

Biotransformation pathway maps in WikiPathways enable direct visualization of drug metabolism related expression changes.

Citation for published version (APA):

Jennen, D. G. J., Gaj, S., Giesbertz, P. J., van Delft, J. H. M., Evelo, C. T. A., & Kleinjans, J. C. S. (2010). Biotransformation pathway maps in WikiPathways enable direct visualization of drug metabolism related expression changes. *Drug Discovery Today*, 15(19-20), 851-858. <https://doi.org/10.1016/j.drudis.2010.08.002>

Document status and date:

Published: 01/01/2010

DOI:

[10.1016/j.drudis.2010.08.002](https://doi.org/10.1016/j.drudis.2010.08.002)

Document Version:

Publisher's PDF, also known as Version of record

Document license:

Taverne

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.



Biotransformation pathway maps in WikiPathways enable direct visualization of drug metabolism related expression changes

Danyel G.J. Jennen^{1,2}, Stan Gaj^{1,2}, Pieter J. Giesbertz¹, Joost H.M. van Delft^{1,2}, Chris T. Evelo^{2,3} and Jos C.S. Kleinjans^{1,2}

¹ Department of Health Risk Analysis and Toxicology, Maastricht University, Universiteitssingel 50, 6229 ER Maastricht, The Netherlands

² Netherlands Toxicogenomics Centre, Universiteitssingel 50, 6229 ER Maastricht, The Netherlands

³ Department of Bioinformatics, BiGCaT, Maastricht University, Universiteitssingel 50, 6229 ER Maastricht, The Netherlands

In recent decades, our knowledge of the genetics and functional genomics of drug-metabolizing enzymes has increased and a wealth of data on drug-related ‘omics’ has become available. Despite the availability of large amounts of biological information on xenobiotic biotransformation, the number of available biotransformation pathway maps that can easily be used for visualization of multiple omics data is limited. Here, we created integrated biotransformation pathway maps suitable for multiple omics analysis using PathVisio. The ease of visualizing data on these maps was demonstrated by using published microarray data from human hepatocyte-like cell models, exemplifying – where a sufficient capacity for metabolizing chemicals is a prerequisite for a suited model – how the biotransformation pathway maps can be used for model selection.

Introduction

Over time, drug metabolism has become more and more important in pharmaceutical research on drug discovery and development [1,2]. Where drug metabolism traditionally investigated the well-defined aspects of absorption, distribution, metabolism and excretion (also known as ADME), its focus has shifted towards areas on the genomics and genetics levels, aiding the early discovery or prediction of adverse effects of new drugs [3]. Advances from the past decade in fields such as pharmacogenetics, pharmacogenomics and toxicogenomics have increased our knowledge of the genetics and genomics of drug-metabolizing enzymes (DMEs), resulting – for example – in new insights in induction and inhibition, substrate specificities and polymorphisms of DMEs [4–6]. This information is useful in the development of novel *in vitro* cell models for the purpose of screening drug candidates for their efficacy and safety because capacity to metabolize chemicals is a prerequisite for such models [7]. For instance, at present, there is an increasing interest in the development of stem-cell-derived models, such as hepatocyte-like cells [8,9], and the metabolic

competence of such novel models is considered of utmost relevance [10].

Although a large amount of information on the biotransformation reactions is available in literature (e.g. Refs. [11–15]) and online pathway database resources, such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) (<http://www.genome.jp/kegg/>) [16] and Reactome (<http://www.reactome.org/>) [17], the number of online biotransformation pathway maps suited for evaluating the metabolic competence of a cell model is limited. Furthermore, these maps cannot easily be updated or used for visualization of transcriptomics, proteomics and metabolomics data. This paper, therefore, will focus not only on the different biotransformation pathways and their availability from different pathway databases but also on their applicability in data analysis and visualization.

Recently, biotransformation pathway maps were constructed using PathVisio (<http://www.pathvisio.org/>) [18], the pathway editor of WikiPathways (<http://www.wikipathways.org/>) [19], and subsequently made available to the community at WikiPathways [20]. The ease of visualizing data onto these maps in PathVisio will be demonstrated by using previously published microarray data on baseline gene expression in important human tissues for biotransformation (i.e. liver, expressing most DMEs, kidney and

Corresponding author: Jennen, Danyel G.J. (danyeljennen@grat.unimaas.nl)

TABLE 1

Overview of availability^a of biotransformation pathways from some pathway databases

	WikiPathways ^b	MetaCore ^c	IPA ^d	Reactome ^e	KEGG ^f	SMPDB ^g	MetaCyc ^h	BioCarta ⁱ
Phase I biotransformation pathways								
Cytochrome P450	FP	CP	CP	R	CP	–	R	–
Flavin-monooxygenase catalytic cycle	FP	–	–	FP	–	–	R	–
Aldo-keto reductase pathway	CP	–	–	R	–	–	R	–
Epoxide hydrolase pathway	CP	R	CP	–	CP	–	R	–
Phase II biotransformation pathways								
Glutathione conjugation	FP	FP	FP	R	FP	FP	R	–
Amino acid conjugation	CP	–	–	R	–	–	R	–
Sulfation/sulfonation	FP	CP	CP	R	–	–	R	–
Acetylation	FP	CP	–	R	–	–	R	–
Glucuronidation	FP	CP	CP	FP	–	–	R	–
Methylation	FP	CP	CP	FP	–	–	R	–

^a The biotransformation pathways are available as FP, full pathway containing all available reaction information; CP, compound-related pathway containing compound specific reaction information; R, reaction showing information of single reactions.

