

# A meta-analysis study of gene expression datasets in mouse liver under PPAR $\alpha$ knockout

KAN HE<sup>1,2</sup>, ZHEN WANG<sup>1,2</sup>, QISHAN WANG<sup>1,2</sup> AND YUCHUN PAN<sup>1,2\*</sup>

<sup>1</sup> School of Agriculture and Biology, Department of Animal Sciences, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China

<sup>2</sup> Shanghai Key Laboratory of Veterinary Biotechnology, Shanghai 200240, People's Republic of China

(Received 22 April 2013; revised 28 May 2013)

## Summary

Gene expression profiling of peroxisome-proliferator-activated receptor  $\alpha$  (PPAR $\alpha$ ) has been used in several studies, but there were no consistent results on gene expression patterns involved in PPAR $\alpha$  activation in genome-wide due to different sample sizes or platforms. Here, we employed two published microarray datasets both PPAR $\alpha$  dependent in mouse liver and applied meta-analysis on them to increase the power of the identification of differentially expressed genes and significantly enriched pathways. As a result, we have improved the concordance in identifying many biological mechanisms involved in PPAR $\alpha$  activation. We suggest that our analysis not only leads to more identified genes by combining datasets from different resources together, but also provides some novel hepatic tissue-specific marker genes related to PPAR $\alpha$  according to our re-analysis.

## 1. Introduction

The transcription factor peroxisome-proliferator-activated receptor  $\alpha$  (PPAR $\alpha$ ) is one of the nuclear-hormone receptors (NRs), which need to be ligand-activated (Issemann & Green, 1990; Bookout *et al.*, 2006; Michalik *et al.*, 2006). In mammals, PPAR $\alpha$  is thought to participate in several biological processes, especially in the nutrient metabolism and inflammation. The expression levels of PPAR $\alpha$  are high in most of tissues with active fatty acid catabolism, such as liver, heart, kidney, brown adipose tissue, muscle, small and large intestines (Issemann & Green, 1990; Braissant *et al.*, 1996; Devchand *et al.*, 1996). Particularly, the liver is a central player in the whole body energy homeostasis by its ability to orchestrate fatty acid and glucose metabolism, and it plays an important role in fatty acid oxidation and this catabolic energy burning is regulated by PPAR $\alpha$  (Reddy & Hashimoto, 2001). Nowadays, the PPAR $\alpha$ -null mouse model is frequently used according to the

study of the critical role for PPAR $\alpha$  in the cellular fasting response (Leone *et al.*, 1999), since we could get the significantly PPAR $\alpha$ -dependent genes by the comparison between wild-type (WT) and knockout (KO) mice. These identified genes may be associated with metabolism or diseases, such as obesity and diabetes. During the past 15 years, numerous studies with PPAR $\alpha$  null mice have demonstrated the critical roles played by this receptor in energy metabolism, hepatic steatosis, inflammation, cardiac pathophysiology, cell-cycle alterations and hepatocarcinogenesis (Lee *et al.*, 2002; Lefebvre *et al.*, 2006). Particularly in recent years, microarray technology has been highly used to map PPAR $\alpha$ -dependent genes and further characterizes PPAR $\alpha$  function in different tissues in genome wide, especially for liver. We could establish gene expression patterns that would facilitate understanding of almost every aspect of cellular and molecular biology related to PPAR $\alpha$ . Since mass databases have been created using a variety of gene expression data by different researchers, many studies have considered gene expression determination by PPAR $\alpha$  activation at the transcriptome level. In these liver studies, PPAR $\alpha$  was shown to be critical for the coordinated transcriptional activation of

\* Corresponding author: School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China. Tel: +86-21-34205836. Fax: +86-21-34206394. E-mail: panyuchun1963@yahoo.com.cn

Table 1. The details of individual study and meta-analysis of the chosen two datasets

| Study   | Time point | Chip         | Sample size | Rep | Down genes | Up genes | Down pathways | Up pathways |
|---------|------------|--------------|-------------|-----|------------|----------|---------------|-------------|
| GSE9786 | 7 d        | Affy 430 2.0 | 16          | 4   | 289        | 198      | 13            | 9           |
| GSE8295 | 5 d        | Affy 430 2.0 | 16          | 4   | 1143       | 624      | 14            | 9           |
| RP      |            |              |             |     | 1695       | 1708     | 13            | 6           |

GSE9786 and GSE8295 were both related to PPAR $\alpha$  in mouse liver using the platform of Affymetrix Mouse 430 2.0 GeneChips, the sample sizes used were both 16 and the numbers of replicates were both 4. However, GSE9786 was contributed by Rosen on 4 February 2008 for 7 days treatment and GSE8295 was contributed by Hooiveld on 24 July 2007 for 5 days treatment. According to individual analysis on the above two with comparison of WT to KO, there were 289 and 1143 significantly down-regulated and 198 and 624 up-regulated genes, respectively, in the study of GSE9786 and GSE8295. The numbers of significantly enriched down- or up-regulated pathways were 14 and 9 for GSE8295 and 13 and 9 for GSE9786. Based on meta-analysis (RP) of the chosen two datasets, there appeared to 1695 down-regulated and 1708 up-regulated genes as well as 13 and 6 enriched down- and up-regulated pathways.

genes involved in lipid catabolism, including cellular fatty acid uptake and activation, mitochondrial  $\beta$ -oxidation, peroxisomal fatty acid oxidation, ketone body synthesis, fatty acid elongation and desaturation and apolipoprotein synthesis (Leone *et al.*, 1999; Mandard *et al.*, 2004). However, the reproducibility of datasets and the credibility of results are both questionable, especially for the fact that the transcription patterns of genes revealed by different databases frequently contradict each other. Therefore, combining the gene-expression datasets generated at different laboratories or conditions in mouse hepatic tissue is needed to identify tissue-specific pathways or genes involved in PPAR $\alpha$  activation. It may be helpful for us to get more powerful and comprehensive knowledge on the biological functions of PPAR $\alpha$  in mice liver.

A non-parametric method named rank product (RP) is one of the most useful statistical models to identify the significant genes in the analysis of microarray experiments, which is based on the calculation of RP from replicate datasets (Breitling *et al.*, 2004; Nemhauser *et al.*, 2004). The RP analysis can also be applied to datasets from multiple origins. Owing to the lack of experimental standards for microarray experiments that leads to heterogeneous datasets, direct comparison is not possible. Instead of using actual expression data, this approach can combine the gene rank from different origins together to select differentially expressed genes. The smaller the RP value, the more important the observed gene is. The significance of the list of up- or down-regulated genes is based on the estimated percentage of false positive (pfp) predictions, which is also known as false discovery rate (FDR). It is true that the meta-analysis of related microarray datasets based on this approach would be powerful to identify regulatory factors required for response to PPAR $\alpha$  KO.

