

Prediction of gene expression in human using rat in vivo gene expression in Japanese Toxicogenomics Project

Martin Otava,^{1,*} Ziv Shkedy,¹ and Adetayo Kasim²

¹Interuniversity Institute for Biostatistics and Statistical Bioinformatics; CenStat; Universiteit Hasselt; Belgium; ²Wolfson Research Institute; Durham University, Queen's Campus; Stockton-on-Tees, UK

Keywords: toxicogenomics, gene expression, drug-induced liver injury, rat-human translatability, multiple contrast test, dose-response relationship

Abbreviations: ANOVA, analysis of variance; DILI, drug-induced liver injury; FARMS, factor analysis for robust microarray summarization; FDR, false discovery rate; FWER, family wise error rate; I/NI, informative/non-informative; MCT, multiple contrast test; TGP, Japanese Toxicogenomics Project

The Japanese Toxicogenomics Project (TGP) provides large amount of data for the toxicology and safety framework. We focus on gene expression data of rat in vivo and human in vitro. We consider two different analyses for the TGP data. The first analysis is based on two-way analysis of variance model and the goal is to detect genes with significant dose-response relationship in both humans and rats. The second analysis consists of a trend analysis at each time point and the goal is to detect genes in the rat in order to predict gene expression in humans. The first analysis leads us to conclusions about the heterogeneity of the compound set and will suggest how to address this issue to improve future analyses. In the second part, we identify, for particular compounds, groups of genes that are translatable from rats to humans, so they can be used for prediction of human in vitro data based on rat in vivo data.

Introduction

The Japanese Toxicogenomics Project (TGP)¹ represents a unique source of information for toxicology and safety studies. The main topic that we address in this paper is related to the prediction of drug-induced liver injury (DILI) in humans using rat in vivo data (henceforth referred to as rat data). The analysis can be viewed from the perspective of translational research. Translation between rat and human data are an important topic,² due to high costs and ethical considerations of experiments that arise if development is moved to humans.³ Gaining strong scientific knowledge in animal models would prevent most risks. Translational research gets attention in all medical fields (e.g., refs. 4 and 5) and genes are a valuable tool in revealing connections across species (e.g., refs. 6 and 7).

In our particular case, successful prediction of a compound being toxic during rat experiments would reduce failure of efficacious compounds during the expensive phase III trials. An average of 10% of marketed drugs is being withdrawn from the market or requiring black-box warnings because of adverse drug reactions or toxicity. Moreover, failures in clinical phase III trial as well as in FDA submission have increased to approximately 50%

in recent years.⁸ In this paper, our aim is to explore the connection between humans and rats in terms of translatability of gene expression. Particularly, our goal is to model effect of dose on the gene expression in human in vitro data (henceforth referred to as human data) using the dose effect on the gene expression estimated from rat in vivo data. Therefore, our method would identify genes that enable translation of toxicity in rats into toxicity in humans.

The core part of the rat data set is gene expression level information across multiple compounds at multiple time points and dose levels. We focus on genes that are orthologous for rats and humans. Most of these genes are already annotated by biological processes or diseases (e.g., refs. 9 and 10). The analysis presented in this paper explores common dose-response pathways between rat and human genome using gene expression. Identifying a subset of genes that show similar dose-response gene expression profiles in rats and humans would support the translation of gene expression from rat in vivo experiments to human experiments. As in the case of DILI, this would enable the prediction of compounds' toxicity in humans using rats' in vivo experiments. The discovery of such genes would create knowledge about underlying mechanisms and connection between species which would

*Correspondence to: Martin Otava; Email: martin.otava@uhasselt.be

Submitted: 09/25/2013; Revised: 05/20/2014; Accepted: 05/30/2014; Published Online: 06/04/2014

Citation: Otava M, Shkedy Z, Kasim A. Prediction of gene expression in human using rat in vivo gene expression in Japanese Toxicogenomics Project. Systems Biomedicine 2014; 2:e29412; <http://dx.doi.org/10.4161/sysb.29412>

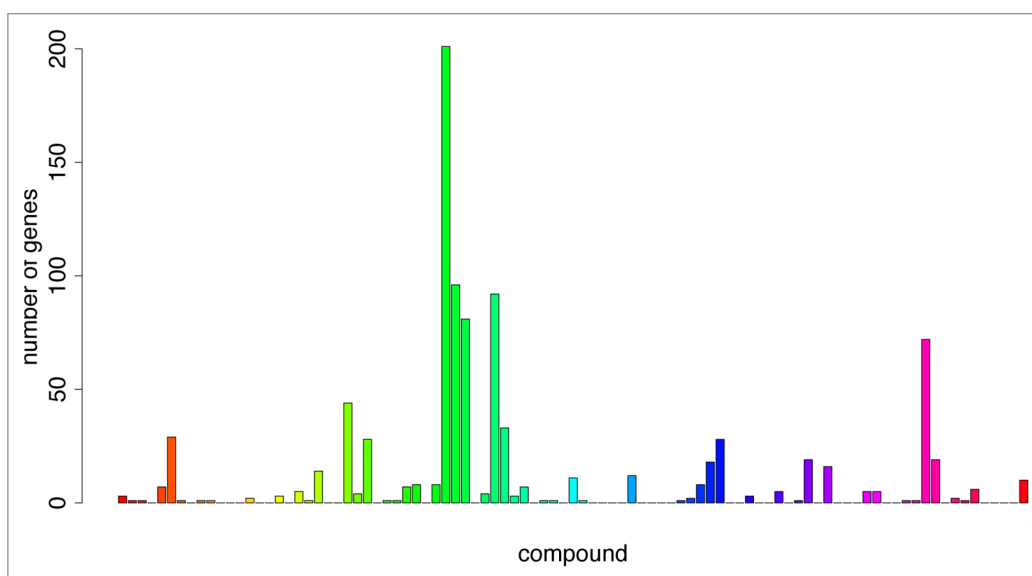


Figure 1. Number of genes with significant interaction in two-way ANOVA model, for both rat and human. The P values are adjusted using Bonferroni's method on significance level of 10%.

significantly improve how rat toxicology is used as a model for human toxicology in the later stages of drug development.