^b <http://www.wikipathways.org/> [19].

^c <http://www.genego.com/>.

^d <http://www.ingenuity.com/>.

^e <http://www.reactome.org/> [17].

^f <http://www.genome.jp/kegg/> [16].

^g <http://www.smpdb.ca/> [24].

^h <http://www.metacyc.org/> [25].

ⁱ <http://www.biocarta.com/genes/index.asp>.

lung) [21]. In addition, case studies on several human hepatocyte-like cell models are presented to exemplify how the biotransformation pathway maps can be used for model selection.

Biotransformation pathways and databases

Biotransformation can be divided into two main phases, phase I and phase II. Phase I biotransformation reactions include oxidation, reduction, hydrolysis, hydration and other relatively rare reactions that cause the introduction of reactive and polar functional groups in the compound, making it more suitable for conjugation reactions. Phase I biotransformation thus functions as a preparation step for phase II biotransformation. In the phase II biotransformation reactions, metabolites are conjugated with (small) endogenous molecules, often resulting in water-soluble metabolites that can be further metabolized or easily excreted from the human body [10,15,21,22]. This further metabolization or excretion is part of the phase III biotransformation reactions [10,15], which are important reactions in the overall process of biotransformation. Phase III biotransformation, however, will not be considered in this paper. In Table 1, pathways of the phase I and phase II biotransformation reactions are listed with their availability from two commercial and six freely available online pathway resources, selected from the online pathway resource list Pathguide (<http://www.pathguide.org/>) [23]. Although Pathguide contains more than 300 resources, each with a short description and a link to the resource homepage, only a few popular pathway resources were selected that contain proprietary pathway maps. Other resources, either commercial or freely available, such as PathArt (Jubilant Biosys Ltd., Bangalore, India; <http://www.jubilantbiosys.com/pathart.html>), Pathway Interaction Database (<http://pid.nci.nih.gov/>) or PANTHER (<http://www.pantherdb.org/>) make use of the resources listed in Table 1 (e.g. KEGG, BioCarta or Reactome). Therefore, we will only focus on the usability of the eight selected pathway databases and their analysis tools with respect to the biotransformation pathways.

From the examined databases, MetaCore (GeneGo, San Diego, CA; <http://www.genego.com/>) and Ingenuity Pathway Analysis (IPA) (Ingenuity Systems, Redwood City, CA, <http://www.ingenuity.com/>) require a license, whereas the other databases are freely available. Six of these pathway resources contain tools that enable visualization of expression data. This is an important step during data analysis because it facilitates a better biological interpretation by directly comparing different conditions and/or concentrations in relation to a specific biological context. WikiPathways, MetaCore, IPA and Reactome provide tools for statistical pathway ranking tests to identify significantly altered pathways by determining whether the changes of the elements (genes, proteins or metabolites) of a given pathway are higher than the average change in the complete dataset (*P*-value or *Z*-score). WikiPathways, Metacore and IPA perform better here because they enable direct visualization of multiple datasets. Data visualization in Reactome is more limited because it is restricted to a single dataset and visualization results in a complex image, making it more difficult to interpret the data. KEGG and The Small Molecule Pathway Database (SMPDB) (<http://www.smpdb.ca/>) [24] are also only capable of visualizing a single dataset but do not perform any statistical pathway ranking tests. BioCarta (<http://www.biocarta.com/genes/index.asp>) provides static images, whereas MetaCyc (<http://www.metacyc.org/>) [25] provides links for each element in its pathways and reactions. The content of the different pathway resources varies from no or hardly any biotransformation pathway (i.e. BioCarta, SMPDB and KEGG), through reactions (i.e. Reactome and MetaCyc), to compound-related and full pathways (i.e. WikiPathways, MetaCore and IPA).

IPA, MetaCore and WikiPathways perform equally well in visualizing expression data, and the latter two have shown to be an asset to each other [26]. In addition, IPA, MetaCore and WikiPathways provide tools to create and edit one's own pathway maps. For MetaCore, however, an extension of the license with the MapEditor software is needed. Furthermore, the maps created

