

OPEN ACCESS

Full open access to this and thousands of other papers at <http://www.la-press.com>.

An Integrated Statistical Approach to Compare Transcriptomics Data Across Experiments: A Case Study on the Identification of Candidate Target Genes of the Transcription Factor PPAR α

Mohammad Ohid Ullah¹, Michael Müller^{1,2} and Guido J.E.J. Hooiveld^{1,2}

¹Nutrition, Metabolism and Genomics group, Division of Human Nutrition, Wageningen University, The Netherlands.

²Netherlands Nutrigenomics Center, TI Food and Nutrition, Wageningen, The Netherlands.

Corresponding author email: guido.hooiveld@wur.nl

Abstract: An effective strategy to elucidate the signal transduction cascades activated by a transcription factor is to compare the transcriptional profiles of wild type and transcription factor knockout models. Many statistical tests have been proposed for analyzing gene expression data, but most tests are based on pair-wise comparisons. Since the analysis of microarrays involves the testing of multiple hypotheses within one study, it is generally accepted that one should control for false positives by the false discovery rate (FDR). However, it has been reported that this may be an inappropriate metric for comparing data across different experiments. Here we propose an approach that addresses the above mentioned problem by the simultaneous testing and integration of the three hypotheses (contrasts) using the cell means ANOVA model. These three contrasts test for the effect of a treatment in wild type, gene knockout, and globally over all experimental groups. We illustrate our approach on microarray experiments that focused on the identification of candidate target genes and biological processes governed by the fatty acid sensing transcription factor PPAR α in liver. Compared to the often applied FDR based across experiment comparison, our approach identified a conservative but less noisy set of candidate genes with same sensitivity and specificity. However, our method had the advantage of properly adjusting for multiple testing while integrating data from two experiments, and was driven by biological inference. Taken together, in this study we present a simple, yet efficient strategy to compare differential expression of genes across experiments while controlling for multiple hypothesis testing.

Keywords: false discovery rate, gene expression, transcriptomics, microarray, transcription factor, peroxisome proliferator-activated receptor alpha, ANOVA

Bioinformatics and Biology Insights 2012:6 145–154

doi: [10.4137/BBI.S9529](https://doi.org/10.4137/BBI.S9529)

This article is available from <http://www.la-press.com>.

© the author(s), publisher and licensee Libertas Academica Ltd.

This is an open access article. Unrestricted non-commercial use is permitted provided the original work is properly cited.



Introduction

Genome-wide transcriptional profiling, or transcriptomics, is extensively used to study how cells respond to certain stimuli or to diagnose and predict clinical outcomes.^{1–4} Transcription factors (TFs) are the key effectors which control gene expression. From a variety of research fields, including nutrition sciences, there is a major interest in characterizing the genes and networks that are controlled by transcription factors. Advances in genome-wide expression profiling methodologies and the availability of model systems offered new, powerful tools to address this.^{5–11}

An effective strategy to elucidate the signal transduction cascades activated by transcription factors is through transcriptional profiling. Transcription profiling can be applied on gain- and loss-of-function TF mutants, and changes in global gene expression that are associated with the various phenotypes could then be used for a comprehensive understanding of TF function.^{6,7,11–13} To this end, transcription factor target genes have to be efficiently and accurately identified from the transcriptomics data set. It is important to realize that from a biological perspective, TF target genes are only those genes that do significantly respond upon treatment with a potent agonist or gain of function, in wild type but not mutant (knockout) models. However, from a statistical inference point of view the identification of biological relevant target genes from such 2×2 factorial experiments is less straight-forward.

It is generally accepted that statistical testing is required to reliably identify differentially expressed genes (reviewed in eg, Allison et al.¹⁴). Moreover, since the statistical analysis of microarrays involves the testing of multiple hypotheses (genes) within one study, it is necessary to control for false positives. A frequently used metric to quantify the level of confidence any particular gene is differentially expressed, that takes into account multiple testing, is the false discovery rate (FDR).¹⁴ Therefore in many studies a cutoff based on the FDR rather than *P*-value is used to select significantly regulated genes within experiments, which subsequently are compared across experiments to identify transcription factor target genes. However, Higdon et al¹⁵ reported that the use of the FDR and its associated *q*-value may result in inconsistent and misleading interpretation of the comparisons across different experiments, especially when the effect sizes

of the experiments vary dramatically, as for example is the case when comparing effects of potent agonists in wild type and TF knockout models.

Therefore, the purpose of the work described in the current paper is to present a strategy that optimally integrates and controls for multiple hypothesis testing using data obtained from two biological systems that respond completely different to a treatment. We outline our approach using one of our datasets on the mouse peroxisome proliferator-activated receptor alpha (PPAR α).¹⁶ PPAR α is a TF belonging to the nuclear receptor superfamily, and is activated by a variety of compounds, including dietary fatty acids and their derivatives as well as synthetic agonists.^{17–19}

Material and Methods

Experimental data

We illustrate our approach (Fig. 1) on one of our publicly available datasets (Gene Expression Omnibus (GEO) accession: GSE8295). This dataset was generated to identify PPAR α target genes in mouse liver,¹⁶ and was also used by Higdon et al¹⁵ to illustrate the inappropriateness of using the FDR as cut-off metric when comparing two transcriptomics experiments with different effect sizes.