Results

Analysis of variance

Figure 1 shows the number of genes with significant interaction effect in both rats and humans and reveals a heterogeneous pattern across compounds. For example, for the compound *sulindac* there are 201 genes with significant interaction effect in both rats and humans while for the compound *perhexiline* there is only 1 gene in common. In total, only 54 compounds had at least one significant gene and only 10 compounds had more than 25 significant genes on the list. An example of one significant gene is shown in Figure 2. There exists a small set of genes that are significant in both rat and humans data consistently across subsets of compounds, even in case of strict multiplicity corrections. For the results presented in this paper we applied Bonferroni correction at significance level of 10%. The subset of compounds, identified through common significant genes, consists of DILI related compound only (if we convert the DILI status into binary variable, by pooling together “most concern” and “less concern” categories). Hence, the significance of the identified genes in rat in vivo could emphasize possible danger of DILI in humans. These genes are typically connected with the liver processes. Table 1 shows one of these genes, noted ASF1A (originally *Asf1a* in rat and ASF1A in human), that is significant for multiple compounds with DILI concern and not for any compound without DILI concern. Other genes from the identified set, FABP1, MCM4, SMC2, TXNRD1, show very similar behavior.

Trend analysis

As mentioned in the previous section, the second analysis consists of trend analysis per time point. An example of gene

complying with monotonicity assumption is shown in Figure 3. Our aim in this section is to predict the dose effect in humans using the dose effect in rat in vivo. All tests are based on MCT and P values are adjusted using Bonferroni correction using significance level of 10%.

At the first stage of the analysis, we identify, in the rat, the time point with the strongest signal. Figure 4 presents the number of genes with significant dose-response relationship per time point. It clearly shows that there are much more significant genes in the last time point, both for rats and humans, than in any other time point. Hence, for the remainder of this section, the dose effects in rats at the last time point are used for prediction. Figure 5 reveals that the number of significant genes in rats does not correspond with the number of significant genes in humans. For several compounds, there are no genes significant both in rats and humans. Hence, we focus on two gene sets: 1) genes significant in rats and, 2) genes significant both in rats and humans.

The dose effect in both rats and humans were estimated using isotonic regression. Only 91 compounds having high dose are considered for the analysis and we used the change in isotonic means of the rat (from the last to the first, i.e., control, dose level) in order to predict the change in isotonic means of the human. The example of resulting gene for the compound *omeprazole* is presented in Figure 6. We can see one of the genes where the translatability of rat data into human data is apparent. The mean at high dose for the rat represents differential expression of almost 6-fold change increase, while isotonic mean for humans shows almost 5 log-fold change increase. Prediction of all dose effects on humans using high dose effect on rats, when only genes significant in rats are used, are explored in Figure 7. As expected, prediction of control dose shows very low correlation, since all values for human control dose should be around zero. However, for higher doses we can see that there are genes with (nearly) the same value of isotonic means both for rat and human. Still,

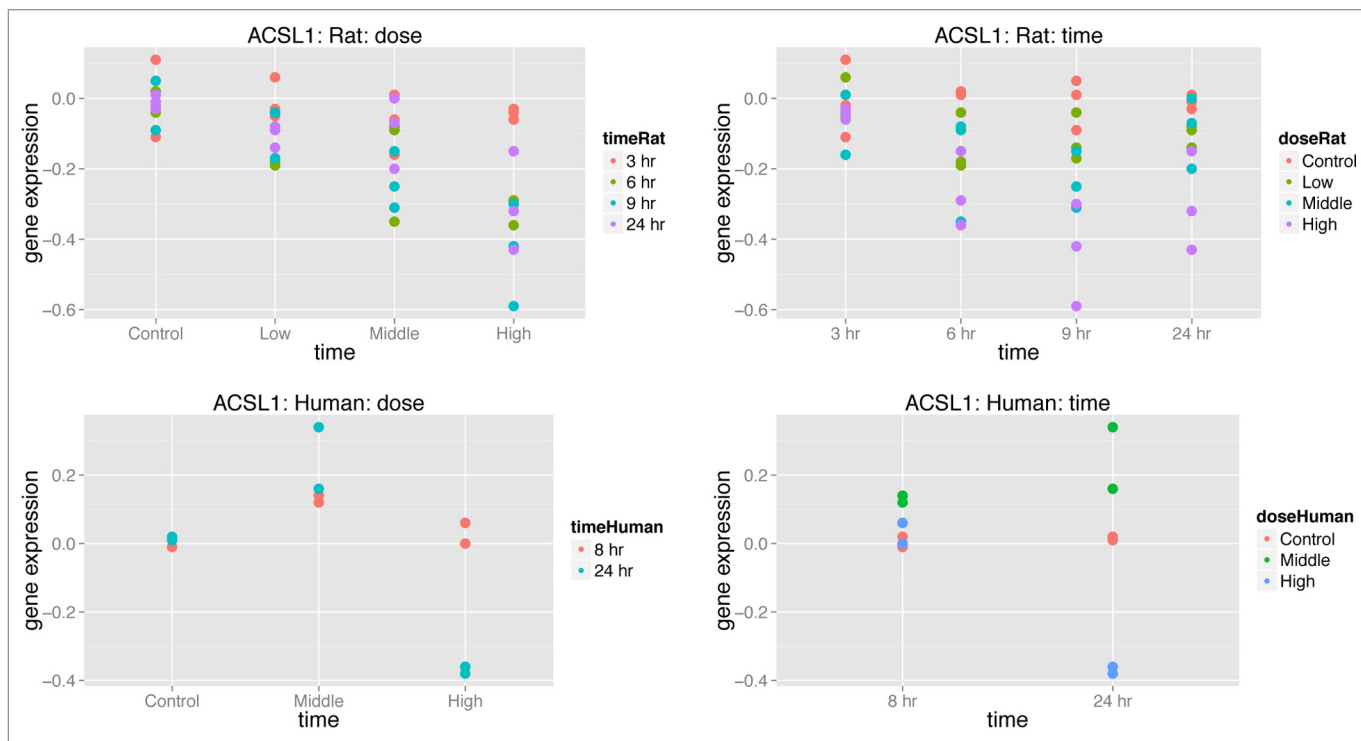


Figure 2. Example of gene with significant interaction in two-way ANOVA model, for both rat and human. Compound *omeprazole* and gene *Acs1* in rat, respectively *ACSL1* in human.

there is large amount of genes centered around zero. However, in **Figure 8**, where only genes that are significant in both rat and human in the last time point, the subset of genes around zero almost disappears. The resulting gene set reveals genes that are both consistently significant across species and translatable between species with respect to fold change induced by high dose of a given compound (*omeprazole* in this case).