## 2. Materials and methods

### (i) Microarray datasets collection and preprocessing

We searched Gene Expression Omnibus (GEO, www.ncbi.nlm.nih.gov/geo/) for the gene expression profiling studies related to PPAR $\alpha$  KO. Data were included in our re-analysis if they met the following conditions: (1) the data are genome-wide; (2) comparison was conducted between PPAR $\alpha$  KO and WT; (3) complete microarray raw or normalized data are available. Finally, we chose the datasets GSE9786 and GSE8295 for our re-analysis studies, which were respectively contributed by Rosen and Hooiveld in different periods, both for mouse liver (Rakhshandehroo *et al.*, 2007; Rosen *et al.*, 2008). Both laboratories measured gene expression under two classes (WT and KO) with similar condition. The details of them were shown in Table 1.

In GSE9786, 129S1/SvImJ WT and PPAR $\alpha$ -null mice were exposed to 3 mg/kg/day PFOA or water for 7 days. Total RNA was isolated from liver samples and gene expression analysed using Affymetrix Mouse 430 2.0 GeneChips. Data from 16 samples, with four mice in each of the four treatment groups, were analysed. In GSE8295, 3–5-month-old male pure bred WT (129S1/SvImJ) and PPAR $\alpha$ -null (129S4/SvJae) mice were used. WT and PPAR $\alpha$ -null mice were treated with the synthetic PPAR $\alpha$  ligand Wy14643 (0.1% w/w) mixed in the food or normal food (control) for 5 days ( $n=4$  per group). Liver total RNA from biological replicates was hybridized onto Affymetrix mouse genome 430 2.0 GeneChip arrays.

For the assessment of the influence of preprocessing on the comparison, the data preprocessing was performed using software packages developed in version 2.6.0 of Bioconductor and R version 2.10 (Gentleman *et al.*, 2004). Each Affymetrix dataset was background adjusted, normalized and log 2 probe-set intensities

calculated using the Robust Multichip Averaging (RMA) algorithm in Affy package (Gautier *et al.*, 2004).

#### (ii) Individual study of each dataset

In our individual study of each dataset, statistical analysis was performed by two-way ANOVA with a Benjamini and Hochberg FDR (BH-FDR=0.05) for multiple testing correction followed by Tukey's *post-hoc* tests (Park *et al.*, 2003). Differentially expressed genes between two neighbour stages were identified by 2-fold changes. Clustering on groups and genes was performed based on the identified genes' expression using the method of Hierarchical clustering. All of these processes were performed using software packages developed in version 2.6.0 of Bioconductor and R version 2.10.0.

#### (iii) Meta-analysis on two datasets

In order to identify regulatory elements required for response to PPAR $\alpha$  KO, we divided the samples in each dataset into two classes. Concretely, for GSE9786, eight samples from GSM247116 to GSM247123 were set to be the first class (WT) and the other eight samples from GSM247124 to GSM247131 were set to be the second class (KO); for GSE8295, eight samples from GSM205784 to GSM205788 and GSM205804 to GSM205816 were set to be the first class (WT) and the other eight samples from GSM205789 to GSM205803 and GSM205817 to GSM205820 were set to be the second class (KO). Our meta-analysis on the chosen two datasets was performed by the RankProd package to identify up- or down-regulated genes under WT against the treatment of PPAR $\alpha$  KO. The cut-off of pfp value was chosen as 0.05.

#### (iv) Pathway analysis

Further classifications of pathway analysis were performed by using the Database for Annotation, Visualization and Integrated Discovery (DAVID) (Dennis *et al.*, 2003; Huang *et al.*, 2009) following the above significance analysis, revealing overrepresented pathways of identified genes associated with PPAR $\alpha$  KO. The pathway with  $P < 0.01$  was identified as the significantly associated one.

### 3. Results and discussion

#### (i) Individual significance analysis results

Through our individual analysis on the above two datasets by two-ANOVA-FDR test with FDR  $\leq 5\%$  and Fold Change (FC)  $\geq 2$  with comparison of WT

to KO, in total 1767 genes were significantly identified in GSE8295 including 1143 down-regulated and 624 up-regulated. In contrast, the total number in GSE9786 was relatively lower as 487 including 289 down-regulated and 198 up-regulated (Table 1). The details of significantly expressed genes were shown in Additional Files 1 and 3 for GSE8295 and GSE9786, respectively. To explore the gene expression patterns of these two datasets, we performed hierarchical clustering on the identified genes across all samples and made a comparison between GSE8295 and GSE9786. As shown in Fig. 1, there were remarkable differences not only in the number of significant genes, but also in the expression patterns. We next sought to determine the pathways associated with the regulation of PPAR $\alpha$  KO in individual dataset, analysis of overrepresented pathways were performed by using DAVID with  $P$ -value cutoff as 0.01. The numbers of significantly enriched down- or up-regulated pathways were 14 and 9 for GSE8295 and 13 and 9 for GSE9786 (Table 1). The full lists of significantly related pathways were shown in Additional Files 2 and 4, respectively, for GSE8295 and GSE9786.

In summary, based on our individual study on the chosen two datasets, there were differences between GSE8295 and GSE9786 in many aspects, such as the number of PPAR $\alpha$ -dependent genes and the significance of related pathways. These differences were possibly due to the quality and quantity of samples chosen in each microarray experiment and the treatment condition may also play essential roles in affecting the accuracy of results. Therefore, a gene expression meta-analysis on these two datasets is necessary to get more reliable results.

#### (ii) Meta-analysis results and comparison with individual study

Through our gene expression meta-analysis of the chosen two datasets based on the RP method, there appeared to be totally 3403 significantly identified genes including 1695 down-regulated genes and 1708 up-regulated genes with the pfp value  $< 0.05$ . The details of all of them were shown in Additional file 5 and the top 10 ranked down- or up-regulated genes were showing in Table 2. On comparison with the results of individual studies, it was noted that the test of meta-analysis lead to more identified genes by combining datasets from two different origins together. In addition, according to the pfp values and  $P$  values of identified genes by our meta-analysis, the significances of down- or up-regulated genes were improved and the results become more powerful.

Among the top 10 ranked up-regulated genes, SULT1E1, one member of the sulphotransferase family whose function are to catalyse the sulphate