Briefly, pure bred wild type (129S1/SvImJ) and PPAR α -null (129S4/SvJae-Pparatm1Gonz/J) mice²⁰ were fed chow or chow supplemented with 0.1% WY14643 (ChemSyn, Lenexa, KS) for 5 days ($n = 4$ mice per group). WY14643, ({4-Chloro-6-[(2,3-dimethylphenyl)amino]-2-pyrimidinyl}-sulfanyl)acetic acid (CAS: 50892-23-4), is a chemical that was developed by the pharmaceutical industry to lower serum cholesterol. It is not used in clinical applications, but it is rather used as prototype chemical to induce peroxisome proliferation. WY14643 is a highly specific and potent agonist for PPAR α and is therefore often used in studies on this nuclear receptor.^{21,22} On the sixth day, mice were anaesthetized and livers were excised. Total RNA was prepared using TRIzol reagent (Invitrogen, Carlsbad, CA) followed by purification using the RNeasy mini kit (Qiagen, Hilden, Germany). RNA integrity was checked by chip analysis (Agilent 2100 bioanalyzer, Agilent Technologies, Amsterdam, The Netherlands) according to the manufacturer's instructions. RNA was judged as suitable for array hybridization only if samples exhibited intact bands corresponding to the

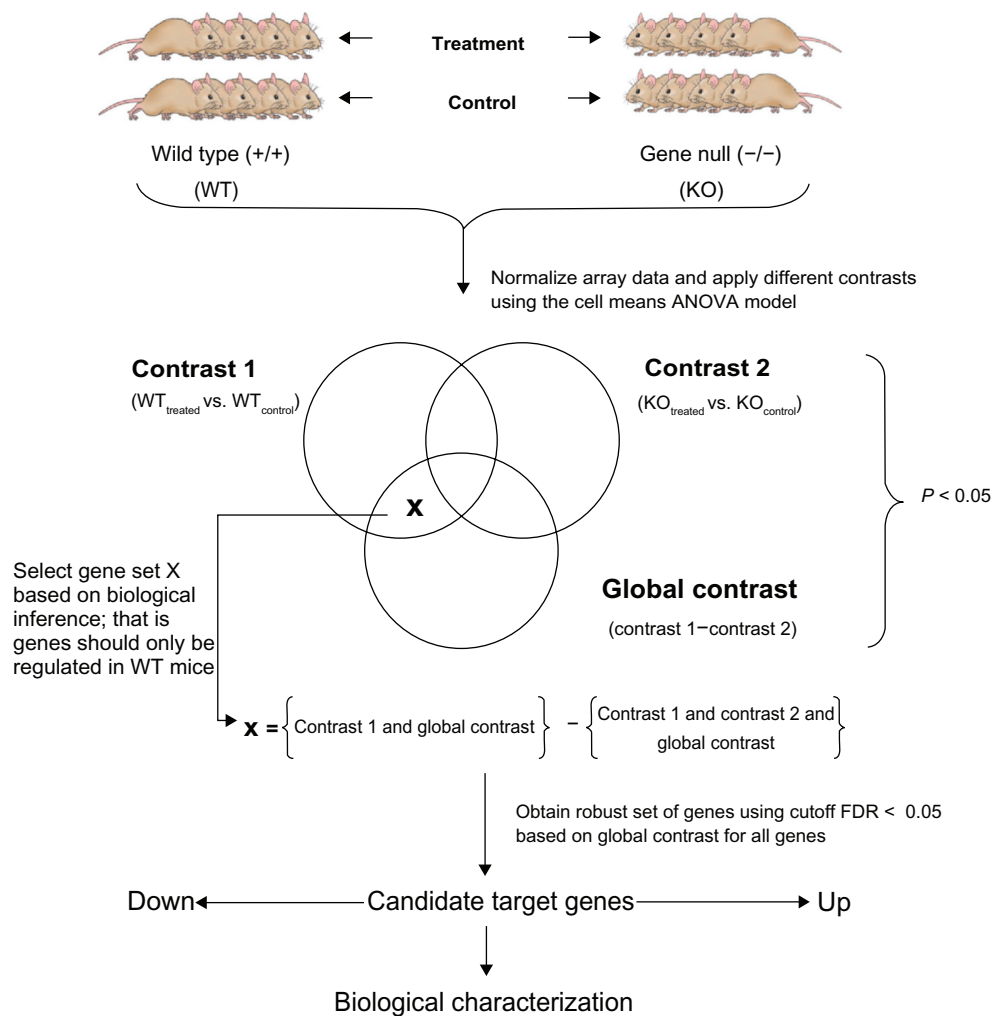


Figure 1. Overview of our integrated strategy.

Notes: After normalization, transcriptome data are analyzed for differentially expressed probesets (genes) using three contrasts (comparisons): Contrast 1, representing probesets regulated by a specific treatment in wild type mice; Contrast 2, representing probesets regulated by the same treatment but in knockout mice, and Global Contrast, representing genes differentially regulated by the treatment between the WT and KO mice. Biologically irrelevant probesets, ie, probesets that are also regulated by the treatment in the KO mice, are discarded, resulting in a set of probesets called X. To correct for multiple testing, FDR values of the probesets in X are calculated using the *P*-values obtained in Global Contrast for all probesets. A robust set of putative target genes regulated by the knocked-out gene is obtained by selecting those probesets from X that fulfill a Global Contrast-based FDR cutoff, eg, $FDR < 0.05$. This set can subsequently be divided in up- and down-regulated genes.

18S and 28S ribosomal RNA subunits, and displayed no chromosomal peaks or RNA degradation products, and had a RNA integrity number (RIN) above 8.0). The Affymetrix GeneChip RNA One cycle Amplification Kit (Affymetrix, Santa Clara, CA) was used to prepare labeled cRNA from 5 μ g of total RNA, which subsequently was hybridized on Affymetrix Mouse Genome 430 2.0 plus arrays. The animal study was approved by the Local Committee for Care and Use of Laboratory Animals.

Cell means ANOVA model

The dataset on the identification of PPAR α target genes in mouse liver has a 2×2 factorial design;

that is factor ‘treatment’ has 2 levels (WT, Control), as has the factor ‘genotype’ (wild type, knockout). Analysis of variance (ANOVA) is commonly used for analyzing data from experiments with multiple categorical factors.^{23,24} To appropriately identify candidate PPAR α target genes, we propose to perform and integrate three comparisons using the cell means ANOVA model.²⁵ For every probeset the model was defined as follows:

$$Y_{ijk} = \mu_{ij} + \epsilon_{ijk}$$

where Y_{ijk} is the expression of a probeset at i^{th} treatment (1 for WT, 2 for Control) in j^{th} strain of



genotype (1 for WT, 2 for KO) and k^{th} replication ($n = 4$), μ_{ij} is the mean value of i^{th} treatment and j^{th} strain of each gene, and ε_{ijk} is a random error term which follows normal distribution with mean = 0 and variance = σ^2 .