Discussion

According to the ANOVA results, the number of significant genes varied among the compounds. This finding is not surprising since the data set contains very distinct compounds, both with respect to their structural properties and biological effects. The data set contains vitamin A next to ibuprofen or nicotinic acid. The analyses presented in this paper suggest that searching for overall differentially expressed genes can fail due to heterogeneity in the data set. Limiting ourselves to smaller subgroups of similar compounds can lead to more efficient analysis and meaningful results. One of such subsets was identified by our analysis, by grouping together 23 compounds with significant gene *ASF1A*. The presence of subgroups of compounds questions the meaningfulness of the goal of identifying genes useful for classification of compounds as DILI. If within given set of compounds would exist latent subgroups of compounds (similar with respect to their overall behavior), then particular genes could be good predictors of DILI in one subgroup, but not necessarily in other subgroups. In other words, genes that can be predictors for DILI

Table 1. Relationship between DILI concern status and simultaneous significance of interaction for both rat and human data for gene *ASF1A*

ASF1A		DILI status	
		No concern	Some concern
Significant interaction	No	8	62
	Yes	0	23

within one subgroup will lose its predictive ability by considering whole data sets with several subgroups of compounds. Besides the fact that the DILI response is highly unbalanced, only 8 compounds out of 93 show “no DILI concern.” Therefore, we propose to use a more specific response variable instead and simultaneously focus on possible identification of subgroups among compounds. These insights lead us to focus on translatability and means prediction in the second part of the analysis.

The second part of the analysis was mainly focused on the translatability of genes between humans and rats. The genes of interest are such that the fold change of their gene expression (precisely its log ratio against control) has similar values in rat and human data and moreover, the dose-response relationship is statistically significant in both species. We have shown that for some compounds, no relevant results were found. This is mostly due to very low overall difference in expressions and high variability. However, for several compounds, we were able to identify gene sets behaving in the desired way. The impact of the finding is clear: the value of gene expression observed in rats can be used as biomarkers for the corresponding gene expression value

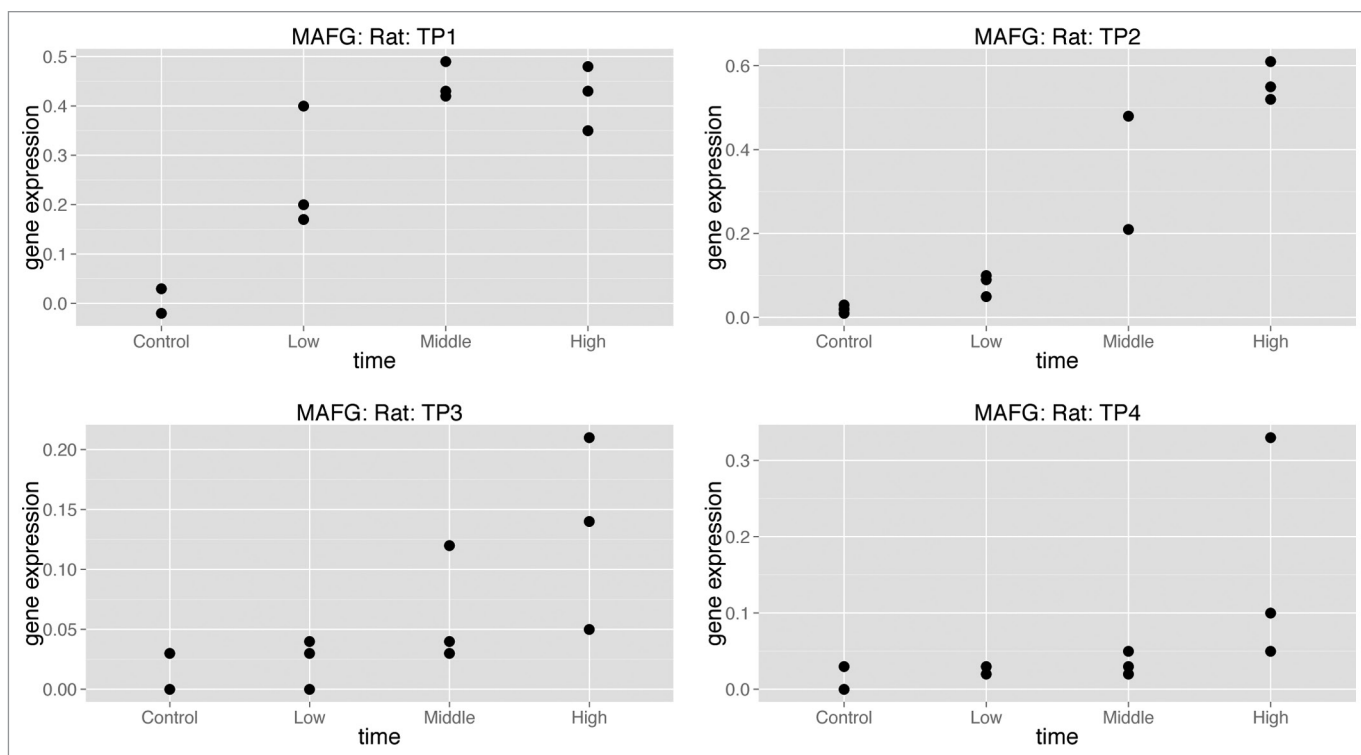


Figure 3. Example of gene with monotone dose-response profile for all time points in rat. Compound *omeprazole* and gene *Mafg* in rat (*MAFG* in human).