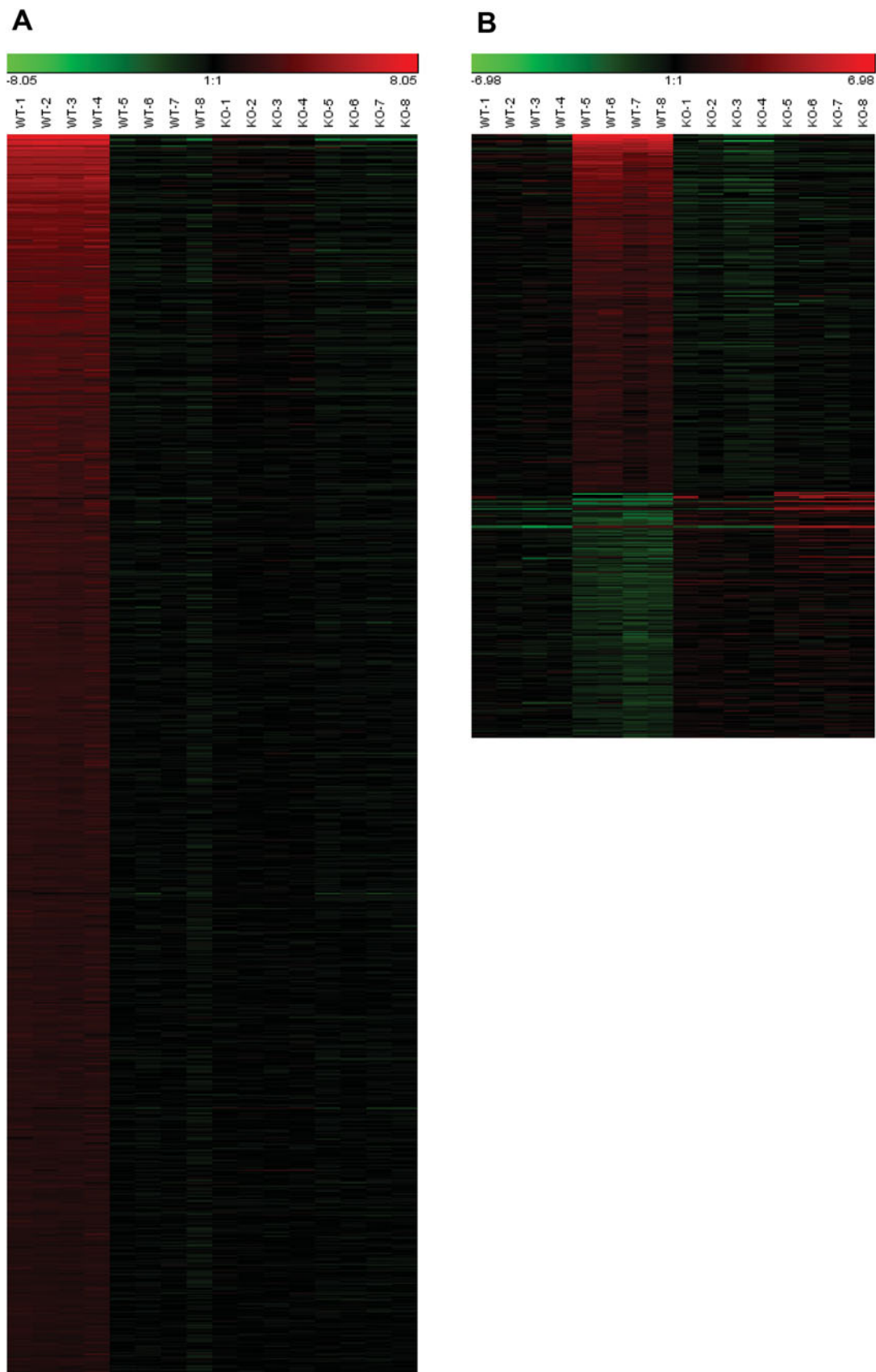


Fig. 1. Heatmaps of the selected two datasets. The heatmaps were showing hierarchical clusterings on the filtered entities using all samples included in GSE8295 (Figure 1A) and GSE9786 (Figure 1B). X-axis represents all samples and their groups included in each dataset; Y-axis represents all the filtered probe sets. Low gene expression was shown in green and high gene expression was in red.

Table 2. The top 10 ranked significant genes in meta-analysis

| ID           | Gene symbol  | RP      | FC:(class1/class2) | pfp | P-value |
|--------------|--------------|---------|--------------------|-----|---------|
| Up genes     |              |         |                    |     |         |
| 1420447_at   | Sult1e1      | 8.7008  | 11.7512            | 0   | 0       |
| 1416318_at   | Serp1b1a     | 17.3453 | 7.7648             | 0   | 0       |
| 1452905_at   | Meg3         | 35.6363 | 6.1294             | 0   | 0       |
| 1452183_a_at | Meg3         | 38.9036 | 5.8094             | 0   | 0       |
| 1434540_a_at | Clta         | 39.6126 | 4.6607             | 0   | 0       |
| 1436713_s_at | Meg3         | 40.3428 | 5.3645             | 0   | 0       |
| 1432517_a_at | Nnmt         | 40.7687 | 5.9662             | 0   | 0       |
| 1439380_x_at | Meg3         | 43.3995 | 5.4641             | 0   | 0       |
| 1443866_at   | Lrtm1        | 47.0388 | 4.7606             | 0   | 0       |
| 1426758_s_at | Meg3         | 92.0453 | 4.143              | 0   | 0       |
| Down genes   |              |         |                    |     |         |
| 1418486_at   | Vnn1         | 4.9119  | 0.0246             | 0   | 0       |
| 1435275_at   | Cox6b2       | 5.3466  | 0.0359             | 0   | 0       |
| 1424853_s_at | LOC100044218 | 8.8218  | 0.0238             | 0   | 0       |
| 1441915_s_at | Plin5        | 13.9657 | 0.0772             | 0   | 0       |
| 1423257_at   | Cyp4a14      | 15.1207 | 0.0343             | 0   | 0       |
| 1424937_at   | Plin5        | 24.5962 | 0.1048             | 0   | 0       |
| 1447845_s_at | Vnn1         | 26.6864 | 0.0441             | 0   | 0       |
| 1421430_at   | Rad5111      | 30.7707 | 0.0907             | 0   | 0       |
| 1420603_s_at | Raet1e       | 38.0251 | 0.1325             | 0   | 0       |
| 1440134_at   | LOC630172    | 39.4119 | 0.0621             | 0   | 0       |

conjugation of many hormones, neurotransmitters, drugs and xenobiotic compounds, was identified at the first rank with the RP value as 8.7008. The human gene of SULT1E1 was reported to encode a protein that transfers a sulpho moiety to and from oestrone and may primarily control levels of oestrogen receptors (Gong *et al.*, 2008). Comparison of murine and human oestrogen sulphotransferase inhibition *in vitro* and *in silico*, it suggested that catalytically inactive binding modes, other than the one observed in the crystal structures, be possible in SULT1E1 (Stjernschantz *et al.*, 2010). Moreover, the function of SULT1E1 was also found to be related to steroid sulphotransferase activity by observing the bacterial expression and characterization of a cDNA for human liver oestrogen sulphotransferase (Falany *et al.*, 1995). These findings may be associated with the regulation of PPAR $\alpha$  KO in lipid metabolism. Serpinb1a also named leukocyte elastase inhibitor a, one member of serine (or cysteine) peptidase inhibitors, was the second ranked one (RP = 17.3453). The regulation of pulmonary innate immunity by serpinb1 was thought to be non-redundant and required to protect two key components, the neutrophil and surfactant protein-D (Benarafa *et al.*, 2007). The most frequently identified gene in our top 10 was Meg3, which was represented by five probe sets all identified as significant. Meg3 named as maternally expressed 3, which is an imprinted gene with preferential expression from the maternal allele and closely linked to or co-expressed with reciprocally imprinted paternally expressed Dlk1 gene. This co-regulation