Formally, the definition of a contrast C is expressed below, using the notation μ_j for the j^{th} treatment mean:

$$C = c_1\mu_1 + c_2\mu_2 + \dots + c_j\mu_j + \dots + c_k\mu_k$$

where, $c_1 + \dots + c_j + \dots + c_k = \sum_{j=1}^k c_j = 0$

As stated before, from a biological perspective, candidate PPAR α target genes are only those genes that do significantly respond upon treatment with the potent PPAR α agonist WY14643 in wild type but not in PPAR α knockout mice. Therefore three different contrasts (comparisons) from this 2×2 factorial experiment were combined to infer the probesets that were significantly and PPAR α -dependently regulated. The different contrasts tested were (Table 1):

Contrast 1: $H_0: \mu_{11} - \mu_{21} = 0$ versus $H_1: \mu_{11} - \mu_{21} \neq 0$, returning all probesets regulated in the wild type mice by the agonist WY;

Contrast 2: $H_0: \mu_{12} - \mu_{22} = 0$ versus $H_1: \mu_{12} - \mu_{22} \neq 0$, returning all probesets regulated in the PPAR α knockout mice by the agonist WY; and

Global Contrast: $H_0: (\mu_{11} - \mu_{21}) - (\mu_{12} - \mu_{22}) = 0$ versus $H_1: (\mu_{11} - \mu_{21}) - (\mu_{12} - \mu_{22}) \neq 0$, returning the overall differential expressed probesets in wild type versus knockout mice groups after treatment with WY compared to control.

The PPAR α -dependently regulated probesets were then identified by extracting those probesets that were only significantly regulated in both Contrast 1

and Global Contrast, and subsequently corrected for multiple testing.

Implementation

All analyses were performed in R,²⁶ using packages from the Bioconductor project.²⁷ Probesets were redefined according to Dai et al.²⁸ In this study, probes were reorganized based on Entrez Gene database, build 36, version 2 (remapped CDF version 12). Our workflow was as follows (note that since we used a remapped chip definition file based on the Entrez Gene database, the terms probeset and gene are used interchangeably):

1. Expression estimates were obtained by GC-robust multiarray (GCRMA) normalization, using the empirical Bayes approach to adjust background.²⁹
2. For each of the three above-mentioned contrasts, differentially expressed probesets (genes) were identified using linear models, as implemented in *limma*.³⁰ For each contrast probesets were selected based on $P < 0.05$.
3. Probesets that were common only in Contrast 1 and the combined Global Contrast were identified. This set of probesets represented only transcription factor regulated genes, and was designated X.
4. Multiple testing was corrected for using a false discovery rate method,³¹ based on the Global Contrast considering the output of all probesets. Probesets in X that satisfied the criterion $FDR < 5\%$ were considered to be transcription factor target genes.

An schematic overview of our implementation is also given in the Figure 1, and the R-code and other required files are available as supplemental material.

Validation

To validate our integrated approach, obtained results (Fig. 2) were compared to results from the across experiment comparison (Fig. 3) using two sets of

Table 1. The contrasts defining the different hypotheses.

μ_{ij}	Levels	Contrast 1 $H_0: \mu_{11} - \mu_{21} = 0$	Contrast 2 $H_0: \mu_{12} - \mu_{22} = 0$	Global contrast $H_0: (\mu_{11} - \mu_{21}) - (\mu_{12} - \mu_{22}) = 0$
μ_{11}	WY, WT	1	0	1
μ_{12}	WY, KO	0	1	-1
μ_{21}	Con, WT	-1	0	-1
μ_{22}	Con, KO	0	-1	1

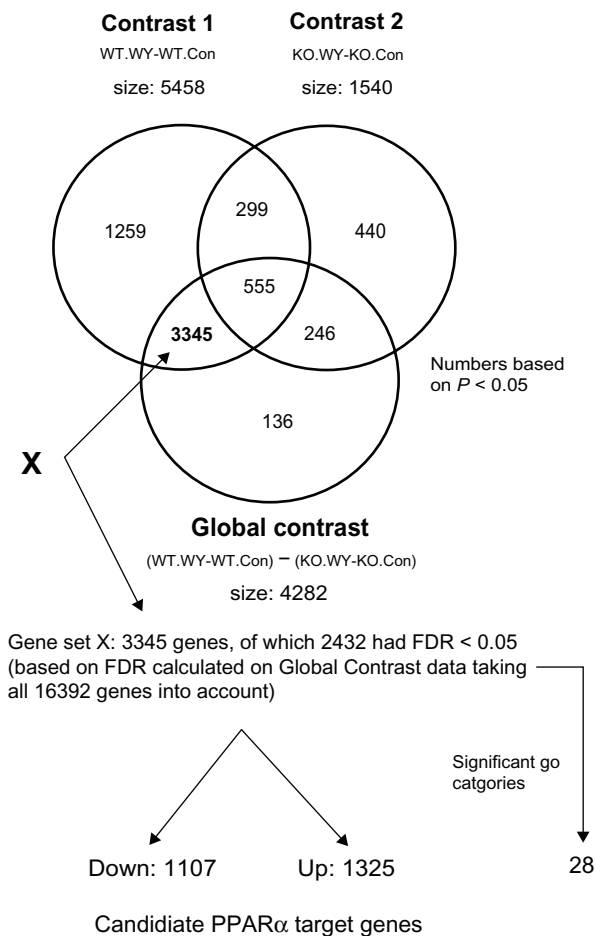


Figure 2. Application of the integrated approach on PPAR α dataset GSE8295.