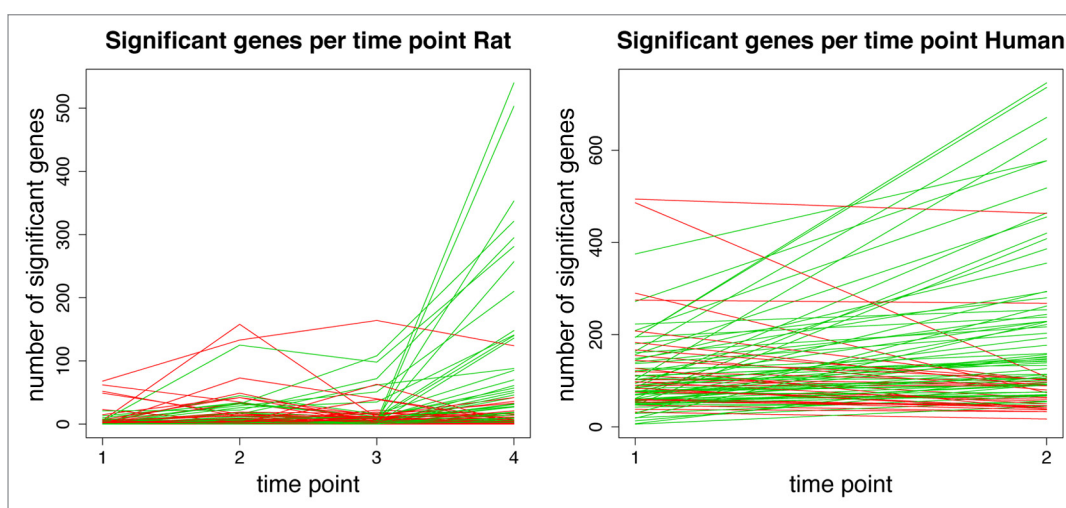


Figure 4. Number of genes with significant dose-response relationship per time point. Green compounds have maximum in last time point, red compounds in any other time point. Rat data results are displayed in left panel, human data in right panel.

in humans. If we are able to connect these genes with particular toxicological process, the signature made by these genes can serve us as early warning mechanism. The reliability of such genes as biomarkers will need to be validated, but the fact that they are significant in both species may highlight a common underlying biological mechanism in both species after exposure to the compounds. This study may provide a leeway into more extensive studies on rats and humans toxicogenomics connectivity in early drug developments.

Materials and Methods

Data sets

The data considered for the analysis presented in this paper consists of 93 compounds that are common in rat in vivo and human experiments and have DILI information available. In total, 4440 Affymetrix microarrays that measured gene expression profiles are available for rats (91 compounds with 48 arrays and 2 compound with 36 arrays) and 1116 arrays are available

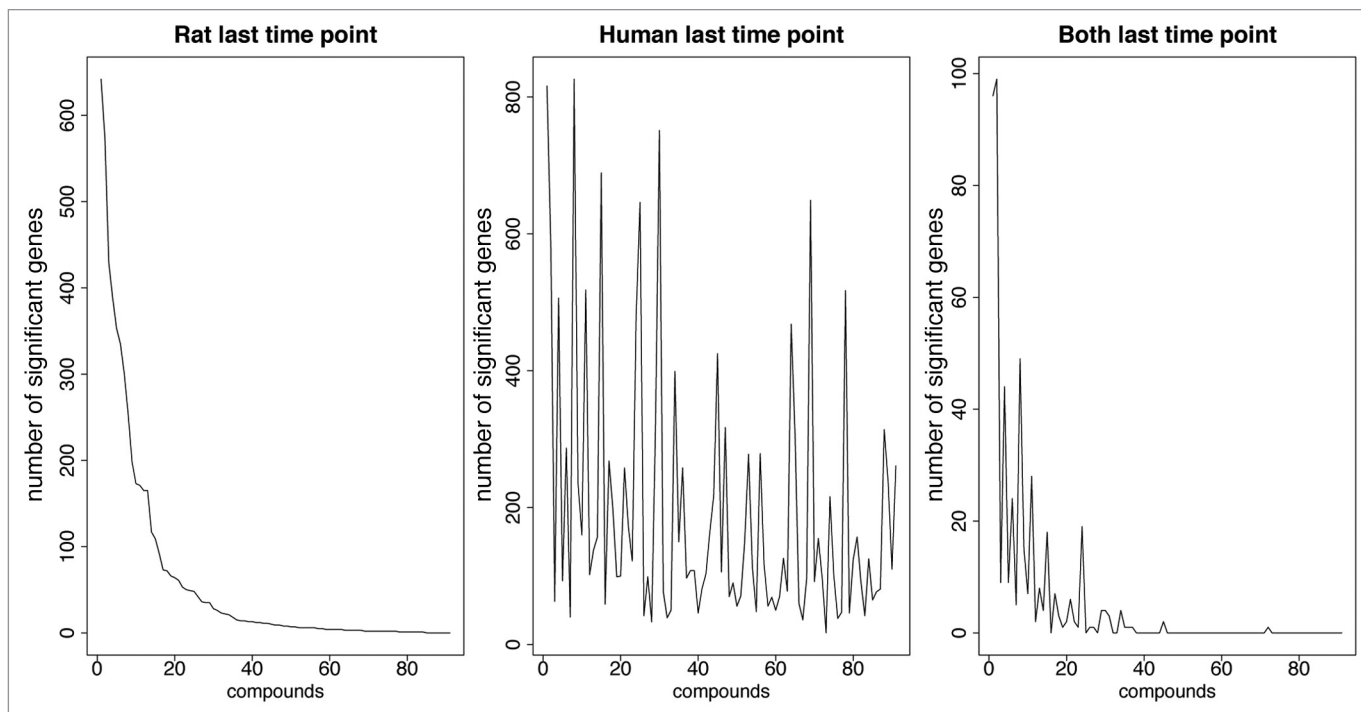


Figure 5. Number of genes with significant dose-response relationship in last time point. Compounds are ordered according to the number of significant genes in rat and ordering is kept across all three panels. Rightmost panel then shows intersection of two figures on the left.