suggests a causative role in the pathologies found in uniparental disomy animals, characterized by defects in skeletal muscle maturation, bone formation, placenta size and organization and prenatal lethality. Meg3 may play an important role in control of vascularization in the brain and may function as a tumour suppressor in part by inhibiting angiogenesis (Gordon *et al.*, 2010). Another top gene Clta named Clathrin, light polypeptide (Lca) is a large, soluble protein composed of heavy and light chains. It functions as the main structural component of the lattice-type cytoplasmic face of coated pits and vesicles which entrap specific macromolecules during receptor-mediated endocytosis. This gene encodes one of two clathrin light chain proteins that are believed to function as regulatory elements. A B-Myb complex containing clathrin and filamin, which was previously reported to be required for normal localization of clathrin at the mitotic spindle stabilizing kinetochore fibres, may contribute to its role in chromosome stability, possibly, tumour suppression (Yamauchi *et al.*, 2008). The seventh ranked up-regulated gene NNMT is named as nicotinamide *N*-methyltransferase, nicotinamide treatment may induce NNMT enzyme activity and led to an increase in adipose tissue homocysteine secretion in tissue culture (Riederer *et al.*, 2009). The last one Lrtm1 named leucine-rich repeats and transmembrane domains 1, a gene of unknown function embedded within CACNA2D3 encoding a voltage-dependent calcium channel, which would be needed for further study in the research of our interest.

Table 3. The significantly identified pathways in meta-analysis

| Pathways  | KEGG maps                                 | Count | P-value  | Genes   |
|---|---|-------|----------|---|
| Up-regulated  |   |       |          |   |
| mmu04610:complement and coagulation cascades          | Immune system                             | 32    | 4·06E-14 | C1s, Serpina1d, C9, F5, Serpind1, C4b, F7, C8a, F9, Vwf, Serpina1c, C4bp, F2r, F13b, Kng1, Cfi, C1qa, C8b, Klkb1, C1r, C6, Serpina1e, C1qb, Hc, Serpina5, F11, Serpinf2, Serping1, C2, F10, Mbl2, Maspl   |
| mmu00980:metabolism of xenobiotics by cytochrome p450 | Xenobiotics biodegradation and metabolism | 20    | 1·05E-05 | Cyp2c54, Ugt2a3, Cyp1a2, Cyp2c50, Cyp2b10, Gstm2, Ugt2b1, Gstm3, Gsta2, Gsta4, Cyp2c70, Gsta1, Cyp2c37, Cyp2c55, Cyp2c65, Adh4, Gstm6, Cyp2c68, Gstm7, Cyp2c44  |
| mmu00260:glycine, serine and threonine metabolism     | Amino acid metabolism                     | 14    | 2·00E-04 | Gcat, Gamt, Gnmt, Sardh, Chka, Cbs, Pipox, Dmgdh, Agxt, Alas2, Agxt2, Bhmt, Pisd, Cth   |
| mmu00590:arachidonic acid metabolism                  | Lipid metabolism                          | 15    | 4·60E-03 | Cyp2c54, Pla2g12a, Cyp2c50, Cyp4f16, Cyp2b10, Ptgds, Cyp2c70, Cyp4f14, Pla2g6, Cyp2u1, Cyp2c37, Cyp2c55, Cyp2c65, Cyp2c68, Cyp2c44  |
| mmu00591:linoleic acid metabolism                     | Lipid metabolism                          | 11    | 6·26E-03 | Cyp2c54, Pla2g12a, Cyp1a2, Cyp2c50, Cyp2c70, Pla2g6, Cyp2c55, Cyp2c37, Cyp2c65, Cyp2c68, Cyp2c44  |
| mmu00910:nitrogen metabolism                          | Energy metabolism                         | 7     | 8·06E-03 | Asns, Car1, Car14, Glc2, Car3, Cth, Hal   |
| Down-regulated  |   |       |          |   |
| mmu00071:fatty acid metabolism                        | Lipid metabolism                          | 28    | 1·93E-18 | Acs15, Acadv1, Cpt1b, Acaa1b, Acat1, Hsd17b10, Aldh9a1, Hsd17b4, Hadhb, Acox1, Acaa1a, Acs11, Acadl, Acs14, Dci, Aldh3a2, Acadm, Acads, Acaa2, Hadha, Cyp4a10, Acs13, Peci, Cpt1a, Ehhadh, Cpt2, Cyp4a14, Hadh                                      |
| mmu03320:PPAR signalling pathway                      | Endocrine system                          | 34    | 2·04E-16 | Slc27a4, Slc27a1, Pltp, Acs15, Fads2, Fabp2, Cpt1b, Angptl4, Slc27a2, Ppara, Cyp7a1, Acox1, Acaa1a, Acs11, Acadl, Acs14, Fabp4, Fabp3, Sorbs1, Acadm, Dbi, Cd36, Fabp1, Cyp4a10, Scd1, Acs13, Cpt1a, Fabp5, Ehhadh, Lpl, Cpt2, Me1, Cyp4a14, Hmgcs2 |
| mmu00280:valine, leucine and isoleucine degradation   | Amino acid metabolism                     | 17    | 3·80E-07 | Acadm, Aldh3a2, Acaa2, Acads, Acaa1b, Acat1, Hsd17b10, Aldh9a1, Hadha, Hsd17b4, Ehhadh, Hadhb, Hmgcl, Acaa1a, Hmgcs2, Hadh, Hmgcs1  |
| mmu01040:polyunsaturated fatty acid biosynthesis      | Lipid metabolism                          | 10    | 3·63E-06 | Fasn, Hadha, Scd1, Elovl5, Fads2, Acox1, Acaa1a, Fads1, Elovl6, Hsd17b12  |
| mmu04110:cell cycle                                   | Cell growth and death                     | 25    | 2·56E-05 | Ccnd3, Gadd45b, Orc5l, Cdc2a, Prkdc, Gadd45g, Mcm4, Ccnd1, Chek1, Cdk4, Mcm5, Cdc20, Mad2l1, Mcm2, Mcm6, Mcm7, Skp1a, Cdc45l, Wee1, Ywhag, Ccnb1, Cdkn1a, Mcm3, Ccnb2, Pttg1  |
| mmu04115:p53 signalling pathway                       | Cell growth and death                     | 17    | 1·28E-04 | Cycs, Ccnd3, Ei24, Ccng2, Gadd45b, Cdc2a, Gadd45g, Perp, Ccnb1, Ddb2, Cdkn1a, Casp8, Ccnd1, Ccnb2, Fas, Cdk4, Chek1   |
| mmu00650:butanoate metabolism                         | Carbohydrate metabolism                   | 14    | 3·23E-04 | Aldh3a2, Acads, Acat1, Hsd17b10, Aldh9a1, Ilvbl, Hadha, Hsd17b4, Aacs, Ehhadh, Hmgcl, Hmgcs2, Hadh, Hmgcs1  |
| mmu00062:fatty acid elongation in mitochondria        | Lipid metabolism                          | 6     | 8·83E-04 | Hsd17b10, Hadha, Hsd17b4, Hadhb, Acaa2, Hadh  |
| mmu00620:pyruvate metabolism                          | Carbohydrate metabolism                   | 12    | 1·11E-03 | Aldh9a1, Acot12, Acaca, Akr1b3, Aldh3a2, Me1, Dlat, Acyp1, Acss2, Acyp2, Glo1, Acat1  |
| mmu00640:propanoate metabolism                        | Carbohydrate metabolism                   | 10    | 1·27E-03 | Aldh9a1, Hadha, Acaca, Ehhadh, Aldh3a2, Acadm, Mlycd, Suclg1, Acss2, Acat1  |
| mmu00930:caprolactam degradation                      | Xenobiotics biodegradation and metabolism | 6     | 2·39E-03 | Hsd17b10, Sirt5, Hadha, Hsd17b4, Ehhadh, Hadh   |
| mmu00564:glycerophospholipid metabolism               | Lipid metabolism                          | 13    | 5·83E-03 | Ptdss2, Ppap2c, Chkb, Ptdss1, Agpat3, Pla2g6, Gpd2, Nat6, Gpd1, Phospho1, Cris1, Etnk1, Gpam  |
| mmu00561:glycerolipid metabolism                      | Lipid metabolism                          | 10    | 9·20E-03 | Aldh9a1, Agpat3, Akr1b3, Aldh3a2, Lpl, Gla, Ppap2c, Lipg, Dgat1, Gpam   |