Notes: Expression estimates were calculated by GCRMA normalization. Differentially expressed probesets were identified using three contrasts using P -value < 0.05 . Contrast 1, representing probesets regulated by the specific PPAR α agonist WY in wild type mice; Contrast 2, representing probesets regulated by WY14643 in PPAR α knockout mice, and Global Contrast, representing probesets differentially regulated by WY14643 between the WT and PPAR α KO mice. Biologically irrelevant probesets, i.e. those 854 probesets that were regulated by WY14643 in both WT and PPAR α KO mice, were discarded, resulting in a set of probesets called X of size 3345 that were only regulated in Contrast 1 and Global Contrast. To correct for multiple testing, FDR values (Benjamini Hochberg procedure) of the probesets in X were calculated based on the P -values for all probesets obtained in Global Contrast. A robust set of candidate PPAR α target genes was obtained by selecting those 2432 probesets from X that had Global Contrast-based FDR value < 0.05 . This set was divided in 1325 up- and 1107 down-regulated probesets.

well-established PPAR α target genes obtained from a recent review (Table 1 from Rakhshandehroo et al³²). These sets are available as supplemental material. The true positive rate (sensitivity) as function of the false positive rate (1-specificity) for different cutoff points was plotted for both the across experiment comparison and our integrated approach using the R-library ROCR.³³ The partial area under the ROC curve was calculated using $P = 0.2$ (thus 1-specificity = 0.2)

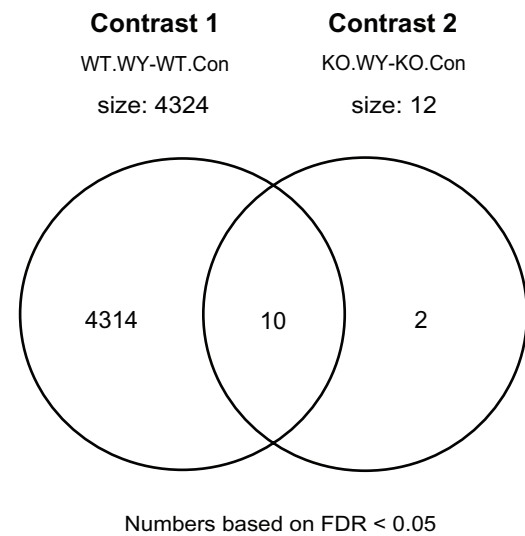


Figure 3. The FDR based across experiment comparison of PPAR α dataset GSE8295.

Notes: Expression estimates were calculated by GCRMA normalization. Differentially expressed probesets were identified in each contrast. Using a FDR value < 0.05 criterion, 4324 probesets were regulated in the wild type experiment (Contrast 1), whereas 12 genes were changed in the knockout experiment (Contrast 2). Of these 12 genes, 10 were also regulated in the wild type experiment. Thus, when comparing across experiments with a FDR value cutoff level of 0.05, 4314 genes were considered PPAR α target genes.

as cutoff. This cutoff value was chosen because for the identification of transcription factor target genes a high specificity is required ($>80\%$) before considering its sensitivity.³⁴ In addition, the biological features that were overrepresented in the lists of candidate PPAR α target genes that were generated on the basis of both approaches were analyzed with the software tool Ontologizer,³⁵ applying the ‘parent-child-union’ (PCU) algorithm and using the biological process ontology of Gene Ontology.

Results and Discussion

Identification of candidate PPAR α target genes

The application of transcriptomics to compare the effects of specific agonists, such as WY14643, in wild type and PPAR α knockout mice is a powerful approach to identify candidate PPAR α target genes.^{12,13} However, when comparing across different experiments the use of FDR cutoff values may result in inconsistent and misleading interpretation of the data.¹⁵ In this study we propose a simple yet effective strategy that avoids comparing probesets across experiments based on FDR values while still controlling for multiple testing. Testing three different

hypotheses (contrasts) for each probeset allowed the robust identification of transcription factor target genes. Since only the interaction effects are of interest for identifying candidate target genes, the cell means ANOVA model was used to infer this 2×2 factorial design.

The number of probesets significantly regulated ($P < 0.05$) upon PPAR α activation by WY14643 in wild type mice (= Contrast 1) equaled to 5458, whereas in PPAR α $-/-$ mice (Contrast 2) this number was only 1540 (Fig. 2).

Such a large difference was expected since the KO mice do not express any functional PPAR α . The Global Contrast, incorporating expression information for all probesets in all groups, identified 4282 significantly regulated probesets ($P < 0.05$) (Fig. 2), representing genes that from an inferential perspective are differentially regulated by WY between the two mouse strains. However, these included genes that for example were only regulated in the KO mice, or were regulated in wild type and, although to a lesser extent, still in KO mice. To filter out these 'biological irrelevant' genes, only probesets that were common in Contrast 1 and Global Contrast were retained, resulting in a set of 3345 probesets, which was called set X. Thus, this set X contained only probesets that from a biological perspective fulfill the criterion of being candidate PPAR α target genes. To correct for multiple testing, FDR values of the 3345 genes in X were calculated based on all 16392 genes in Global Contrast, since in this comparison statistical inference was simultaneously adjusted for both experiments in wild type and knockout mice. Finally, a robust set of PPAR α target genes was obtained by selecting those 2432 probesets from set X that fulfilled the criterion $FDR < 0.05$ (Fig. 2). Of these, 1325 probesets were induced and 1107 probesets were suppressed.

For comparison, we also generated a list of candidate PPAR α target genes that were generated on the basis of directly comparing the wild type and knockout experiment using a FDR cutoff (Fig. 3). Note that this frequently used approach is criticized¹⁵ and that it is in essence identical to the analysis strategy published and interpreted by Rakhshandehroo et al¹⁶, except that these authors also employed a fold change cutoff. Using a FDR cutoff of 0.05, we identified 4324 probesets that were regulated in the wild type experiment (Contrast 1), whereas 12 probesets were

changed in the knockout experiment (Contrast 2). Of these 12 probesets, 10 were also regulated in the wild type experiment. Thus, the FDR based comparison of these two experiments identified 4314 probesets that should be considered PPAR α target genes.