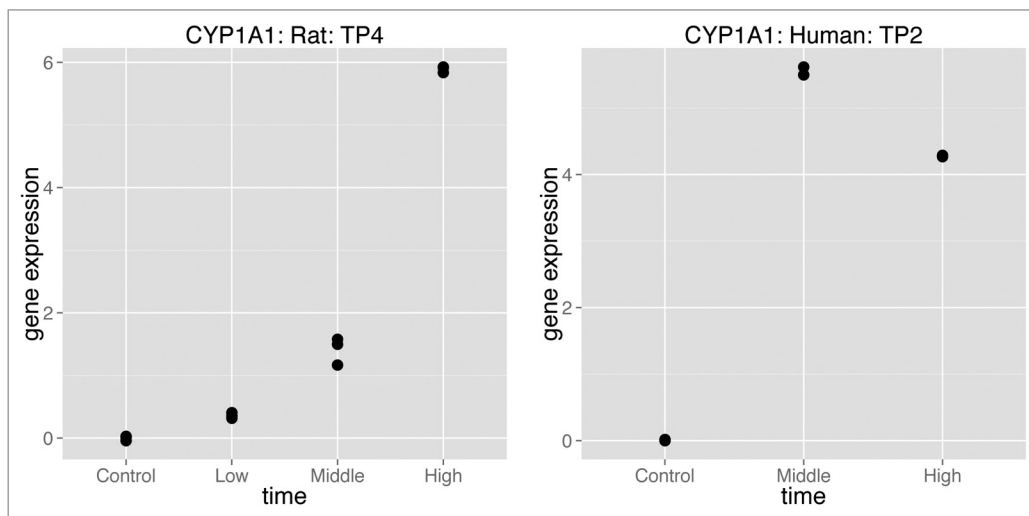


Figure 6. Example of gene translatable between rat and human for compound *omeprazole*: gene *Cyp1a1* in rat, respectively *CYP1A1* in human.

for humans (12 arrays per compound). We consider only genes that are orthologous for rats and humans. Further, we filter the genes using the I/NI calls criterion.¹¹ The preprocessed and filtered data set consists of 4359 genes. Response is computed as log ratio of the gene expression level against mean of expression levels under control dose (vehicle). The gene expression values are based on FARMS¹² summarized data.

For rat in vivo data, there are for each compound 48 arrays, 3 biological replicates measured at 4 dose levels (including control), each at 4 different time points (except 2 compounds with

36 arrays missing highest dose). In this study, human data set comprises 12 arrays per compound, 2 replicates measured at 3 doses and 2 time points. For the analysis presented in this paper we use the ordinal dose levels, i.e., low, middle or high as provided in original data sets.

Exploratory analysis: Analysis of variance approach

For the exploratory analysis, a gene specific linear model with dose and time as covariates is used. Interaction between covariates is also included. Let Y_{ijk} denotes the gene expression level for the i^{th} compound ($i = 1, \dots, 93$), j^{th} gene ($j = 1, \dots, 4359$)

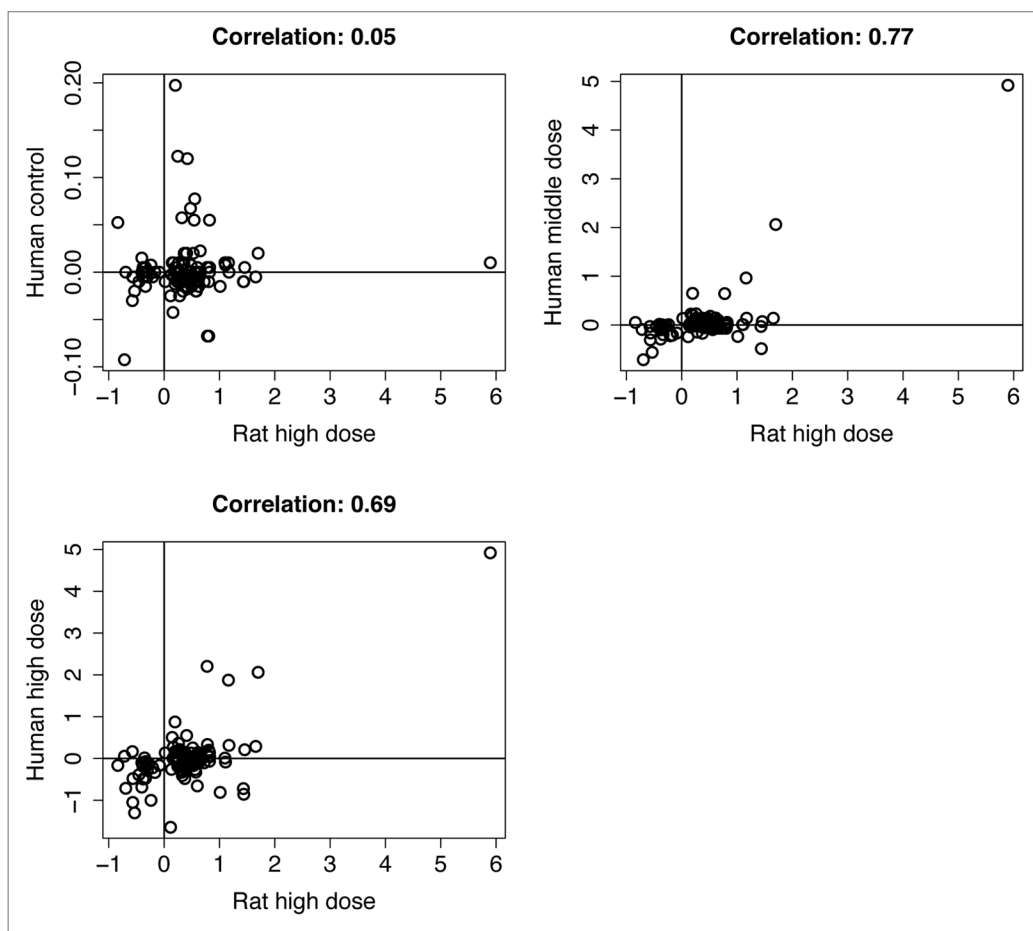


Figure 7. Dose effect for the compound *omeprazole*: estimated isotonic mean in particular dose in human against estimated isotonic mean in high dose in rat, both for last time point. Genes with significant dose-response relationship for rat in last time point (significance in human is not considered).