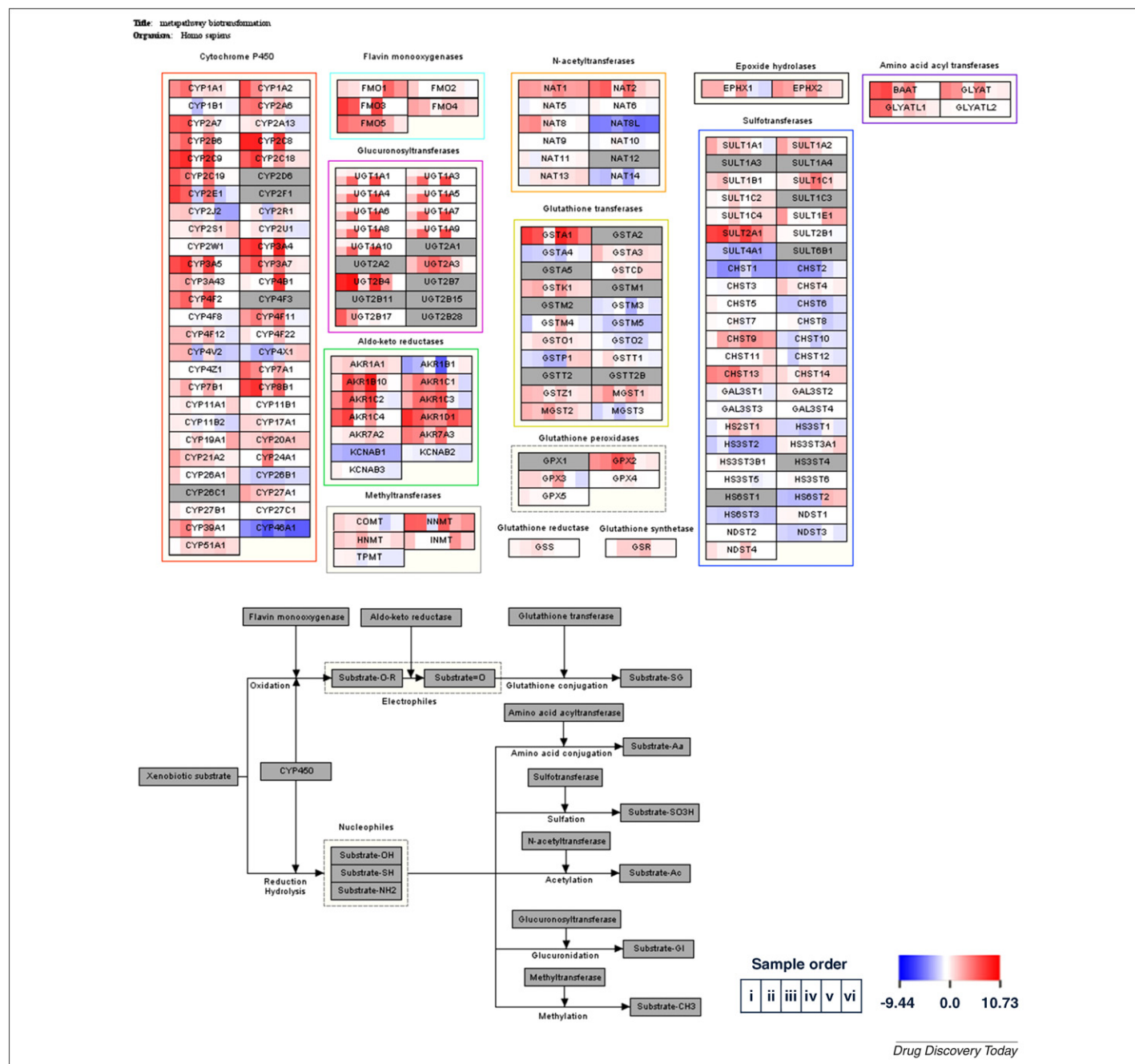


FIGURE 2 Meta pathway for biotransformation showing basal gene expression data from (i) human liver tissue, (ii) primary human hepatocytes, (iii) HepG2, (iv) HepaRG, (v) hepatocyte-like cells derived from human embryonic stem cells and (vi) hepatocyte-like cells derived from human induced pluripotent stem cells. The expression data consist of log2 differences between each liver cell type and the common reference sample from human brain, whereby overexpression is shown in red and underexpression in blue. Gray indicates that no data was available.

in MetaCore and IPA can only be added to the local pathway database.

Thus, based on the comparison between the different pathway resources and their tools, WikiPathways was chosen for the further analyses and evaluations presented in this paper because WikiPathways provides the tools to construct and edit pathway maps, to visualize expression data for multiple ‘omics’ and to perform statistical pathway ranking tests and because it is open to the community.

Pathway development in WikiPathways

Human phase I and phase II biotransformation pathway maps were constructed using PathVisio [18], the pathway editor of WikiPathways [19]. PathVisio is a tool for editing and displaying biological pathways. It provides a basic palette of objects and annotations that represent specific biological processes. Genes, proteins and metabolites can be directly mapped to biological annotations from multiple public databases through the identifier synonym database at WikiPathways using the

BridgeDb identifier mapping framework (<http://www.bridgedb.org/>) [27].

The pathway content of the human phase I and phase II biotransformation pathway maps in various target organs was generated using information obtained from several resources: pharmacological and/or toxicological literature; online biological resources listed in Table 1, such as KEGG [16] and Reactome [17]; and functional databases, such as the gene database GeneCards (<http://www.genecards.org/>) [28] and the enzyme database BRENDA (<http://www.brenda-enzymes.org/>) [29].

