due to chance alone. If the association between the input list and biological function has a p-value less than 0.05 the function and/or disease category was considered significant.

Function	Gene symbol	p-value	number of genes
Lipid Metabolism			
beta-oxidation of palmitic acid	CYP19A1, ESRRA	9.0E-5	2
secretion of testosterone	CTSG, TNF	2.2E-4	2
release of eicosanoid	CCL11, PLCG1, PRKCA, TNF	3.3E-4	4
release of arachidonic acid	PLCG1, PRKCA, TNF	1.6E-3	3
production of diacylglycerol	PLCGI, TNF	6.6E-4	2

production of testosterone	AMH, TNF	1.5E-4	2
hydrolysis of phosphatidylinositol 4,5-diphosphate	PLCG1, PRKACA	9.7E-4	2
hydrolysis of phosphatic acid	PLCG1, PRKACA, TNF	2.4E-3	3
synthesis of acylglycerol	FASN, PLAU	1.1E-3	2
formation of eicosanoid	CCL11, TNF	2.0E-3	2
Carbohydrate metabolism			
uptake of d-glucose	PLAU, PRKCA, PTPN1, TNF	1.4E-4	4
uptake of 2-deoxyglucose	PRKCA, PTPN1, TNF	5.4E-4	3
synthesis of carbohydrate	PLAU, PRKCA, PTPN1, TNF	3.2E-3	3

TABLE 6A **Enriched biological functions.** Comparison of the IPA results from the present and previous study (rosiglitazone study) identified the functions 'accumulation and production of lipid' as well as 'uptake of monosaccharide and D-glucose' as significantly enriched in the adipose tissue data. The 'release of eicosanoid' function was significantly enriched across the skeletal muscle data sets of both studies. The direction of change (up or down regulation) for overlapping individual genes (in bold) is indicated with an arrow.

Function	Gene symbol (Rosiglitazone study)	p-value	Gene symbol (Ciprofibrate study)	p-value
Adipose tissue				
accumulation of lipid	CCL2, DBI, IL1RN, **LPL(↓)**, PPARD, **PPARG(↓)**, TNF	4.2E-6	BCL2, IL1B, **LPL(↑)**, **PPARG(↑)**, PTEN, TIMP1	5.1E-5
production of lipid	CCL2, F2, IL1RN, P2RX7, **PLAA(↓)**, RALA, TNF	1.5E-5	EDN1, GCK, IL1B, **PLAA(↑)**, SYK, TGFB1	1.5E-4
uptake of monosaccharide	CSF1, CTF1, IRS1(↓), LPL(↓), PPARG(↓), PRKCB1, SLC2A4, TNF	6.9E-7	AGTR1, EDN1, GCK, GNAS, IL1B, **IRS1(↑)**, **LPL(↑)**, **PPARG(↑)**, PTEN, UCP3	2.8E-9
uptake of D-glucose	CSF1, CTF1, **IRS1(↓)**, **LPL(↓)**, **PPARG(↓)**, SLC2A4, TNF	9.1E-7	IL1B, **IRS1(↑)**, **LPL(↑)**, **PPARG(↑)**, UCP3	1.9E-4
diabetes associated	HLA-DRA, ICAM1, **IRS1(↓)**, **LPL(↓)**, **PPARG(↓)**, SLC2A4, SOD1, TNF	2.4E-6	AGTR1, B2M, GCK, **IRS1(↑)**, **LPL(↑)**, **PPARG(↑)**, PTEN, TGFB1, UCP3	2.0E-7
Skeletal muscle				
release of eicosanoid	ALOX5AP, CLEC11A	0.012	CCL11, PLCG1, PRKCA, TNF	3.3E-4

TABLE 6B **Enriched biological pathways.** Comparison of the IPA results from the present (ciprofibrate study) and previous study (rosiglitazone study) identified the 'PPAR' and 'IL-6' canonical pathways as most significantly enriched in the adipose tissue and skeletal muscle data sets, respectively. The direction of change (up or down regulation) for overlapping individual genes (in bold) is indicated with an arrow.