The number of FDR based selected probesets was about twice as large as the list of probesets obtained using our integrated approach (4314 versus 2432 probesets). Comparison of these two sets of candidate genes revealed that almost all (i.e. 99%) of the probesets obtained by our integrated approach were also identified when using a FDR cutoff (Fig. 4). This indicates that while Global Contrast is more conservative it will identify similar if not identical biological features (see also section on validation).

It is important to realize that the results of statistical hypothesis testing are never free of error. Two types of error are distinguished: type I error, i.e. rejecting the null hypothesis when it is in fact true, and type II error, i.e. not rejecting the null hypothesis when in fact the alternative hypothesis is true. In other words, occurrence of the former leads to inclusion of false positives whereas the latter leads to inclusion of false negatives. Consequently, we cannot exclude that the set of 1911 probesets that were discarded by Global Contrast contained false negatives that otherwise would have been retained. However, especially within the context of genome-wide screening studies

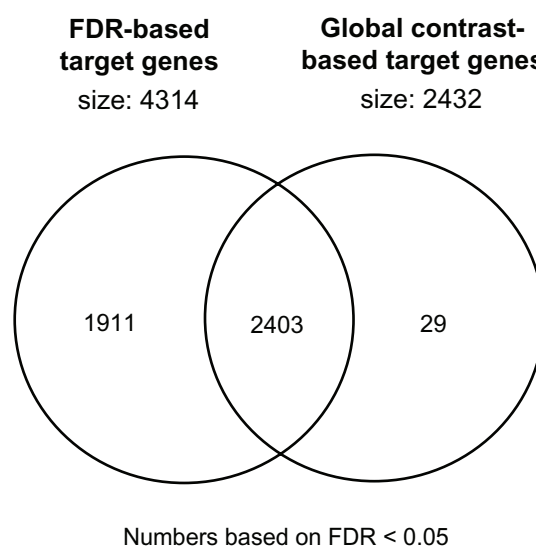


Figure 4. Venn diagram of the identified candidate PPAR α target genes obtained by our integrated approach or the FDR based across experiment comparison of PPAR α dataset GSE8295.

Note: Almost all (99%) of the candidate target genes identified by our proposed approach were also identified in the FDR based across experiment comparison.

for candidate genes, we believe that limiting type I error is of primary concern, and that of type II error is of secondary importance. Thus, to err on the safe side we prefer to control for false positives rather than for false negatives. Moreover, the probesets that were discarded by Global Contrast were characterized by a relatively low effect size compared to the probesets that were still included. The mean of the absolute coefficients (\log_2 of the fold-change) of the excluded probesets was 0.36 (equating to a mean fold change of 1.28), and was 0.87 (mean FC = 1.83) for the included probesets. Taken together, we showed that compared to the FDR based across experiment comparison our approach identified a conservative set of more robustly regulated candidate PPAR α target genes. We believe this is advantageous because a clear overview of candidate genes and corresponding biological process normally is aimed for.

Validation

To compare the performance of our integrated approach with that of the FDR based across experiments comparison, we first performed a sensitivity versus specificity analysis. To this end two benchmark sets of well-established PPAR α target genes were selected from a review that summarized the latest literature on this topic.³² We created two benchmark sets; one set containing only 32 genes, and another set containing 189 genes. The smaller benchmark set contained only genes that were demonstrated to be bona

fide PPAR α target genes in both human and mouse liver and that do contain a functional PPAR response element (PPRE) in the regulatory regions. The larger benchmark set contained all genes that were demonstrated to be PPAR α -dependently regulated in mouse liver but for which no functional PPRE has yet been identified. We next plotted the true positive rate (sensitivity) as function of the false positive rate (1-specificity) for different cutoff points for both our integrated approach and the across experiment comparison (Fig. 5).

Even though our approach identified a conservative list of candidate genes, we observed that it performed very similar to the across experiment comparison in identifying known PPAR α target genes, which was also reflected by almost identical partial area under the ROC curve (pAUC; $P = 0.2$) for both methods. Values were 0.129 and 0.121, respectively for the across experiment comparison and our integrated approach when the smaller set of 32 PPAR α target genes was used, whereas these numbers were 0.128 and 0.124, respectively, for the larger set of 189 putative PPAR α target genes.

Next we detected and compared the biological features that were overrepresented in the lists of candidate PPAR α target genes that were either generated by our approach or the across experiment comparison. Enriched biological processes were identified by overrepresentation analysis based on Gene Ontology (GO) categories, which is a generally accepted procedure to

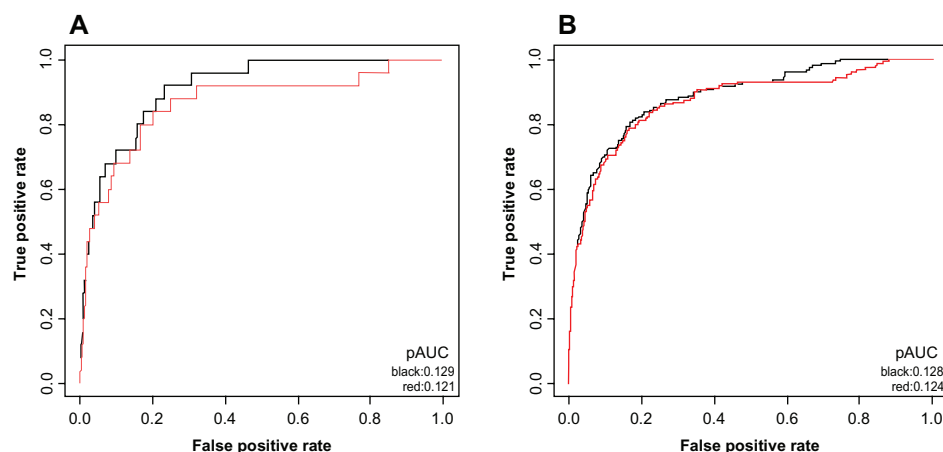


Figure 5. Sensitivity versus specificity of our proposed method and the across experiment comparison.