Most of the biotransformation pathway maps in WikiPathways have been approved by GenMAPP (<http://www.genmapp.org/>) [30] and will be included in the curated pathway archive of GenMAPP. As an example, the small pathway map of the catalytic cycle of mammalian flavin-monooxygenases (FMOs) is shown in Figure 1a. A pathway description and literature references are attached to this pathway. Furthermore, each element, or 'Data-Node', in the pathway is linked to an identifier. As such, each FMO is directly linked to an Ensembl gene identifier [31]. In addition, cross-references to other gene identifiers and protein identifiers are connected using an Ensembl-based derby database mapping in the BridgeDb framework. Furthermore, metabolites are linked to a unique identifier, as indicated in Figure 1a. This provides the possibility of mapping not only gene expression data to this and other biotransformation pathway maps but also any other gene-based data set – such as proteomics, DNA methylation and metabolomics data – thereby enabling multiple omics integration. In total, the biotransformation pathway maps contain 317 Data-Nodes representing 187 genes or proteins and 130 metabolites. Next, the biotransformation pathway maps were combined and integrated into the meta pathway for biotransformation (Figures 1b and 2), which shows all the genes or proteins involved in biotransformation. Although the meta pathway can be used for visualization purposes, it is primarily intended for statistical evaluation.

Microarray data pre-processing

For visualization purposes, microarray data from five data sets were obtained from the online microarray data repository Gene Expression Omnibus (GEO) of the US National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/geo/>) [32] and from the microarray data repository ArrayExpress of the European Bioinformatics Institute (<http://www.ebi.ac.uk/microarray-as/ae/>) [33].

(i) GEO accession GSE3526 contains baseline gene expression data from normal human kidney cortex and medulla, liver and lung (collected post-mortem) obtained from three or four donors (GSM80686 to GSM80689 for kidney cortex, GSM80731 to GSM80734 for kidney medulla, GSM80728 to GSM80730 and GSM80739 for liver, and GSM80710, GSM80707 and GSM80712 for lung) [34]. (ii) GEO accession GSE14897 contains data from hepatocyte-like cells derived from three independent cultures of human embryonic stem cells (GSM372147 to GSM372149) and from hepatocyte-like cells derived from three independent cultures of human induced pluripotent stem cells (GSM372154 to GSM372156) [9]. (iii) GEO accession GSE11942 contains data from primary human hepatocytes (PHH) obtained from four donors (GSM301603 to GSM301606) [35]. (iv) GEO accession GSE5350

contains expression data of the Ambion First Choice Human Brain Reference RNA collected from one experiment with samples from five replicates used in the MAQC project (GSM122779 to GSM122783) [36]. This data set was used as a common reference for further analysis. (v) ArrayExpress accession E-MEXP-2458 contains data from the hepatoma cell lines HepG2 (48 h solvent samples) and HepaRG (48 h solvent samples), each from three independent cultures [37].

All raw data sets were generated on the same microarray platform (i.e. Affymetrix Human Genome U133 Plus 2.0 GeneChip arrays). These data were re-annotated to the MBNI Custom CDF-files (http://brainarray.mbni.med.umich.edu/Brainarray/Database/CustomCDF/genomic_curated_CDF.asp) [38] and RMA normalized [39] using the NuGOExpressionFileCreator, an enhanced version of the standard ExpressionFileCreator module that is present in GenePattern [40]. The resulting 17 788 probe sets represent 17 726 unique genes and 62 internal controls. For each gene, the ratio between the mean intensity per tissue or cell type and the mean intensity of the common reference was calculated; log₂ transformed and, subsequently, visualized on the biotransformation pathway maps using PathVisio.

Microarray data visualization of human organ tissues

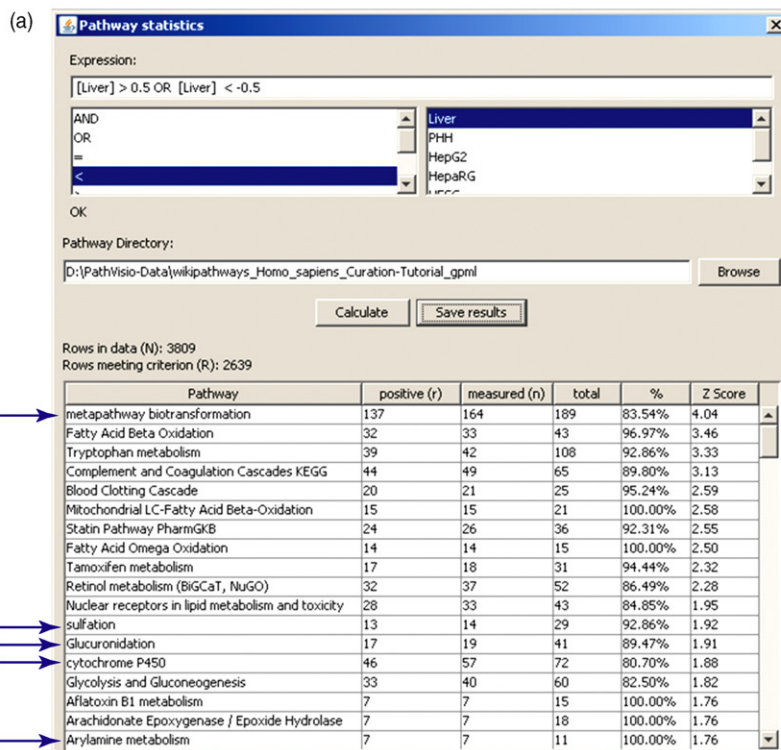
The liver is considered the most important organ in drug metabolism because one of its main functions is to break down and synthesize compounds. Most DMEs are expressed in the liver at relatively high levels [21,41]. The kidney and the lung also are important metabolically active organs that, furthermore, play a part in the excretion of metabolized drugs [41]. Therefore, considering their major role in drug metabolism, liver, kidney and lung were selected for microarray data visualization on the biotransformation pathway maps. The reference microarray data originated from the human brain. Although the brain is also considered to be an important drug-metabolizing organ, it expresses most DMEs at low levels [21].