and k^{th} observation ($k = 1, \dots, 48$ or 36) based on time-dose combinations. To test possible dose effect, time effect and as well as their interaction, a two-way analysis of variance (ANOVA) model is used:

$$Y_{ijk} = \alpha_{0ij} + \beta_{Dij} Dose_{ijk} + \beta_{Tij} Time_{ijk} + \gamma_{ij} Dose_{ijk} * Time_{ijk} + \varepsilon_{ijk}.$$

Parameters α_0 , β_D , β_T , γ are gene (within compound) specific and the measurement error ε_{ijk} is considered to follow a Gaussian distribution $\varepsilon_{ijk} \sim N(0, \sigma_{ij}^2)$. The parameter vectors β_{Dij} , β_{Tij} , γ_{ij} represent the dose, time and interaction effects. In practice they comprise parameters representing particular level of explanatory variables [e.g., $\beta_{Dij} = (\beta_{DijCONTROL}, \beta_{DijDOSE1}, \beta_{DijDOSE2}, \beta_{DijDOSE3})$]. Note that the two-way ANOVA model specified above is fitted as a gene specific model within each compound. Testing if the parameters differ from null gives us an indication if the gene is differentially expressed for given compound, or not. However, gene specific omnibus test based on F-distribution can also be used to test if there is any significant effect at all.

Whatever test is used, multiplicity adjustment have to be applied due to extensive number of tests performed (4359 per compound). Correction for multiplicity was applied within each

compound. In general, either Family Wise Error Rate (FWER)¹³ or False Discovery Rate (FDR)¹⁴ can be used. Controlling FWER translates into level of certainty that there is no false positive finding among all our findings, but controlling FDR assumes there is at least one false positive finding while controlling for portion of false positive among all findings. Hence, FWER is a more conservative method than FDR. In our analysis, we apply Bonferroni method to control FWER to prevent false positives entering later stages of the analysis.

The whole procedure is conducted for both rat in vivo and human data. Only those genes that are found to be significant (according to test we choose) for both humans and rats are kept for further analysis. The resulting lists of significant genes are compared across compounds to identify genes that are significantly expressed in multiple compounds. Indicators of significance of a particular gene can be compared with DILI status of compounds to find out if the genes' appearance is connected with potential danger for the liver. In general, any information about compounds can be used in this stage and can be compared with indicator of genes significance. For example, pathological data available in the study can be used, information about compound chemical structure or grouping of compounds based on their phenotypic effect.

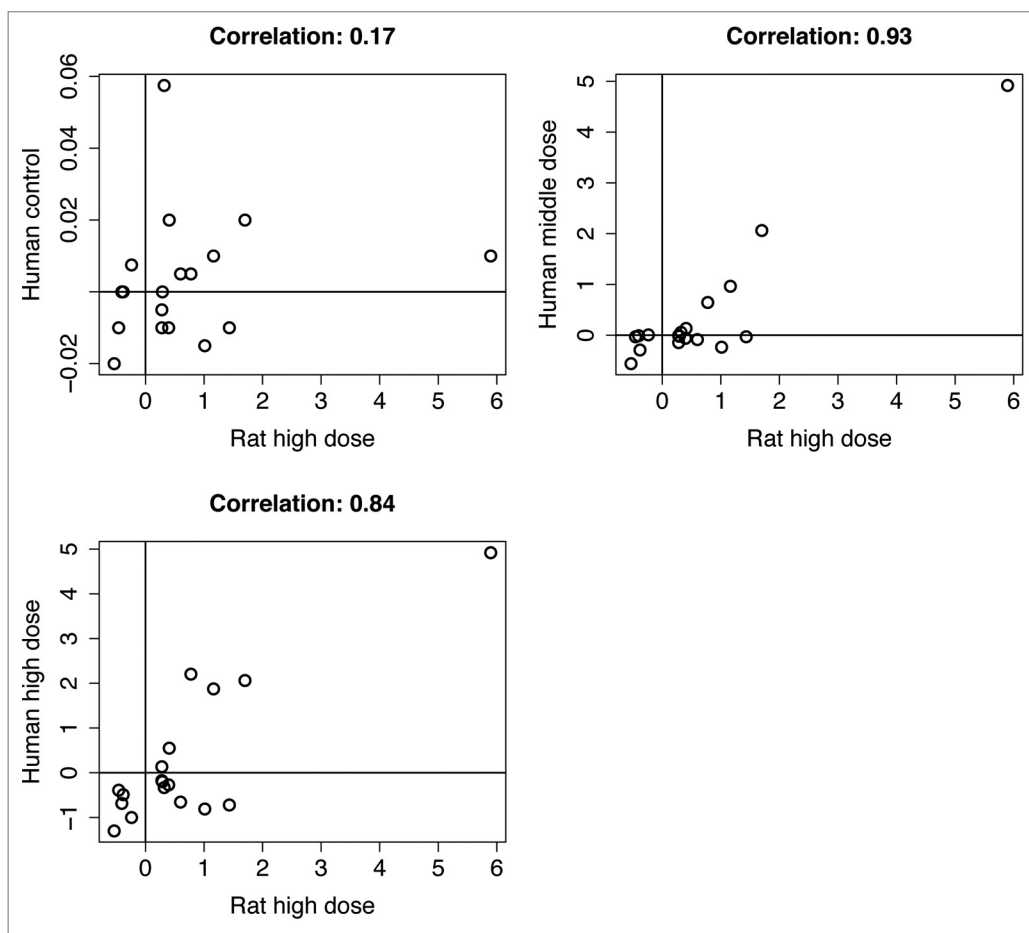


Figure 8. Dose effect for the compound *omeprazole*: estimated isotonic mean in particular dose in human against estimated isotonic mean in high dose in rat, both for last time point. Genes with significant dose-response relationship for both rat and human in last time point.