Notes: The sensitivity versus specificity was analyzed using two benchmark lists of established PPAR α target genes derived from literature. **(Panel A)** ROC curve for both methods using a set of 32 benchmark genes that were demonstrated to be PPAR α target genes in both human and mouse liver and that do contain a functional PPAR response element (PPRE) in the regulatory regions. **(Panel B)** ROC curve for both methods using a set of 189 benchmark genes that were demonstrated to be PPAR α -dependently regulated in mouse liver but for which no functional PPRE has been identified yet. Red lines: ROC curves of our integrated approach; Black lines: ROC curves of the across experiment comparison.



A							B								
GO ID	Name	NSP	P-Value	Adj. P-Value	Rank	Pop. Count	Study Count	GO ID	Name	NSP	P-Value	Adj. P-Value	Rank	Pop. Count	Study Count
GO:000152	metabolic process	B	4.46e-42	7.67e-41	1	6526	2105	GO:000152	metabolic process	B	6.37e-23	1.09e-20	1	6526	1199
GO:0042180	cellular ketone metabolic process	B	3.41e-22	3.09e-20	2	519	265	GO:0042180	cellular ketone metabolic process	B	3.95e-18	3.36e-16	2	519	170
GO:0006082	organic acid metabolic process	B	3.28e-20	1.88e-18	3	509	257	GO:0006082	organic acid metabolic process	B	1.05e-16	5.36e-15	3	509	164
GO:0051186	cofactor metabolic process	B	1.24e-18	5.35e-17	4	191	119	GO:0051186	cofactor metabolic process	B	7.07e-16	2.99e-14	4	191	81
GO:0006629	lipid metabolic process	B	1.93e-13	6.63e-12	5	681	295	GO:0006629	lipid metabolic process	B	6.44e-11	1.93e-09	5	681	182
GO:0009987	cellular process	B	1.43e-12	4.10e-11	6	9424	2694	GO:0009308	amine metabolic process	B	6.84e-11	1.93e-09	6	308	91
GO:0044237	cellular metabolic process	B	4.23e-12	1.04e-10	7	5651	1798	GO:0009987	cellular process	B	1.52e-09	3.67e-08	7	9424	1547
GO:0009308	amine metabolic process	B	8.18e-12	1.76e-10	8	308	141	GO:0006091	generation of precursor metabolites and energy	B	6.75e-08	1.09e-06	8	230	74
GO:0006091	generation of precursor metabolites and energy	B	1.27e-10	2.43e-09	9	230	119	GO:0006519	cellular amino acid and derivative metabolic pro...	B	5.79e-08	1.09e-06	9	281	87
GO:0006519	cellular amino acid and derivative metabolic pro...	B	2.29e-09	3.94e-08	10	281	136	GO:0051234	establishment of localization	B	8.79e-08	1.49e-06	10	2362	442
GO:0051234	establishment of localization	B	9.54e-09	1.49e-07	11	2362	742	GO:0044255	cellular lipid metabolic process	B	1.90e-07	2.93e-06	11	467	127
GO:0015031	protein transport	B	1.26e-07	1.83e-06	12	660	260	GO:0005975	carbohydrate metabolic process	B	4.04e-07	5.68e-06	12	417	112
GO:0005975	carbohydrate metabolic process	B	3.18e-07	4.20e-06	13	85	39	GO:0051179	localization	B	2.59e-06	3.37e-05	13	2714	468
GO:0044255	cellular lipid metabolic process	B	1.86e-06	2.28e-05	14	467	196	GO:0050817	coagulation	B	0.000154	0.00185	14	85	23
GO:0051179	localization	B	3.46e-06	3.97e-05	15	2714	818	GO:0006518	peptide metabolic process	B	0.000237	0.00267	15	40	17
GO:0005975	carbohydrate metabolic process	B	9.23e-06	9.92e-05	16	417	169	GO:0050896	response to stimulus	B	0.000276	0.00292	16	2159	380
GO:0006396	RNA processing	B	5.08e-05	0.000514	17	367	116	GO:0044237	cellular metabolic process	B	0.000311	0.00309	17	5651	1009
GO:0005878	regulation of body fluid levels	B	6.58e-05	0.000629	18	123	45	GO:0006725	cellular aromatic compound metabolic process	B	0.000601	0.00564	18	132	39
GO:0006869	lipid transport	B	0.000...	0.00566	19	116	53	GO:0048771	tissue remodeling	B	0.00152	0.0136	19	70	18
GO:0007031	peroxisome organization	B	0.00273	0.0235	20	16	11	GO:0050878	regulation of body fluid levels	B	0.00162	0.0137	20	123	27
GO:0006725	cellular aromatic compound metabolic process	B	0.00362	0.0297	21	132	57	GO:0009225	nucleotide-sugar metabolic process	B	0.00264	0.0205	21	15	7
GO:0006996	organelle organization	B	0.00461	0.0361	22	1020	328	GO:0015031	protein transport	B	0.00267	0.0205	22	660	148
GO:0050896	response to stimulus	B	0.00482	0.0361	23	2159	625	GO:0006869	lipid transport	B	0.00302	0.0222	23	116	34
GO:0006805	xenobiotic metabolic process	B	0.00533	0.0382	24	17	11	GO:0007005	mitochondrion organization	B	0.00330	0.0232	24	92	28
GO:0005976	polysaccharide metabolic process	B	0.00607	0.0405	25	90	37	GO:0005976	polysaccharide metabolic process	B	0.00389	0.0283	25	90	24
GO:0007005	mitochondrion organization	B	0.00611	0.0405	26	92	41	GO:0006396	RNA processing	B	0.00410	0.0286	26	367	60
GO:0009225	nucleotide-sugar metabolic process	B	0.00668	0.0425	27	15	9	GO:0007031	peroxisome organization	B	0.00443	0.0274	27	16	8
GO:0016043	cellular component organization	B	0.00862	0.0530	28	1981	572	GO:0006081	cellular aldehyde metabolic process	B	0.00454	0.0274	28	20	9
GO:0051604	protein maturation	B	0.00980	0.0581	29	92	40	GO:0006996	organelle organization	B	0.0181	0.105	29	1020	192
GO:0007033	vacuole organization	B	0.0120	0.0660	30	30	16	GO:0016043	cellular component organization	B	0.0196	0.111	30	1981	330
GO:0006518	peptide metabolic process	B	0.0124	0.0660	31	40	20	GO:0006805	xenobiotic metabolic process	B	0.0210	0.114	31	17	7
GO:0015893	drug transport	B	0.0124	0.0660	32	21	12	GO:0006508	proteolysis	B	0.0249	0.132	32	510	106
GO:0065003	macromolecular complex assembly	B	0.0127	0.0660	33	391	135	GO:0051604	protein maturation	B	0.0269	0.138	33	92	24
GO:0048771	tissue remodeling	B	0.0171	0.0867	34	70	23	GO:0006818	hydrogen transport	B	0.0452	0.224	34	78	21
GO:0051301	cell division	B	0.0200	0.0984	35	312	106	GO:0006323	DNA packaging	B	0.0472	0.228	35	58	15
GO:0006081	cellular aldehyde metabolic process	B	0.0265	0.127	36	20	11	GO:0016044	cellular membrane organization	B	0.0532	0.250	36	305	61
GO:0018116	cytokine production	B	0.0309	0.143	37	162	45	GO:0046483	heterocycle metabolic process	B	0.0588	0.269	37	687	138
GO:0051640	organelle localization	B	0.0326	0.147	38	51	22	GO:0006300	oxygen and reactive oxygen species metabolic...	B	0.0631	0.281	38	38	11
GO:0007059	chromosome segregation	B	0.0446	0.197	39	52	21	GO:0051301	cell division	B	0.0768	0.333	39	312	61
GO:0048285	organelle fission	B	0.0501	0.215	40	210	78	GO:0015893	drug transport	B	0.0806	0.341	40	21	7