The log₂ ratios between each tissue sample and the reference brain sample were visualized on the biotransformation pathway maps. For kidney, two distinct anatomical regions were used because these showed unique and highly distinctive patterns of gene expression [42]. Figure 1 shows the baseline expression of the FMO pathway and meta pathway for biotransformation for all tissue samples. FMO1 is highly expressed in the kidney, FMO2 in the lung, and FMO3 and FMO5 in the liver. FMO4 shows the lowest expression of all FMOs, but still the highest expression is found in the kidney, closely followed by the liver. The visualization of the basal gene expression in the investigated tissue samples of the meta pathway for biotransformation shows a clearly higher expression of most biotransformation-related CYP genes in liver. CYP1B1 is an exception, however, because it is hardly expressed in liver, whereas lung demonstrates elevated expression. In addition, CYP1A1 shows higher expression in lung than in liver and kidney. Furthermore, in the other biotransformation pathway maps, differences in basal gene expression from the examined tissues are observed, showing overexpression (mainly in liver) of several phase II genes (e.g. SULT2A1, UGT2B4, AKR1D1 and BAAT).

The baseline expression profiles of each biotransformation gene from the human tissues were compared with those displayed by the BioGPS database from the Genomics Institute of the Novartis

Research Foundation (GNF) (<http://biogps.gnf.org/>) [43]. BioGPS is a gene portal providing gene and protein information from different (online) sources, such as Gene Atlas expression profiles from the GNF [44]. These Gene Atlas basal expression profiles are displayed for 79 human tissue samples per single gene present on the Affymetrix Human Genome U133A GeneChip array. The observations on the biotransformation pathway maps correspond well with the basal expression profiles of each biotransformation gene as displayed by the BioGPS database. Indeed, the expression profiles in BioGPS confirm that FMO1 is specific for kidney, FMO2

for lung and FMO3 and FMO5 for liver, but the expression profile of FMO4 differs slightly from our analysis. In BioGPS, the highest expression is also found for kidney, but the expression of FMO4 is much lower in liver, whereas in our analysis, basal gene expression in kidney is just slightly higher than in liver. With regard to the CYP and phase II genes, baseline expression profiles from our analysis are similar to those from BioGPS. Differences in expression between our analysis and those from BioGPS might be due to differences in tissue samples, as well as microarray analysis approach (e.g. original Affymetrix probe sets versus re-annotated



(b)

log2 ratio of 0.5	Ranking	Positive (r)	Measured (n)	Total	%	Z-Score	Over-expressed genes
Liver	1	137	164	189	83.54%	4.04	110
PHH	1	126	164	189	76.83%	4.17	100
HepG2	27	110	164	189	67.07%	1.52	75
HepaRG	1	128	164	189	78.05%	5.43	96
HESC	118	92	164	189	56.10%	-0.92	55
IPS	83	99	164	189	60.37%	0.22	55
log2 ratio of 2							
Liver	1	76	164	189	46.34%	9.49	69
PHH	1	75	164	189	45.73%	8.61	66
HepG2	11	43	164	189	26.22%	2.83	28
HepaRG	1	76	164	189	46.34%	9.81	63
HESC	92	24	164	189	14.63%	-0.46	17
IPS	113	21	164	189	12.80%	-1.06	13

Drug Discovery Today

FIGURE 3

Pathway ranking test by PathVisio. (a) A screen shot of the ranking output of the normal liver tissue sample with an expression level at an absolute log2 ratio >0.5. Arrows indicate the biotransformation pathways. (b) The ranking of the meta pathway for biotransformation for the different liver cell models with an expression level at an absolute log2 ratio of 0.5 or 2.0. PHH, primary human hepatocytes; HESC, hepatocyte-like cells derived from human embryonic stem cells; IPS, hepatocyte-like cells derived from human induced pluripotent stem cells. Further explanation can be found in the main text of the article.

probe sets). Although in BioGPS, expression profiles are shown for 79 human tissue samples at once, these profiles are only shown for a single gene. In PathVisio, by contrast, expression profiles of multiple genes and their relationship are visualized. It should be noted that BioGPS has plug-ins for pathway databases including WikiPathways [43], showing in which pathway the specific genes are involved; however, in BioGPS, no actual visualization of expression data is possible.