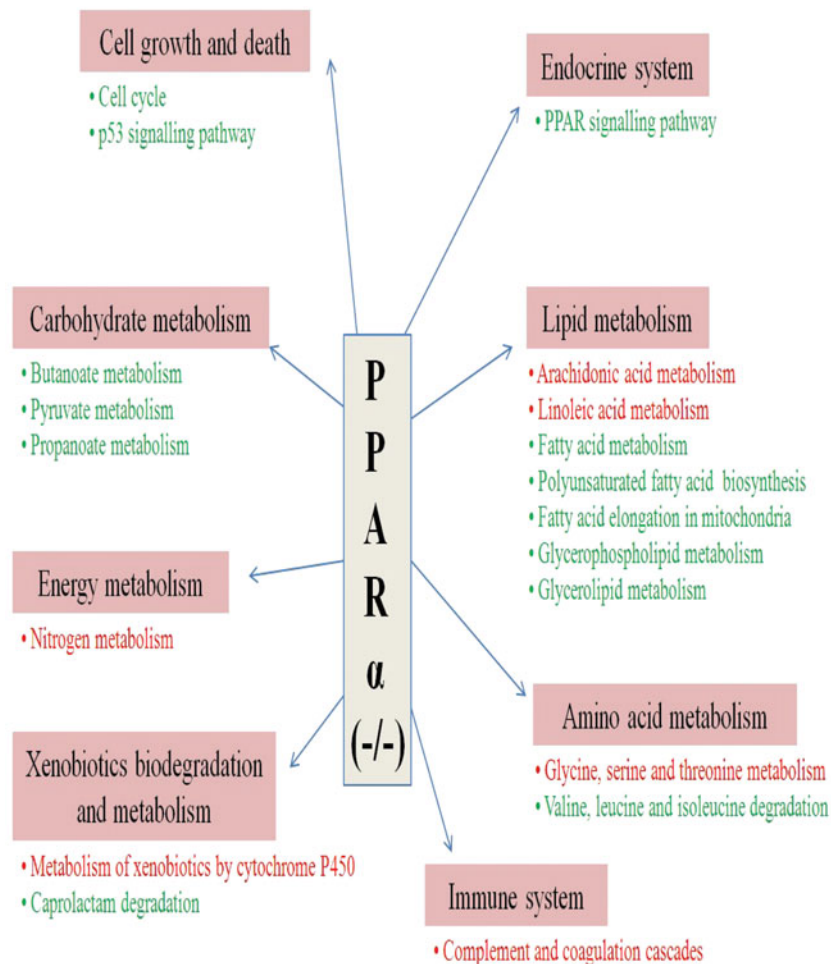


Fig. 2. PPAR $\alpha$  regulates a variety of biological processes in mouse liver. Summary of functional implication of PPAR $\alpha$  activation as assessed by our meta-analysis based on KEGG pathways enrichment. The names in shaded boxes represent eight identified KEGG pathway maps, following with significant pathways' names, up-regulated in red and down-regulated in green under PPAR $\alpha$ (-/-).

On the other hand, among the top 10 ranked down-regulated genes, the gene of *Vnn1* named vanin 1 was identified twice with the represented probe sets at the first rank (RP=4.9119) and the seventh rank (RP=26.6864). Mouse vanin 1 was reported to be involved in many biological processes in mouse, from thymus homing to sexual development and indicated to be cytoprotective for islet  $\beta$  cells and regulates the development of type 1 diabetes (Di Leandro *et al.*, 2008; Roisin-Bouffay *et al.*, 2008). The second one was *Cox6b2* named cytochrome *c* oxidase subunit 6B2 participating in some disease pathways with the function of cytochrome *c* oxidase activity (Pagliarini *et al.*, 2008); however, the exact knowledge about regulation of this gene on lipid metabolism is limited. Another identified gene *Cyp4a14* was a member of the family of cytochrome p450 enzymes (CyPs), which are major phase-I xenobiotic-metabolizing enzymes. The previous study on mice liver indicated that perfluorodecanoic acid (PFDA) may increase *Cyp4A14* mRNA expression in WT mice, but much less in

PPAR $\alpha$ -null mice (Cheng & Klaassen, 2008). Moreover, a recent study suggests that the *Cyp4a14*-deficient mouse may be a useful model for evaluation of NO/20-HETE interactions, the production of which in the kidney is thought to be involved in the control of renal vascular tone and tubular sodium and chloride reabsorption (Fidelis *et al.*, 2010). The gene *Plin5* named perilipin 5 was also identified twice in our study. The perilipin is an adipose differentiation-related protein, tail-interacting protein of 47 kd (PAT) proteins that are thought to regulate intracellular lipid droplets (LDs) turnover by modulating lipolysis (Bell *et al.*, 2008). Mouse *Rad5111* is a radiation-inducible gene that regulates cell-cycle progression. The previous study based on disruption in embryonic stem cells by homologous recombination suggested that *Rad5111* would play a role in cell proliferation and early mice embryonic development, perhaps through interaction with p53 (Shu *et al.*, 1999). The gene of *Raet1e* named as retinoic acid early transcript 1E may promote tumour surveillance