Pathway	Gene symbol (Rosiglitazone study)	p-value	Gene symbol (Ciprofibrate study)	p-value
Adipose tissue				
PPAR signalling	PDGFRB, PPARD, **PPARG(↓)**, TNF	5.1E-3	FOS, IL1B, NFKB2, NFKBIB, PPARG(↑)	1.4E-3
B Cell receptor signalling	BTK, MAP2K3, PRKCB1, RAC1, RPS6KB1	7.4E-3	FCGR2B, NFKB2, NFKBIB, PTEN, SYK	1.6E-2
PDGF signalling	PDGFRB, PRKCB1, SPHK1	0.021	FOS, PDGFC, SRF	0.034
p38 MAPK signalling	MAP2K3, **TGFB2(↓)**, TNF	0.029	IL1B, MAPK11, SRF, TGFB1, **TGFB2(↑)**	1.1E-3
IGF-1 signalling	IGFBP6, **IRS1(↓)**, RPS6KB1	0.036	FOS, **IRS1(↑)**, PRKACA, PRKAR2B, SRF	1.6E-3
Skeletal muscle				
IL-6 signalling	CD14, **CSNK2A1(↓)**, SRF, STAT3, TNFRSF1B	8.7E-4	**CSNK2A1(↓)**, CYP19A1, IL6ST	1.6E-3
PDGF signalling	**CSNK2A1(↓)**, SRF, SPHK2	3.0E-3	**CSNK2A1(↓)**, PLCG1	1.2E-2
EGF signalling	**CSNK2A1(↓)**, SRF, STAT3	8.7E-3	**CSNK2A1(↓)**, PLCG1, PRKCA, RAF1, SOS1	2.6E-3

In skeletal muscle, several biological functions, and canonical pathways implicated in PPAR and insulin signalling as well as negative modulation of insulin signalling were identified, including the Interleukin-6 (IL6), Chemokine, and Extracellular Signal-regulated Kinase Mitogen Activated Protein Kinase (ERK-MAPK) pathway (table 6b and figure 4b; p. 221). In addition, the chemokine tumour necrosis factor alpha (TNFα), which is recognized as an important negative modulator of molecular insulin action [16], was identified as central node in gene network 1 (figure 3; p. 220).

Enriched biological functions and pathways

Comparison of the IPA results from the present and previous study identified several biological functions and canonical pathways that were significantly enriched in both adipose tissue and skeletal muscle data sets (tables 6a and 6b). More specifically, 'accumulation and production of lipid' as well as 'uptake of monosaccharide and D-glucose' functions, were significantly enriched in the adipose tissue data sets whereas only the 'release of eicosanoid function' was enriched in the skeletal muscle data sets. In addition, the PPAR and IL-6 canonical pathways were identified as most significantly enriched in the adipose tissue and skeletal muscle data sets, respectively (table 6b).

Six individual genes within these pathways and functional categories were significantly enriched (tables 6a and 6b; overlapping genes are indicated in bold). Of these genes only one (casein kinase 2, alpha 1 polypeptide (CSNK2A1) had the same direction of differential expression (up- or down regulation) in both studies (table 6b).

Discussion

In this study we failed to identify genes in adipose tissue, skeletal muscle tissue and PBLs of T2DM patients and HVs that were significantly altered in expression after 3 weeks treatment with ciprofibrate *vs.* placebo (table 2). We did identify several genes that were significant differentially expressed between the T2DM patients and HVs at baseline (table 2). In addition, subsequent exploratory IPA identified several putative disease related gene networks (figures 2a; p. 218, 2b; p. 219 and 3; p. 220), biological functions and canonical pathways in adipose tissue and skeletal muscle (tables 4 and 5). Comparison of the IPA results from the present and our previous microarray study identified several biological functions and pathways that were significantly 'enriched' across the adipose tissue and skeletal muscle data sets (tables 6a and 6b). However, only one of the six individually 'enriched' genes exhibited the same direction of differential expression (table 6b).