Microarray data visualization of human liver cell models

PHH are considered the most relevant *in vitro* model that resembles the human liver *in situ* [45]. The usability of PHH is limited, however, because of the difficulty of obtaining sufficient donor material and because of the large variability between the donors. Hepatocytes derived from human stem cells, by contrast, are expected to become exceptionally useful as a human *in vitro* system for studying drug metabolism and toxicity [7,8]. Other available *in vitro* models are the frequently used HepG2 [46–48] and the more recently developed HepaRG [49,50], which have both been previously compared to PHH and normal liver tissue [37,48–55].

Here, we investigate the above cell models (i.e. PHH, HepG2, HepaRG and stem-cell-derived hepatocytes) for their baseline expression of the biotransformation genes and compare these with comparable data from normal human liver. In the analysis, hepatocyte-like cells derived from two different types of human progenitor stem cells were used (i.e. H9 human embryonic stem cells and induced pluripotent stem cells obtained from foreskin fibroblasts). Both stem cell lines were cultured using ‘standard’ conditions and after differentiation showed hepatocyte-like characteristics (e.g. several hepatic functions, including accumulation of glycogen, accumulation of lipid, active uptake of low-density lipoprotein and synthesis of urea, as well as several morphological characteristics associated with hepatocytes) [9].

In a similar manner to that used for the human tissue samples, log₂ ratios of each liver cell model sample against the reference brain sample were visualized onto the meta pathway for biotransformation (Figure 2). The expression pattern of the biotransformation genes from the liver sample and primary hepatocytes seem comparable. Also for HepaRG, the expression pattern of the phase I and phase II genes is similar to that of the liver sample, whereas for HepG2, the pattern is different. These findings correspond well with our previous analysis, in which a hierarchical clustering analysis of the basal gene expression showed that PHH and HepaRG are more closely related to liver tissue than HepG2 is [37]. In the two stem-cell-derived hepatocytes, however, only a few biotransformation genes (e.g. FMO4, NAT1, AKR1D1 and several glutathione transferases) show expression that is comparable with those from the normal liver sample and primary hepatocytes. Based on these results, the two stem cell models do not seem to be applicable in research on drug metabolism and toxicity because they lack expression of many of the biotransformation genes. The low expression of these genes might be due to the standard culture conditions used, however, and could be improved under the right conditions [9]—as shown in other studies, in which the expressions of CYP1A1 and CYP3A4 in stem-cell-derived hepatocytes were comparable to those of primary hepatocytes [8].

Pathway ranking

In addition to the direct visualization of the basal expression of the biotransformation genes from the different liver cell models, a statistical pathway ranking test was performed using PathVisio. For this, the performance of each individual liver cell model was evaluated based on the ranking of the biotransformation pathways and in particular the meta pathway for biotransformation. A high ranking of the biotransformation pathways is expected for those models showing an increase in basal expression levels of the phase I and phase II genes. In the statistical pathway ranking test, significantly altered pathways are identified by counting how many genes on each pathway meet user-defined criteria and comparing this to the expected number of genes that meet the criteria to calculate a Z-score. A Z-score >1.65 (1 tail) corresponds with a significant P-value <0.05. As a criterion for the pathway ranking test, an absolute log₂ ratio of each liver cell model sample against the reference brain sample >0.5 or >2 was selected. Figure 3a shows a screenshot of the ranking output of the normal liver tissue sample with an expression level at an absolute log₂ ratio >0.5. The top-ranked pathway is the meta pathway for biotransformation. It contains 164 genes present in the microarray data, of which 137 have an absolute log₂ ratio >0.5. Most of these 137 genes are overexpressed in liver compared to the reference brain sample, which is in agreement with what is known from literature [21,41]. In addition, four more biotransformation pathways had a Z-score >1.65 (Figure 3b).

Further statistical pathway ranking for all liver cell models, using absolute log₂ ratios >0.5 or >2 as criteria, shows that the meta pathway for biotransformation is also ranked number 1 for PHH and HepaRG (Figure 3b). Next in line would be HepG2 because it shows the meta pathway for biotransformation to be significant (Z-score >2.83) for the absolute log₂ ratio >2. For both stem cell models, however, the meta pathway for biotransformation was not significantly ranked.

With respect to the number of overexpressed genes using the above selection criteria (Figure 3b), the liver cell models can be placed in the same order as for the ranking of the meta pathway for biotransformation: liver, PHH, HepaRG > HepG2 > stem cell models.

Concluding remarks and future perspectives

The usefulness of the presented biotransformation pathway maps lies in the easy visualization of the expression of multiple genes and proteins or changed amounts of metabolites in one go. In this article, we clearly illustrated the use of a meta pathway for biotransformation in the analysis and interpretation of microarray data from various human tissues and *in vitro* cell models. In this analysis, the stem cell models underperform and thus further development is needed for these models to become suitable for screening drug candidates. This was also indicated by Si-Tayeb *et al.* [9].