Figure 6. Significantly enriched Gene Ontology categories found in the two lists of candidate PPAR α target genes. **Notes:** Enriched biological processes were identified in the two lists of candidate PPAR α targets genes generated by the across experiment comparison (Panel A), or our integrated approach (Panel B). All significant probesets identified by the respective methodologies were used as input. The 'parent-child-union' algorithm was applied followed by the Benjamini-Hochberg correction for multiple testing to identify enriched GO categories. In both lists the same underlying biology was identified. **Abbreviations:** NSP, name space (sub ontology); B, Biological process.

achieve this.¹⁴ To this end the software tool Ontologizer was used,³⁵ applying the 'parent-child-union' (PCU) algorithm. This algorithm takes the graph structure of GO into account, thereby reducing false-positive and biologically misleading results.³⁶ We ran the PCU algorithm on the biological process ontology of GO. Using the Benjamini-Hochberg correction for multiple testing and a cutoff of 0.05, 27 significantly enriched categories (out of 172 annotated categories) were returned in the list of 4314 putative PPAR α target genes obtained by comparing the wild type and knockout experiment (Fig. 6A).

Similarly, using the same criteria 28 significantly enriched categories (out of 169 annotated categories) were scored in the list of 2432 genes generated by our integrated approach (Fig. 6B). Twenty-five identified enriched biological processes were identical in both sets of genes. As expected, many processes that were enriched have been functionally demonstrated to be controlled by PPAR α , including cellular ketone metabolic process, lipid metabolic process, cellular amino acid and derivative metabolic process, peroxisome organization, and mitochondrion organization.^{17,37,38}

Thus, despite the drastically reduced number of candidate PPAR α target genes identified by our approach, GO enrichment analysis demonstrated a very similar functional characterization of these genes, again demonstrating the validity of our strategy.

Conclusions

Taken together, in this study we present a simple, yet efficient strategy to compare genes across experiments that controls for multiple testing and is able to properly detect differentially expressed genes. Compared to the conventional used FDR based across experiment comparison, our approach is more conservative and less noisy. Our approach is in particular suitable to identify candidate target genes of a transcription factor or signaling route from functional genomics experiments, but can be applied to any genomics experiment in which the effects of a treatment are compared between two genotypes.

Funding

This work was partially supported by the Wageningen University program for Systems Biology to MM (IP/OP SysBio).

Competing Interests

The authors disclose no potential conflicts of interest.

Author Contributions

Conceived and designed the experiments: MOU, MM, GH. Analysed the data: MOU, GH. Wrote the first draft of the manuscript: MOU. Made critical revisions and approved final version: MM, GH. All authors reviewed and approved of the final manuscript.

Disclosures and Ethics

As a requirement of publication author(s) have provided to the publisher signed confirmation of compliance with legal and ethical obligations including but not limited to the following: authorship and contributorship, conflicts of interest, privacy and confidentiality and (where applicable) protection of human and animal research subjects. The authors have read and confirmed their agreement with the ICMJE authorship and conflict of interest criteria. The authors have also confirmed that this article is unique and not under consideration or published in any other publication, and that they have permission from rights holders to reproduce any copyrighted material. Any disclosures are made in this section. The external blind peer reviewers report no conflicts of interest.