The main objective of this study was to identify changes in gene expression profiles (i.e. molecular biomarker patterns) in human tissues that could serve as 'mechanism-based' efficacy parameters for the action of PPARα-agonist class drugs *in vivo*. To this end, we conducted an intensive clinical study that investigated the pharmacological effects of ciprofibrate (as prototype PPARα-agonist class) *vs.* placebo treatment on global gene expression profiles in relevant, easily accessible target tissues (i.e. skeletal muscle and adipose tissue) and PBLs of T2DM patients and HVs. In parallel, we investigated (and demonstrated) the efficacy of ciprofibrate using conventional lipid parameters and markers of inflammation. We found that spotted cDNA microarray technology was unable to detect changes in skeletal muscle, adipose tissue and PBLs gene expression profiles that reflect the pharmacological effects of ciprofibrate at the transcriptional level.

This study was a follow-up on a previous investigation, which evaluated the effects of rosiglitazone (PPARγ agonist) on global gene expression profiles in skeletal muscle, adipose tissue and PBLs of T2DM patients and HVs. In that study we failed to identify significant changes in gene expression profiles after 6 weeks of treatment with (in that case) rosiglitazone *vs.* placebo. Potential factors influencing the interpretation of the results were related to print batch number and relatively high drop-out rate of the subjects, mainly due to hyperglycaemia in the placebo group following the withdrawal of oral hypoglycaemic medication. These issues in combination with the many variables that need to be taken into account in clinical studies, made 'gold standard' permutation based analyses like significance analysis of microarrays (SAM) [18], which are capable of generating or working with a (pre-set) false discovery rate (FDR), unsuitable. In the present study, all patients remained on stable doses of oral hypoglycaemic medication. In addition, the present study was less hampered by missing data due to low quality tissue samples (RNA degradation or RNA loss from handling).

The results from the present study are similar to the findings from our previous investigation with the PPARγ agonist rosiglitazone and indicate that, even in this adapted study design, the signal-to-noise ratio is probably too low to adequately identify pharmacological effects of PPAR agonists on global gene expression levels using spotted cDNA

microarrays. As pointed out before, this could be due to the fact that many small and subtle changes in the expression of PPAR modulated genes give rise to the observed (more robust) downstream effects observed on several biochemical markers in plasma. Furthermore, it appears that the natural (intra- and inter-individual) variation of gene expression in time, and the less tightly controlled experimental conditions in human studies, may well be important factors that add noise to the data and obscure the biological signal.

The notion that adipose tissue is not considered to be a major target for PPARα agonists may explain why no (direct) effects on gene expression profiles were observed in this particular tissue. Although PPARα has been shown to be expressed in adipose tissue using real-time RT-PCR, the mRNA levels in this specific tissue were much lower than in liver tissue [19]. However, this notion does not appear to hold true for skeletal muscle, since PPARα is reported to be substantially expressed in this tissue type [20;21]. In addition, we confirmed the expression of PPARα using real-time RT-PCR in all three tissues in a previous experiment (data not shown).

Alternatively, our assumption with regard to the working mechanism of PPAR agonists *in vivo* may be incomplete. Theoretically, pivotal 'downstream' effects could be attributable to other mechanisms in addition to the effects on transcription, i.e. post-transcriptional effects. In fact, data from a recent publication by Rasouli et al. support this notion by demonstrating that in subjects with impaired glucose tolerance the increased plasma levels of adiponectin in response to pioglitazone (a PPARγ agonist) were not associated with increased adiponectin gene expression in adipose tissue [22]. Their data suggest that pioglitazone increases plasma adiponectin levels by PPARγ-independent post-transcriptional regulation in contrast to the previously assumed transcriptional regulation.