Furthermore, we emphasize that the pathway maps can be used to investigate, for example, drug-induced gene expression changes over time or from different concentrations in an *in vitro* cell model. As the elements in the biotransformation pathway maps contain not only gene identifiers but also protein and metabolite identifiers, their use can be extended towards analysis of proteomics and metabolomics data. Integration of the different omics will help to

improve the understanding of drug metabolism and aid in the investigation of new drugs.

Finally, the biotransformation pathway maps are not just available for human: mouse and rat biotransformation pathway maps have also been developed. This will be useful for inter-species comparisons. Recently, the usage of the rat biotransformation

pathway maps by visualizing tamoxifen and aflatoxin B1 expression data from Iconix Biosciences (Entelos, Foster City, CA) [3] has been presented at the Benelux Bioinformatics Conference (BBC-09, Liège, Belgium, 14–15 December 2009) [20] and is available at Nature Precedings (<http://precedings.nature.com/documents/4575/version/1>).

References

- Testa, B. (2009) Drug metabolism for the perplexed medicinal chemist. *Chem. Biodivers.* 6, 2055–2070
- Yengi, L.G. *et al.* (2007) The evolving role of drug metabolism in drug discovery and development. *Pharm. Res.* 24, 842–858
- Natsoulis, G. *et al.* (2008) The liver pharmacological and xenobiotic gene response repertoire. *Mol. Syst. Biol.* 4, 175
- Bolt, H.M. and Hengstler, J.G. (2009) Predicting drug metabolism-dependent toxicity. *Arch. Toxicol.* 83, 635–638
- Decker, M. *et al.* (2009) Mammalian epoxide hydrolases in xenobiotic metabolism and signalling. *Arch. Toxicol.* 83, 297–318
- Pelkonen, O. *et al.* (2008) Inhibition and induction of human cytochrome P450 enzymes: current status. *Arch. Toxicol.* 82, 667–715
- Guguen-Guillouzo, C. *et al.* (2010) Stem cell-derived hepatocytes and their use in toxicology. *Toxicology* 270, 3–9
- Brolen, G. *et al.* (2010) Hepatocyte-like cells derived from human embryonic stem cells specifically via definitive endoderm and a progenitor stage. *J. Biotechnol.* 145, 284–294
- Si-Tayeb, K. *et al.* (2010) Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells. *Hepatology* 51, 297–305
- Coecke, S. *et al.* (2006) Metabolism: a bottleneck in *in vitro* toxicological test development. The report and recommendations of ECVAM workshop 54. *Altern. Lab. Anim.* 34, 49–84
- Cashman, J.R. and Zhang, J. (2006) Human flavin-containing monooxygenases. *Annu. Rev. Pharmacol. Toxicol.* 46, 65–100
- Coughtrie, M.W. (2002) Sulfation through the looking glass—recent advances in sulfotransferase research for the curious. *Pharmacogenomics J.* 2, 297–308
- Guengerich, F.P. (2008) Cytochrome p450 and chemical toxicology. *Chem. Res. Toxicol.* 21, 70–83
- Meyer, U.A. (1994) Polymorphism of human acetyltransferases. *Environ. Health Perspect.* 102 (Suppl. 6), 213–216
- Testa, B. and Kramer, S.D. (2008) The biochemistry of drug metabolism—an introduction: part 4. Reactions of conjugation and their enzymes. *Chem. Biodivers.* 5, 2171–2336
- Wixon, J. and Kell, D. (2000) The Kyoto encyclopedia of genes and genomes—KEGG. *Yeast* 17, 48–55
- Matthews, L. *et al.* (2009) Reactome knowledgebase of human biological pathways and processes. *Nucleic Acids Res.* 37, D619–D622
- van Iersel, M.P. *et al.* (2008) Presenting and exploring biological pathways with PathVisio. *BMC Bioinform.* 9, 399
- Pico, A.R. *et al.* (2008) WikiPathways: pathway editing for the people. *PLoS Biol.* 6, e184
- Giesbertz, P. *et al.* (2009) Application of WikiPathways and PathVisio for pathway interpretation of toxicogenomics data. *Book of Abstracts Benelux Bioinformatics Conference (BBC-09)*, Liège, Belgium, 14–15 December 2009
- Testa, B. and Kramer, S.D. (2006) The biochemistry of drug metabolism—an introduction: part 1. Principles and overview. *Chem. Biodivers.* 3, 1053–1101
- Liska, D.J. (1998) The detoxification enzyme systems. *Altern. Med. Rev.* 3, 187–198
- Bader, G.D. *et al.* (2006) Pathguide: a pathway resource list. *Nucleic Acids Res.* 34, D504–D506
- Frolkis, A. *et al.* (2010) SMPDB: The Small Molecule Pathway Database. *Nucleic Acids Res.* 38, D480–D487
- Caspi, R. *et al.* (2010) The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nucleic Acids Res.* 38, D473–D479