References

1. Brown PO, Botstein D. Exploring the new world of the genome with DNA microarrays. *Nat Genet.* 1999;21(Suppl 1):33–7.
2. Stuart JM, Segal E, Koller D, Kim SK. A gene-coexpression network for global discovery of conserved genetic modules. *Science.* 2003;302(5643):249–55.
3. Segal E, Friedman N, Kaminski N, Regev A, Koller D. From signatures to models: understanding cancer using microarrays. *Nat Genet.* 2005; 37 Suppl:S38–45.
4. Quackenbush J. Microarray analysis and tumor classification. *N Engl J Med.* 2006;354(23):2463–72.
5. Muller M, Kersten S. Nutrigenomics: goals and strategies. *Nat Rev Genet.* 2003;4(4):315–22.
6. Kirmizis A, Farnham PJ. Genomic approaches that aid in the identification of transcription factor target genes. *Exp Biol Med.* 2004;229(8):705–21.
7. Gregory BD, Belostotsky DA. Whole-genome microarrays: applications and technical issues. *Methods Mol Biol.* 2009;553:39–56.
8. Babu MM, Lang B, Aravind L. Methods to reconstruct and compare transcriptional regulatory networks. *Methods Mol Biol.* 2009;541:163–80.
9. Brien GL, Bracken AP. Transcriptomics: unravelling the biology of transcription factors and chromatin remodelers during development and differentiation. *Semin Cell Dev Biol.* 2009;20(7):835–41.
10. Petricka JJ, Benfey PN. Reconstructing regulatory network transitions. *Trends Cell Biol.* 2011;21(8):442–51.
11. Gorte M, Horstman A, Page RB, et al. Microarray-based identification of transcription factor target genes. *Methods Mol. Biol.* 2011;754:119–41.
12. Bunger M, Hooiveld GJ, Kersten S, Muller M. Exploration of PPAR functions by microarray technology—a paradigm for nutrigenomics. *Biochim Biophys Acta.* 2007;1771(8):1046–64.
13. Woods CG, Heuvel JP, Rusyn I. Genomic profiling in nuclear receptor-mediated toxicity. *Toxicol. Pathol.* 2007;35(4):474–94.
14. Allison DB, Cui X, Page GP, Sabripour M. Microarray data analysis: from disarray to consolidation and consensus. *Nature Reviews Genetics.* 2006; 7(1):55–65.
15. Higdon R, van Belle G, Kolker E. A note on the false discovery rate and inconsistent comparisons between experiments. *Bioinformatics.* 2008; 24(10):1225–8.
16. Rakhshandehroo M, Sanderson LM, Matilainen M, et al. Comprehensive analysis of PPARalpha-dependent regulation of hepatic lipid metabolism by expression profiling. *PPAR Res.* 2007;2007:26839.
17. Desvergne B, Wahli W. Peroxisome proliferator-activated receptors: nuclear control of metabolism. *Endocr Rev.* 1999;20(5):649–88.
18. Sampath H, Ntambi JM. Polyunsaturated fatty acid regulation of gene expression. *Nutr Rev.* 2004;62(9):333–9.
19. Germain P, Staels B, Dacquet C, Spedding M, Laudet V. Overview of nomenclature of nuclear receptors. *Pharmacol Rev.* 2006;58(4):685–704.
20. Lee SS, Pineau T, Drago J, et al. Targeted disruption of the alpha isoform of the peroxisome proliferator-activated receptor gene in mice results in abolishment of the pleiotropic effects of peroxisome proliferators. *Mol Cell Biol.* 1995;15(6):3012–22.
21. Issemann I, Green S. Activation of a member of the steroid hormone receptor superfamily by peroxisome proliferators. *Nature.* 1990;347(6294):645–50.
22. Willson TM, Brown PJ, Sternbach DD, Henke BR. The PPARs: from orphan receptors to drug discovery. *J Med Chem.* 2000;43(4):527–50.
23. Kerr MK, Martin M, Churchill GA. Analysis of variance for gene expression microarray data. *J Comput Biol.* 2000;7(6):819–37.
24. Ayroles JF, Gibson G. Analysis of variance of microarray data. *Methods Enzymol.* 2006;411:214–33.
25. Kutner MH, Nachtsheim CJ, Neter J, Li W. *Applied Linear Statistical Models.* New York: McGraw-Hill, 2005.
26. Ihaka R, Gentleman RR: a language for data analysis and graphics. *J Comput Graph Stat.* 1996;5:299–314.
27. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* 2004;5(10):R80.
28. Dai M, Wang P, Boyd AD, et al. Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.* 2005;33(20):e175.
29. Wu Z, Irizarry RA, Gentleman R, Martinez-Murillo F, Spencer F. A model-based background adjustment for oligonucleotide expression arrays. *J Am Stat Assoc.* 2004;99:909–17.
30. Smyth G. limma: Linear Models for Microarray Data. In: *Bioinformatics and Computational Biology Solutions Using R and Bioconductor.* Edited by Gentleman R, et al. New York: Springer New York; 2005:397–420.
31. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Royal Stat Soc Ser B.* 1995;57:289–300.
32. Rakhshandehroo M, Knoch B, Muller M, Kersten S. Peroxisome proliferator-activated receptor alpha target genes. *PPAR Res.* 2010.
33. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCr: visualizing classifier performance in R. *Bioinformatics.* 2005;21(20):3940–1.
34. Huber W, Scholtens D, Hahne F, Heydebreck A. Differential Expression. In: *Bioconductor Case Studies.* Edited by Hahne F, et al. New York: Springer New York; 2008:89–102.
35. Bauer S, Grossmann S, Vingron M, Robinson PN. Ontologizer 2.0—a multifunctional tool for GO term enrichment analysis and data exploration. *Bioinformatics.* 2008;24(14):1650–1.
36. Grossmann S, Bauer S, Robinson PN, Vingron M. Improved detection of overrepresentation of Gene-Ontology annotations with parent child analysis. *Bioinformatics.* 2007;23(22):3024–31.
37. Kersten S, Mandard S, Escher P, et al. The peroxisome proliferator-activated receptor alpha regulates amino acid metabolism. *FASEB J.* 2001;15(11):1971–8.
38. Mandard S, Muller M, Kersten S. Peroxisome proliferator-activated receptor alpha target genes. *Cell Mol Life Sci.* 2004;61(4):393–416.



Supplementary Materials

Supplementary 1. R code to perform the described analysis.

Supplementary 2. Sample description file for use with limma.

Supplementary 3. Excel file containing the two benchmark gene sets.