Main data analysis: Trend analysis approach

A trend analysis is a common analysis in toxicology. The aim of such analysis is to identify a subset of genes for which a monotone relationship with an increasing dose of a compound can be detected.¹⁵ Such an assumption of monotonicity allows us to gain power and it is scientifically reasonable. For toxicological studies, this assumption is typically used, since toxic effect usually gets stronger with increasing dose. Monotone means are computed for each gene using isotonic regression method.¹⁶⁻¹⁸ Isotonic regression pools together the means that violate assumption of monotonicity and makes these means equal. **Figure 9** shows examples of the isotonic means $\mu = \mu_0, \mu_1, \mu_2, \mu_3$ for an experiment with control dose and three active dose levels.

Hence, within the second modeling approach the null hypothesis of no dose effect is tested against an ordered alternative in the following way:

$$H_0 : \mu_0 = \mu_1 = \mu_2 = \mu_3 \quad \text{OR} \quad H_0 : \mu_0 = \mu_1 = \mu_2 = \mu_3$$

$$H_1 : \mu_0 \leq \mu_1 \leq \mu_2 \leq \mu_3 \quad \text{OR} \quad H_1 : \mu_0 \geq \mu_1 \geq \mu_2 \geq \mu_3$$

with at least one inequality strict. We start with simple ANOVA model:

$$Y_{ijkl} = \mu_{0ijl} + \delta_{ijl}Dose + \varepsilon_{ijkp}$$

where i stays for compound, j for gene, l for specific time point and k for observation within each time point (within gene, within compound). The vector of parameters $(\delta_{ijl} = \delta_{1ijl}, \delta_{2ijl}, \delta_{3ijl})$ represents the change of the mean in particular dose (compared with control dose) and parameters are either non-negative or non-positive (according to direction of monotonicity assumption). The measurement error follows a Gaussian distribution, $\varepsilon_{ijkp} \sim N(0, \sigma_{ijl}^2)$. An advantage of the model is absence of any parametrical assumption on dose response relationship shape. Dose-specific means are modeled separately, connected through deltas values.

The analysis is done per compound and per time point (and separately for human and rat). A multiple contrast test with Marcus' monotone contrast (MCT)¹⁹ is used to identify significant genes. The MCT is designed to cover all the possible parameter space of each model while using as few tests as possible (and so keeping power as high as possible). It comprises of several single contrast tests, while different combination of contrasts can also be used. We follow implementation arising from Marcus' test statistic²⁰ proposed for MCT by Bretz.²¹ Multiplicity adjustment is conducted within each compound and time point combination

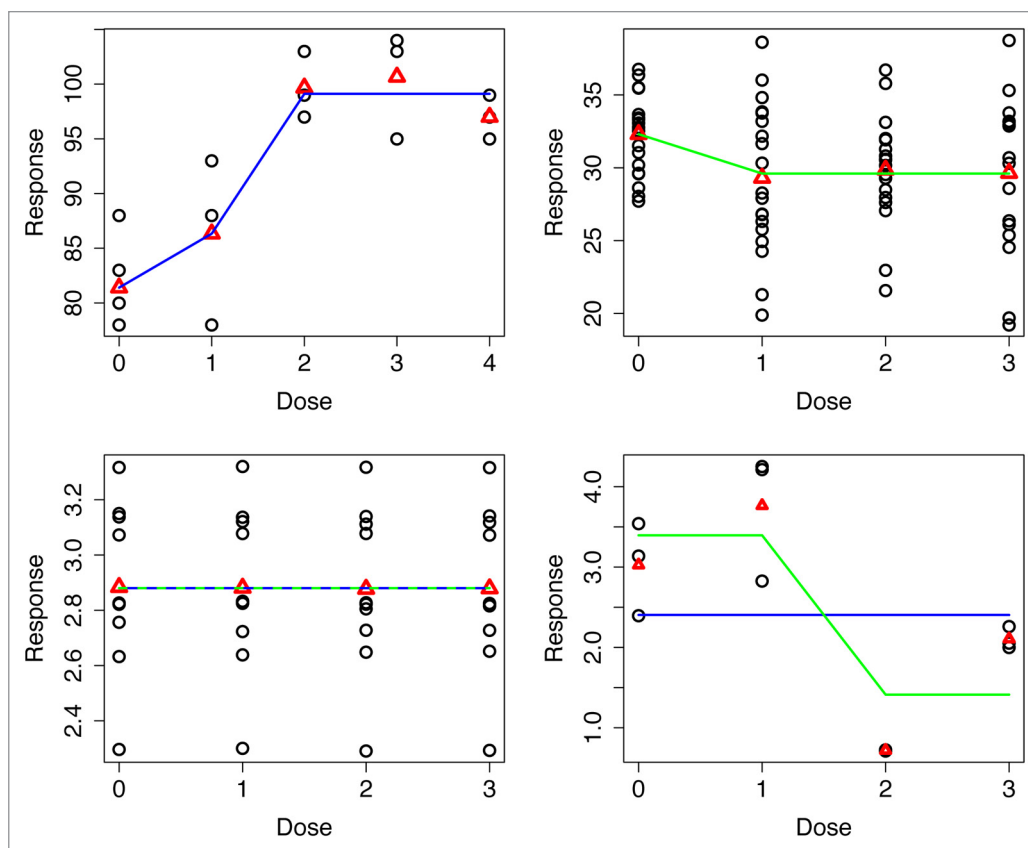


Figure 9. Examples of isotonic regression. Red triangles represent sample means and blue and green lines resulting isotonic means, under either upward or downward monotone assumption.

using FWER approach (with Bonferroni correction) within a gene and FDR adjustment across the genes.²²

Finally, for each compound and time point combination, we create lists of genes that show significant dose-response relationship. The time points with highest rate of significant genes (if such exist) are identified and we focus on them. Then, genes are listed that show significant dose-response relationship for such time points simultaneously in both rats and humans. For a particular gene on the resulting list, isotonic means at all doses are estimated and their values are compared between humans and rats. Hence, we can identify such genes in rats that can be used in order to predict the gene expression level in humans.