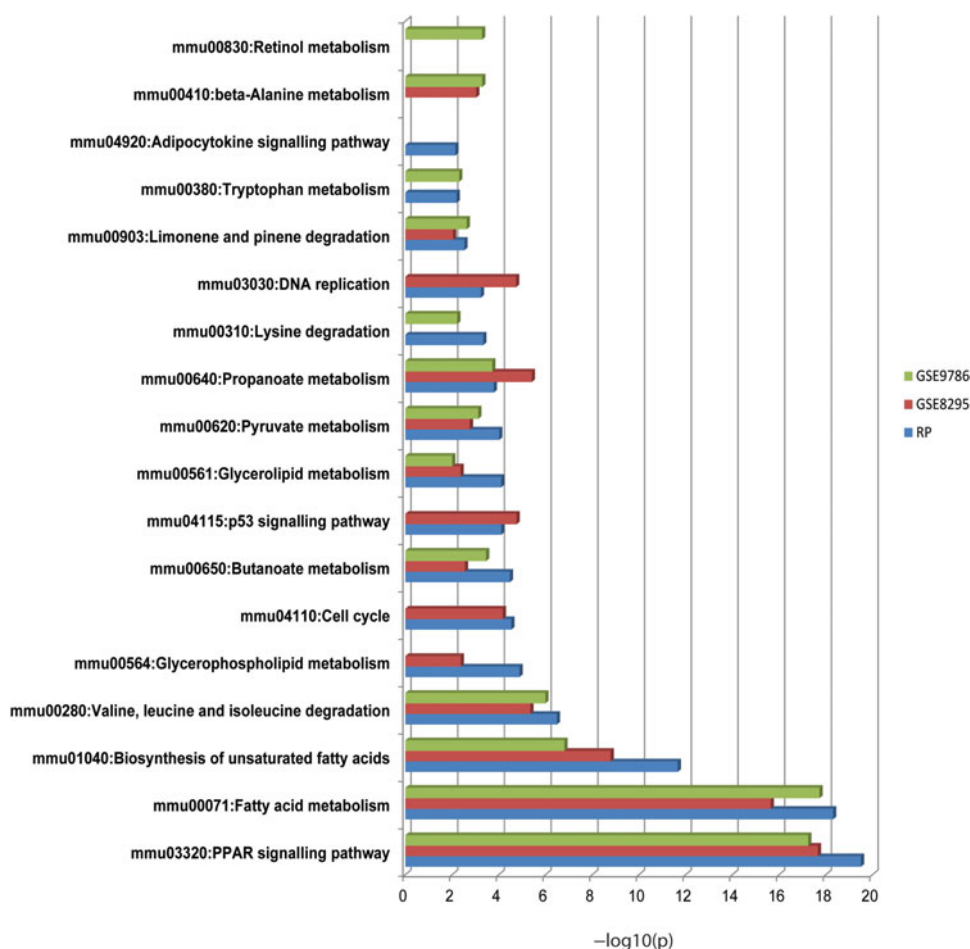


Fig. 3. Comparison of the significance of down-regulated pathways identified among individual analysis and meta-analysis. The chart was showing the significance of down-regulated pathways identified under each analysis. X-axis represents  $-\log(P)$  values; Y-axis represents the names of each identified pathways. The result of individual analysis of GSE9786 was marked in green, GSE8295 was marked in red and the meta-analysis was marked in blue.

and autoimmunity by engaging the activating receptor NKG2D on natural killer (NK) cells and T cells (Oppenheim *et al.*, 2005). Two hypothetical protein LOC100044218 and LOC630172 were both reported here as the top ones, whose annotation status are not on current assembly.

The subsequent pathway analysis revealed that there were 13 down-regulated and six up-related pathways participating in regulating PPAR $\alpha$  activation. The whole information of significantly identified pathways in our meta-analysis are shown in Table 3, including the pathway names, the count of genes in each identified pathway, the significance ( $P$ -value) of identified pathways and the related genes in each pathway. According to our analysis based on KEGG pathway database (<http://www.genome.jp/kegg/>), we can map our identified pathways into different functional classification. For up-regulated pathways, the top significant pathway of complement and coagulation cascades ( $P=4.06 \times 10^{-14}$ ) was mapped to immune system, metabolism of xenobiotics by cytochrome p450 ( $1.05 \times 10^{-5}$ ) was mapped

to xenobiotics biodegradation and metabolism, glycine, serine and threonine metabolism ( $2.00 \times 10^{-4}$ ) was mapped to amino acid metabolism, nitrogen metabolism ( $8.06 \times 10^{-3}$ ) was mapped to energy metabolism, both arachidonic acid metabolism ( $4.60 \times 10^{-3}$ ) and linoleic acid metabolism ( $6.26 \times 10^{-3}$ ) were mapped to lipid metabolism. For down-regulated pathways, besides five pathways including fatty acid metabolism ( $1.93 \times 10^{-18}$ ), polyunsaturated fatty acid biosynthesis ( $3.63 \times 10^{-6}$ ), fatty acid elongation in mitochondria ( $8.83 \times 10^{-4}$ ), glycerophospholipid metabolism ( $5.83 \times 10^{-3}$ ) and glycerolipid metabolism ( $9.20 \times 10^{-3}$ ) in the classification of lipid metabolism, the pathway of valine, leucine and isoleucine degradation ( $3.80 \times 10^{-7}$ ) in the classification of amino acid metabolism, the pathway of caprolactam degradation ( $2.39 \times 10^{-3}$ ) in the classification of xenobiotics biodegradation and metabolism, more classifications were found including PPAR signalling pathway ( $2.04 \times 10^{-16}$ ) of endocrine system, two pathways of cell cycle ( $2.56 \times 10^{-5}$ ) and p53 signalling pathway ( $1.28 \times 10^{-4}$ ) in the