Comparison of the T2DM and HVs groups at baseline identified several potentially disease related genes that were significantly differentially expressed in muscle tissue, adipose tissue and PBLS after adjustment for multiple comparisons (table 2). In addition, subsequent exploratory IPA analyses identified several biological functions and canonical pathways that have previously been implicated in the pathophysiology of T2DM. However, the overlap of significant individual genes within the Bonferroni corrected differentially expressed genes list of the present and previous study (data not shown) as well as the overlap of genes within the top 200 focus genes list in the exploratory analysis (table 6a and 6b) was very small. Moreover,

the direction of change (up- or down regulation) for all but one of the overlapping 'enriched' genes identified in the exploratory IPA analysis, exhibited differential expression in opposite directions (table 6a and 6b). Theoretically, this discrepancy could be related to differences in baseline conditions of the subjects between the two studies. In fact, in our previous (rosiglitazone) study the patients were taken off their regular oral hypoglycaemic medication while the patients in the current study remained on stable doses of hypoglycaemic medication. Therefore, metabolic derangement in the previous study could have resulted in different (noisier) gene profiles compared to the ones obtained from the (relatively) stable T2DM patients in the present study. In addition, the oral hypoglycaemic treatment itself could have resulted in changes in gene expression profiles. On the other hand, using a similar study design, identical biopsy methods, (CDNA) microarray platform and statistical workup, one would have expected to find considerable overlap in putative disease related genes between our two studies. These observations question the ability of CDNA microarray technology to provide reproducible results at the individual gene level and thus to provide valid findings. Nonetheless, as demonstrated by the pathway enrichment analysis comparing data from the present and our previous microarray study, one appears to gain more differentiating power for disease exploration by comparing two independent cohorts at the functional and pathway level instead of looking at individual genes. In addition, the reason for lack of overlap between the top differentially expressed genes of our two independent studies with similar populations may not only be inherent to the sensitivity and specificity of the (CDNA) microarray platform, but may also relate to the vast heterogeneity of metabolic diseases like T2DM. Hence, pathway-level analyses probably represent the best possible option to capture the essence of the coordinating disease-related changes.

In this light, a promising new approach described by Mootha et al, in which subsets of predefined genes in a common metabolic pathway (oxidative phosphorylation) in skeletal muscle were compared between T2DM patients and HVs, yielded encouraging results [23]. In their approach the hypothesis generating potential of microarray technology is traded for more statistical robustness, which may be the preferable approach if the goal is to attain reproducible and thus valid results from microarray experiments in the setting of metabolic diseases. However, also for this new approach, future experiments have to point out if these findings can be validated in an independent cohort.

In conclusion, as previously shown for the PPARγ agonist rosiglitazone, the results from the present study indicate that gene expression profiling using spotted CDNA microarrays coupled with sophisticated bioinformatic analyses is unable to capture the broad molecular response to drug intervention with a PPARα agonist *in vivo*. Consequently, we conclude that in the tested study designs and wet lab design settings of our studies, this technology is unable to provide a broad array of 'mechanism-based' biomarkers for the pharmacological effects of PPAR agonists in humans. However, although the technology could not detect consistent disease related changes at the individual gene level, it may still have value for disease exploration by making use of biological function and pathway-level analyses.