- Hebels, D.G. *et al.* (2009) Molecular signatures of N-nitroso compounds in Caco-2 cells: implications for colon carcinogenesis. *Toxicol. Sci.* 108, 290–300
- van Iersel, M.P. *et al.* (2010) The BridgeDb framework: standardized access to gene, protein and metabolite identifier mapping services. *BMC Bioinform.* 11, 5
- Safran, M. *et al.* (2002) GeneCards 2002: towards a complete, object-oriented, human gene compendium. *Bioinformatics* 18, 1542–1543
- Chang, A. *et al.* (2009) BRENDA, AMENDA and FRENDA the enzyme information system: new content and tools in 2009. *Nucleic Acids Res.* 37, D588–D592
- Salomonis, N. *et al.* (2007) GenMAPP 2: new features and resources for pathway analysis. *BMC Bioinform.* 8, 217
- Hubbard, T.J. *et al.* (2009) Ensembl 2009. *Nucleic Acids Res.* 37, D690–D697
- Sayers, E.W. *et al.* (2009) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 37, D5–D15
- Parkinson, H. *et al.* (2009) ArrayExpress update—from an archive of functional genomics experiments to the atlas of gene expression. *Nucleic Acids Res.* 37, D868–D872
- Roth, R.B. *et al.* (2006) Gene expression analyses reveal molecular relationships among 20 regions of the human CNS. *Neurogenetics* 7, 67–80
- Reyermann, S. and Borlak, J. (2008) Topoisomerase II inhibition involves characteristic chromosomal expression patterns. *BMC Genomics* 9, 324
- Shi, L. *et al.* (2006) The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat. Biotechnol.* 24, 1151–1161
- Jennen, D.G. *et al.* (2010) Comparison of HepG2 and HepaRG by whole genome gene expression analysis for the purpose of chemical hazard identification. *Toxicol. Sci.* 115, 66–79
- Dai, M. *et al.* (2005) Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.* 33, e175
- Irizarry, R.A. *et al.* (2003) Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4, 249–264
- De Groot, P.J. *et al.* (2008) NuGO contributions to GenePattern. *Genes Nutr.* 3, 143–146
- Benedetti, M.S. *et al.* (2009) Drug metabolism and pharmacokinetics. *Drug Metab. Rev.* 41, 344–390
- Higgins, J.P. *et al.* (2004) Gene expression in the normal adult human kidney assessed by complementary DNA microarray. *Mol. Biol. Cell* 15, 649–656
- Wu, C. *et al.* (2009) BioGPS: an extensible and customizable portal for querying and organizing gene annotation resources. *Genome Biol.* 10, R130
- Su, A.I. *et al.* (2004) A gene atlas of the mouse and human protein-encoding transcriptomes. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6062–6067
- Castell, J.V. *et al.* (2006) Hepatocyte cell lines: their use, scope and limitations in drug metabolism studies. *Expert Opin. Drug Metab. Toxicol.* 2, 183–212
- Hsu, I.C. *et al.* (1993) p53 gene mutation and integrated hepatitis B viral DNA sequences in human liver cancer cell lines. *Carcinogenesis* 14, 987–992
- Knasmuller, S. *et al.* (1998) Use of metabolically competent human hepatoma cells for the detection of mutagens and antimutagens. *Mutat. Res.* 402, 185–202
- Wilkening, S. *et al.* (2003) Comparison of primary human hepatocytes and hepatoma cell line HepG2 with regard to their biotransformation properties. *Drug Metab. Dispos.* 31, 1035–1042
- Guillouzo, A. *et al.* (2007) The human hepatoma HepaRG cells: a highly differentiated model for studies of liver metabolism and toxicity of xenobiotics. *Chem. Biol. Interact.* 168, 66–73
- Kanebratt, K.P. and Andersson, T.B. (2008) HepaRG cells as an *in vitro* model for evaluation of cytochrome P450 induction in humans. *Drug Metab. Dispos.* 36, 137–145
- Harris, A.J. *et al.* (2004) Comparison of basal gene expression profiles and effects of hepatocarcinogens on gene expression in cultured primary human hepatocytes and HepG2 cells. *Mutat. Res.* 549, 79–99
- Kanebratt, K.P. and Andersson, T.B. (2008) Evaluation of HepaRG cells as an *in vitro* model for human drug metabolism studies. *Drug Metab. Dispos.* 36, 1444–1452
- Liguori, M.J. *et al.* (2008) Trovafloxacin-induced gene expression changes in liver-derived *in vitro* systems: comparison of primary human hepatocytes to HepG2 cells. *Drug Metab. Dispos.* 36, 223–233
- Olsavsky, K.M. *et al.* (2007) Gene expression profiling and differentiation assessment in primary human hepatocyte cultures, established hepatoma cell lines, and human liver tissues. *Toxicol. Appl. Pharmacol.* 222, 42–56
- Turpeinen, M. *et al.* (2009) Functional expression, inhibition and induction of CYP enzymes in HepaRG cells. *Toxicol. In Vitro* 23, 748–753