References

1. Uehara T, Ono A, Maruyama T, Kato I, Yamada H, Ohno Y, Urushidani T. The Japanese toxicogenomics project: application of toxicogenomics. *Mol Nutr Food Res* 2010; 54:218-27; PMID:20041446; <http://dx.doi.org/10.1002/mnfr.200900169>
2. McGonigle P, Ruggeri B. Animal models of human disease: challenges in enabling translation. *Biochem Pharmacol* 2014; 87:162-71; PMID:23954708; <http://dx.doi.org/10.1016/j.bcp.2013.08.006>
3. Hobin JA, Deschamps AM, Bockman R, Cohen S, Dechow P, Eng C, Galey W, Morris M, Prabhakar S, Raj U, et al. Engaging basic scientists in translational research: identifying opportunities, overcoming obstacles. *J Transl Med* 2012; 10:72; PMID:22500917; <http://dx.doi.org/10.1186/1479-5876-10-72>
4. Andrews NA. Skeletal regulation of glucose metabolism: challenges in translation from mouse to man. *IBMS BoneKEy* 2013;10:353; <http://dx.doi.org/10.1038/bonekey.2013.87>
5. Mestas J, Hughes CCW. Of mice and not men: differences between mouse and human immunology. *J Immunol* 2004; 172:2731-8; PMID:14978070; <http://dx.doi.org/10.4049/jimmunol.172.5.2731>
6. Seok J, Warren HS, Cuenca AG, Mindrinos MN, Baker HV, Xu W, Richards DR, McDonald-Smith GP, Gao H, Hennessy L, et al.; Inflammation and Host Response to Injury, Large Scale Collaborative Research Program. Genomic responses in mouse models poorly mimic human inflammatory diseases. *Proc Natl Acad Sci U S A* 2013; 110:3507-12; PMID:23401516; <http://dx.doi.org/10.1073/pnas.1222878110>
7. Rye MS, Bhutta MF, Cheeseman MT, Burgner D, Blackwell JM, Brown SDM, Jamieson SE. Unraveling the genetics of otitis media: from mouse to human and back again. *Mamm Genome* 2011; 22:66-82; PMID:21107580; <http://dx.doi.org/10.1007/s00335-010-9295-1>
8. Arrowsmith J. Trial watch: phase III and submission failures: 2007-2010. *Nat Rev Drug Discov* 2011; 10:87; PMID:21283095; <http://dx.doi.org/10.1038/nrd3375>
9. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al.; The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat Genet* 2000; 25:25-9; PMID:10802651; <http://dx.doi.org/10.1038/75556>

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Martin Otava and Ziv Shkedy gratefully acknowledge the support from the IAP Research Network P7/06 of the Belgian State (Belgian Science Policy). Martin Otava gratefully acknowledge the financial support of the Research Project BOF11DOC09 of Hasselt University.

10. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, Lerner J, Brunet JP, Subramanian A, Ross KN, et al. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science* 2006; 313:1929-35; PMID:17008526; <http://dx.doi.org/10.1126/science.1132939>
11. Talloen W, Clevert DA, Hochreiter S, Amaratunga D, Bijns L, Kass S, Göhlmann HWH. I/NI-calls for the exclusion of non-informative genes: a highly effective filtering tool for microarray data. *Bioinformatics* 2007; 23:2897-902; PMID:17921172; <http://dx.doi.org/10.1093/bioinformatics/btm478>
12. Hochreiter S, Clevert DA, Obermayer K. A new summarization method for Affymetrix probe level data. *Bioinformatics* 2006; 22:943-9; PMID:16473874; <http://dx.doi.org/10.1093/bioinformatics/btl033>
13. Hochberg Y, Tamhane AC. Multiple comparison procedures. New York: Wiley; 1987. p. 480
14. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B Met* 1995; 57:289-300
15. Lin D, Shkedy Z, Yekutieli D, Amaratunga D, Bijns L, eds. Modeling dose-response microarray data in early drug development experiments using R: order-restricted analysis of microarray data. Berlin: Springer; 2012. p. 282
16. Barlow RE, Bartholomew DJ, Bremner MJ, Brunk HD. Statistical inference under order restriction. New York: Wiley; 1972. p. 388
17. Robertson T, Wright FT, Dykstra RL. Order Restricted Statistical Inference. New York: Wiley; 1988. p. 544
18. Shkedy Z, Amaratunga D, Aerts M. Estimations under order restrictions. In: Lin D, Shkedy Z, Yekutieli D, Amaratunga D, Bijns L, editors. Modeling dose-response microarray data in early drug development experiments using R: order-restricted analysis of microarray data. Berlin: Springer; 2012. p. 11-27
19. Mukerjee H, Robertson T, Wright FT. Comparison of Several Treatments with a Control Using Multiple Contrasts. *J Am Stat Assoc* 1987; 82:902-10; <http://dx.doi.org/10.1080/01621459.1987.10478515>
20. Marcus R. The powers of some tests of the quality of normal means against an ordered alternative. *Biometrika* 1976; 63:177-83
21. Bretz F. An extension of the Williams trend test to general unbalanced linear models. *Comput Stat Data Anal* 2006; 50:1735-48; <http://dx.doi.org/10.1016/j.csda.2005.02.005>
22. Lin D, Hothorn LA, Djira GD, Bretz F. Multiple contrast tests for testing dose-response relationships under order-restricted alternatives. In: Lin D, Shkedy Z, Yekutieli D, Amaratunga D, Bijns L, editors. Modeling dose-response microarray data in early drug development experiments using R: order-restricted analysis of microarray data. Berlin: Springer; 2012. p. 233-247