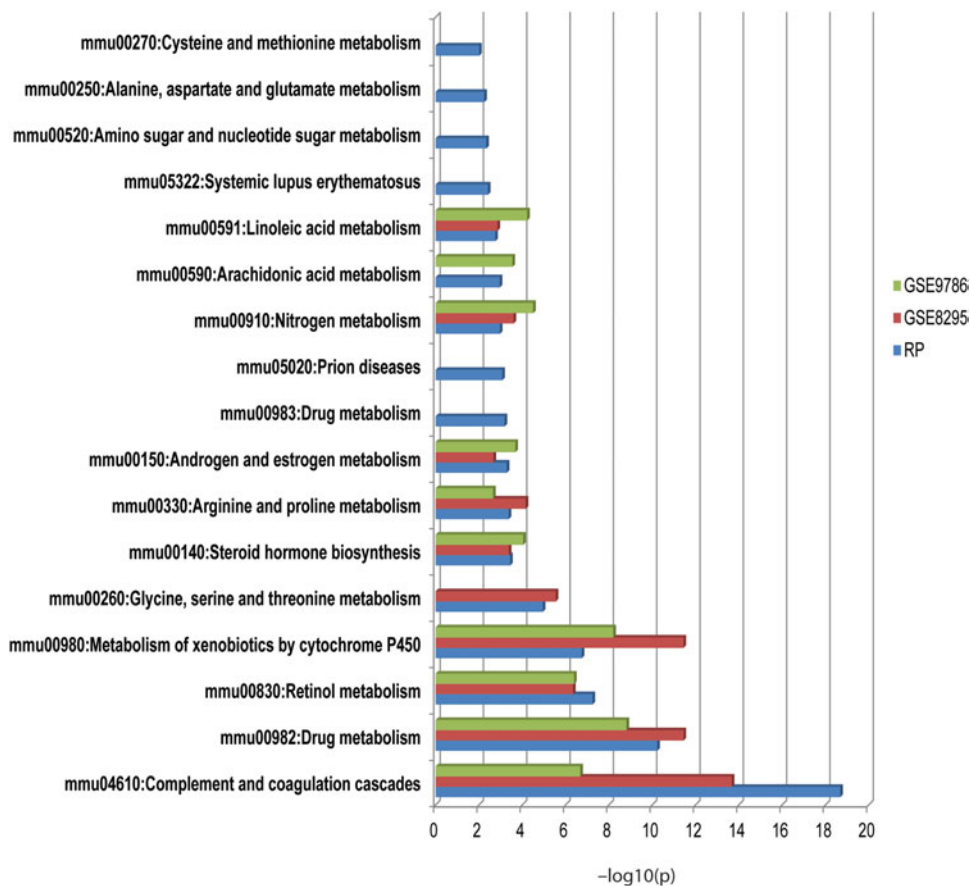


Fig. 4. Comparison of the significance of up-regulated pathways identified among individual analysis and meta-analysis. The chart was showing the significance of up-regulated pathways identified under each analysis. X-axis represents  $-\log(P)$  values; Y-axis represents the names of each identified pathways. The result of individual analysis of GSE9786 was marked in green, GSE8295 was marked in red and the meta-analysis (based on the method of RP) was marked in blue.

classification of cell growth and death, three pathways including butanoate metabolism ( $3.23 \times 10^{-4}$ ), pyruvate metabolism ( $1.11 \times 10^{-3}$ ), and propanoate metabolism ( $1.27 \times 10^{-3}$ ) in the classification of carbohydrate metabolism. A variety of biological processes regulated by PPAR $\alpha$  in mouse liver from our re-analysis results were summarized in Fig. 2.

On comparison with individual results of pathway analysis, there were remarkable differences not only in the terms but also in the significance of identified pathways. The information on comparison of the significance of down- or up-regulated pathways identified among individual analysis and meta-analysis was shown in Figs 3 and 4. Actually, there appeared to be some common pathways among these studies, such as two up-regulated pathways including complement and coagulation cascades and metabolism of xenobiotics by cytochrome p450 and seven down-regulated pathways including PPAR signalling pathway, fatty acid metabolism, propanoate metabolism, the pathway of valine, leucine and isoleucine degradation, pyruvate metabolism, butanoate metabolism and glycerolipid metabolism. However, their significances were not the same. In addition, the

results based on meta-analysis exhibited some novel associated pathways such as polyunsaturated fatty acid biosynthesis, fatty acid elongation in mitochondria and caprolactam degradation, related to diverse metabolisms.

In order to identify the gene expression patterns of PPAR $\alpha$ , a type of meta-analysis based on the method of RP was employed to combine datasets from different origins (datasets of GSE9786 and GSE8295) which were contradicted in some aspect of analysis results. As a result, it increases the power of the identification of differentially expressed genes and significant pathways in this study. Some novel hepatic tissue-specific marker genes as well as pathways related to PPAR $\alpha$  were detected. These may be helpful for us to further uncover the regulation mechanisms of PPAR $\alpha$  in mice liver.

Our work is funded by the National Natural Science Foundation of China (grant numbers 30871782, 31072003, 3100992, 31101706, 31272414), the Ministry of Science and Technology of the People's Republic of China (grant number 2006AA10Z1E3), National Programmes for High Technology Research and Development of China (863 programme) (grant number 2008AA101009), the National

Key Technology R&D Programme (grant number 2008BADB2B11) and the National 948 Project of China (2012-Z26, 2011-G2A).

#### 4. Declaration of interest

The authors declare that they have no competing interests.

#### 5. Supplementary material

The online data can be found available at <http://journals.cambridge.org/GRH>

#### Authors' contributions

KH designed the study, collected the datasets from databases and analysed the data, then prepared the original draft of the manuscript. ZW, QW and YP designed the study and reviewed the manuscript. All authors read and approved the final manuscript.