REFERENCE LIST

1 Lockhart DJ, Dong H, Byrne MC *et al.* Expression monitoring by hybridization to high-density oligonucleotide arrays. *Nat. Biotechnol.* 1996; **14**:1675-1680

2 't Veer LJ, Dai H, van de Vijver MJ *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; **415**:530-536

3 DeRisi J, Penland L, Brown PO *et al.* Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat.Genet.* 1996; **14**:457-460

4 Golub TR, Slonim DK, Tamayo P *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* 1999; **286**:531-537

5 Valk PJ, Verhaak RG, Beijen MA *et al.* Prognostically useful gene-expression profiles in acute myeloid leukemia. *N.Engl.J Med* 2004; **350**:1617-1628

6 Gunther EC, Stone DJ, Gerwien RW, Bento P, Heyes MP. Prediction of clinical drug efficacy by classification of drug-induced genomic expression profiles in vitro. *Proc.Natl.Acad.Sci. U.S.A* 2003; **100**:9608-9613

7 Gunther EC, Stone DJ, Rothberg JM, Gerwien RW. A quantitative genomic expression analysis platform for multiplexed in vitro prediction of drug action. *Pharmacogenomics.J* 2005; **5**:126-134

8 Natsoulis G, El Ghaoui L, Lanckriet GR *et al.* Classification of a large microarray data set: algorithm comparison and analysis of drug signatures. *Genome Res* 2005; **15**:724-736

9 DePrimo SE, Wong LM, Khatry DB *et al.* Expression profiling of blood samples from an SU5416 Phase III metastatic colorectal cancer clinical trial: a novel strategy for biomarker identification. *BMC.Cancer* 2003; **3**:3

10 Palotas A, Puskas LG, Kitajka K *et al.* Altered response to mirtazapine on gene expression profile of lymphocytes from Alzheimer's patients. *Eur.J Pharmacol* 2004; **497**:247-254

11 Palotas A, Puskas LG, Kitajka K *et al.* The effect of citalopram on gene expression profile of Alzheimer lymphocytes. *Neurochem.Res* 2004; **29**:1563-1570

12 van Doorn M, Kemme M, Ouwens M *et al.* Biopsy and RNA extraction procedures of muscle and adipose tissue for microarray gene-expression profiling.In: *Biochips as pathways to drug discovery* eds Carmen E, Hardiman G, CRC, 2006; **7**:109-122

13 Rocke DM, Durbin B. Approximate variance-stabilizing transformations for gene-expression microarray data. *Bioinformatics.* 2003; **19**:966-972

14 Shaw KJ, Miller N, Liu X *et al.* Comparison of the changes in global gene expression of Escherichia coli induced by four bactericidal agents. *J Mol Microbiol.Biotechnol.* 2003; **5**:105-122

15 Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics* 1989; **45**:255-268

16 Perseghin G, Petersen K, Shulman GI. Cellular mechanism of insulin resistance: potential links with inflammation. *Int J Obes.Relat Metab Disord.* 2003; **27 Suppl 3**:S6-11

17 Lehrke M, Lazar MA. The many faces of PPARgamma. *Cell* 2005; **123**:993-999

18 Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc.Natl.Acad.Sci.U.S.A* 2001; **98**:5116-5121

19 Tsuchida A, Yamauchi T, Takekawa S *et al.* Peroxisome Proliferator-Activated Receptor (PPAR){alpha} Activation Increases Adiponectin Receptors and Reduces Obesity-Related Inflammation in Adipose Tissue: Comparison of Activation of PPAR{alpha}, PPAR{gamma}, and Their Combination. *Diabetes* 2005; **54**:3358-3370

20 Neve BP, Fruchart JC, Staels B. Role of the peroxisome proliferator-activated receptors (PPAR) in atherosclerosis. *Biochem Pharmacol* 2000; **60**:1245-1250

21 Lefebvre P, Chinetti G, Fruchart JC, Staels B. Sorting out the roles of PPAR alpha in energy metabolism and vascular homeostasis. *J Clin Invest* 2006; **116**:571-580

22 Rasouli N, Yao-Borengasser A, Miles LM, Elbein SC, Kern PA. Increased Plasma Adiponectin in Response to Pioglitazone Does Not Result from Increased Gene Expression. *Am J Physiol Endocrinol.Metab* 2005

23 Mootha VK, Lindgren CM, Eriksson KF *et al.* PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat.Genet.* 2003; **34**:267-273.