#### References

- Bell, M., Wang, H., Chen, H., McLenithan, J. C., Gong, D. W., Yang, R. Z., Yu, D., Fried, S. K., Quon, M. J., Londos, C. & Sztalryd, C. (2008). Consequences of lipid droplet coat protein downregulation in liver cells: abnormal lipid droplet metabolism and induction of insulin resistance. *Diabetes* **57**, 2037–2045.
- Benarafa, C., Priebe, G. P. & Remold-O'Donnell, E. (2007). The neutrophil serine protease inhibitor serpinb1 preserves lung defense functions in *Pseudomonas aeruginosa* infection. *Journal of Experimental Medicine* **204**, 1901–1909.
- Bookout, A. L., Jeong, Y., Downes, M., Yu, R. T., Evans, R. M. & Mangelsdorf, D. J. (2006). Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network. *Cell* **126**, 789–799.
- Braissant, O., Foufelle, F., Scotto, C., Dauca, M. & Wahli, W. (1996). Differential expression of peroxisome proliferator-activated receptors (PPARs): tissue distribution of PPAR- $\alpha$ , - $\beta$ , and - $\gamma$  in the adult rat. *Endocrinology* **137**, 354–366.
- Breitling, R., Armengaud, P., Amtmann, A. & Herzyk, P. (2004). Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS Letters* **573**, 83–92.
- Cheng, X. G. & Klaassen, C. D. (2008). Perfluorocarboxylic acids induce cytochrome p450 enzymes in mouse liver through activation of PPAR- $\alpha$  and CAR transcription factors. *Toxicological Sciences* **106**, 29–36.
- Dennis, G., Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C. & Lempicki, R. A. (2003). DAVID: database for annotation, visualization, and integrated discovery. *Genome Biology* **4**, P3.
- Devchand, P. R., Keller, H., Peters, J. M., Vazquez, M., Gonzalez, F. J. & Wahli, W. (1996). The PPAR $\alpha$ -leukotriene B<sub>4</sub> pathway to inflammation control. *Nature* **384**, 39–43.
- Di Leandro, L., Maras, B., Schinina, M. E., Dupre, S., Koutris, I., Martin, F. M., Naquet, P., Galland, F. & Pitari, G. (2008). Cystamine restores GSTA3 levels in Vanin-1 null mice. *Free Radical Biology and Medicine* **44**, 1088–1096.
- Falany, C. N., Krasnykh, V. & Falany, J. L. (1995). Bacterial expression and characterization of a cDNA for human liver estrogen sulfotransferase. *Journal of Steroid Biochemistry and Molecular Biology* **52**, 529–539.
- Fidelis, P., Wilson, L., Thomas, K., Villalobos, M. & Oyekan, A. O. (2010). Renal function and vasomotor activity in mice lacking the Cyp4a14 gene. *Experimental Biology and Medicine* **235**, 1365–1374.
- Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. (2004). Affy – analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* **20**, 307–315.
- Gentleman, R. C., Carey, V. J., Bates, D. M., Bolstad, B., Dettling, M., Dudoit, S., Ellis, B., Gautier, L., Ge, Y., Gentry, J., Hornik, K., Hothorn, T., Huber, W., Iacus, S., Irizarry, R., Leisch, F., Li, C., Maechler, M., Rossini, A. J., Sawitzki, G., Smith, C., Smyth, G., Tierney, L., Yang, J. Y. & Zhang, J. (2004). Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* **5**, R80.
- Gong, H., Jarzynka, M. J., Cole, T. J., Lee, J. H., Wada, T., Zhang, B., Gao, J., Song, W. C., DeFranco, D. B., Cheng, S. Y. & Xie, W. (2008). Glucocorticoids antagonize estrogens by glucocorticoid receptor-mediated activation of estrogen sulfotransferase. *Cancer Research* **68**, 7386–7393.
- Gordon, F. E., Nutt, C. L., Cheunsuchon, P., Nakayama, Y., Provencher, K. A., Rice, K. A., Zhou, Y., Zhang, X. & Klibanski, A. (2010). Increased expression of angiogenic genes in the brains of mouse Meg3-Null embryos. *Endocrinology* **151**, 2443–2452.
- Huang, D. W., Sherman, B. T. & Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**, 44–57.
- Issemann, I., Green, S. (1990). Activation of a member of the steroid hormone receptor superfamily by peroxisome proliferators. *Nature* **347**, 645–650.
- Lee, Y., Yu, X., Gonzales, F., Mangelsdorf, D. J., Wang, M. Y., Richardson, C., Witters, L. A. & Unger, R. H. (2002). PPAR  $\alpha$  is necessary for the lipopenic action of hyperleptinemia on white adipose and liver tissue. *Proceedings of the National Academy of Sciences USA* **99**, 11848–11853.
- Lefebvre, P., Chinetti, G., Fruchart, J. C. & Staels, B. (2006). Sorting out the roles of PPAR  $\alpha$  in energy metabolism and vascular homeostasis. *Journal of Clinical Investigation* **116**, 571–580.
- Leone, T. C., Weinheimer, C. J. & Kelly, D. P. (1999). A critical role for the peroxisome proliferator-activated receptor  $\alpha$  (PPAR $\alpha$ ) in the cellular fasting response: the PPAR $\alpha$ -null mouse as a model of fatty acid oxidation disorders. *Proceedings of the National Academy of Sciences USA* **96**, 7473–7478.
- Mandard, S., Muller, M. & Kersten, S. (2004). Peroxisome proliferator-activated receptor  $\alpha$  target genes. *Cellular and Molecular Life Sciences* **61**, 393–416.
- Michalik, L., Auwerx, J., Berger, J. P., Chatterjee, V. K., Glass, C. K., Gonzalez, F. J., Grimaldi, P. A., Kadowaki, T., Lazar, M. A., O'Rahilly, S., Palmer, C. N., Plutzky, J., Reddy, J. K., Spiegelman, B. M., Staels, B. & Wahli, W. (2006). International Union of Pharmacology. LXI. Peroxisome proliferator-activated receptors. *Pharmacological Reviews* **58**, 726–741.
- Nemhauser, J. L., Mockler, T. C. & Chory, J. (2004). Interdependency of brassinosteroid and auxin signaling in Arabidopsis. *PLoS Biology* **2**, 1460–1471.
- Oppenheim, D. E., Roberts, S. J., Clarke, S. L., Filler, R., Lewis, J. M., Tigelaar, R. E., Girardi, M. & Hayday, A. C. (2005). Sustained localized expression of ligand

- for the activating NKG2D receptor impairs natural cytotoxicity *in vivo* and reduces tumor immunosurveillance. *Nature Immunology* **6**, 928–937.
- Pagliarini, D. J., Calvo, S. E., Chang, B., Sheth, S. A., Vafai, S. B., Ong, S. E., Walford, G. A., Sugiana, C., Boneh, A., Chen, W. K., Hill, D. E., Vidal, M., Evans, J. G., Thorburn, D. R., Carr, S. A. & Mootha, V. K. (2008). A mitochondrial protein compendium elucidates complex I disease biology. *Cell* **134**, 112–123.
- Park, T., Yi, S. G., Lee, S., Lee, S. Y., Yoo, D. H., Ahn, J. I. & Lee, Y. S. (2003). Statistical tests for identifying differentially expressed genes in time-course microarray experiments. *Bioinformatics* **19**, 694–703.
- Rakhshandehroo, M., Sanderson, L. M., Matilainen, M., Stienstra, R., Carlberg, C., de Groot, P. J., Muller, M. & Kersten, S. (2007). Comprehensive analysis of PPARalpha-dependent regulation of hepatic lipid metabolism by expression profiling. *PPAR Research* **2007**, 26839.
- Reddy, J. K. Hashimoto, T. (2001). Peroxisomal beta-oxidation and peroxisome proliferator-activated receptor alpha: an adaptive metabolic system. *Annual Review of Nutrition* **21**, 193–230.
- Riederer, M., Erwa, W., Zimmermann, R., Frank, S. & Zechner, R. (2009). Adipose tissue as a source of nicotinamide N-methyltransferase and homocysteine. *Atherosclerosis* **204**, 412–417.
- Roisin-Bouffay, C., Castellano, R., Valero, R., Chasson, L., Galland, F. & Naquet, P. (2008). Mouse vanin-1 is cytoprotective for islet beta cells and regulates the development of type 1 diabetes. *Diabetologia* **51**, 1192–1201.
- Rosen, M. B., Lee, J. S., Ren, H., Vallanat, B., Liu, J., Waalkes, M. P., Abbott, B. D., Lau, C. & Corton, J. C. (2008). Toxicogenomic dissection of the perfluorooctanoic acid transcript profile in mouse liver: evidence for the involvement of nuclear receptors PPAR alpha and CAR. *Toxicological Sciences* **103**, 46–56.
- Shu, Z. G., Smith, S., Wang, L. J., Rice, M. C. & Kmiec, E. B. (1999). Disruption of muREC2/RAD51L1 in mice results in early embryonic lethality which can be partially rescued in a p53(–/–) background. *Molecular and Cellular Biology* **19**, 8686–8693.
- Stjernaschantz, E., Reinen, J., Meinel, W., George, B. J., Glatt, H., Vermeulen, N. P. E. & Oostenbrink, C. (2010). Comparison of murine and human estrogen sulfotransferase inhibition *in vitro* and *in silico* – implications for differences in activity, subunit dimerization and substrate inhibition. *Molecular and Cellular Endocrinology* **317**, 127–140.
- Yamauchi, T., Ishida, T., Nomura, T., Shinagawa, T., Tanaka, Y., Yonemura, S. & Ishii, S. (2008). A B-Myb complex containing clathrin and filamin is required for mitotic spindle function. *EMBO Journal* **27**, 1852–1862.