Nuclear receptor variation in mice

Beratis, Alexander, Iacob, Diana, Boyanova, Dr. Desislava[1], Mewes, Prof. Dr. Hans-Werner[1]

Abstract

Nuclear receptors (NRs) are a large family of ligand-activated transcription factors, that bind directly to DNA to regulate the expression of target genes. They regulate critical functions in cell control, inflammation, fibrosis and tumor formation and are involved in metabolism, development and reproduction. Nuclear receptors influence the metabolism and signalling processes in the cells by changing the expression of target genes and are associated with numerous pathologies such as cancer, cardiovascular disease, and reproductive abnormalities. This paper presents the investigative results of knockout phenotypes and genetic variation for mouse NRs. Based on an assembly of all known mouse SNPs in the mouse NR genes, the phenotype information for genetic knockouts and genetic variation data was compiled from public databases. Knockout phenotypes were extracted from the Mouse Genome Informatics (MGI) database, while the Mouse Phenome Database (MPD) provides SNPs from various mouse strains, which can be correlated with extreme phenotypes measured in these mouse strains. The goal of this analysis is to find NR-associated SNPs in mice that influence changes in biological parameters such as body weight, body fat and other phenotypic traits. Furthermore, these findings will be coupled to phenotypes observed in human metabolic pathways and thus provide an insight into the implications of the investigated SNPs in human disease and drug development processes.

Keywords

Nuclear receptors — SNPs — Gene variation

[1]*Institute of Bioinformatics and Systems Biology, Helmholtz Zentrum MŸnchen, German Research Center for Environmental Health,*

Introduction

Looking across the evolutionary patterns between mouse and human, numerous research experiments and gene regulation studies have shown striking similarities regarding certain processes and systems in the two organisms. The mouse presents up to 95% genome similarity to humans and is thus often being used as a model organism when investigating anatomical, physiological or genetical markers in humans. Practically, mice are small, have an accelerated life cycle and represent a cost-effective alternative to genetic research and drug development for human diseases. Also, the majority of the genes responsible for complex diseases are shared between mice and humans, enhancing the chances of successfully identifying patterns in mice which would reveal human disease phenotypes [1].

This paper makes use of the publicly available data in the Mouse Genome Informatics database [2] and Mouse Phenome Database [3], respectively, in order to highlight changes in various biological parameters in mice under the influence of the NR-associated SNPs. Furthermore this enables the study of the human biology and disease, by mapping these findings to genotype - phenotype associations in humans found in the HMDB (Human Metabolome Database) [4], [5], EHMN (Edinburgh Human Metabolic Network) [6], Recon X (Reconstruction of Human Metabolism) [7] and KEGG (Kyoto Encyclopedia of Genes and Genomes) [8], [9].

1. Methods

Both the process of consolidating a genotype - phenotype map, as well as the subsequent analysis of extreme phenotypes observed in mice rely on four biological elements: mouse strains, gene names / symbols, nuclear receptors, SNPs. The correlation between the position of a nuclear receptor and various SNPs on a gene can result in

a specific phenotype. Due to increased specificity rate of the phenotype associated with a particular nuclear receptor, different NR and SNPs associations will result in different phenotypes for the same gene. Similarly, diverging mouse strains present non-identical phenotypes for the genetic parameters.

The analysis presented in this paper was based upon a dataset consisting of the 49 nuclear receptors of mouse [10] and their associated genes (242 unique gene Ids, 49 gene names / symbols); for a full list of gene names, see Appendix 6). The gene names were further on used to compile additional information regarding the position, associated SNPs and phenotypes for the nuclear receptors in mice.

1.1 MGI - Mouse Genome Informatics

Mouse Genome Informatics[1] [2] is a free, online database for the laboratory mouse and provides access to information about integrated genetics and associations between specific phenotypes and their corresponding alleles. It contains over 24000 genes and their protein sequences and approximately 48000 genotypes and phenotype annotations. For this research, the MGI database was solely used for building a connection between the nuclear receptor genes in the mouse and the associated phenotypes dependent on miscellaneous strains.

1.2 MPD - Mouse Phenome Database

Mouse Phenome Database[2] [3] includes annotations of measured data on the laboratory mouse strains and populations, as well as SNPs and phenotypes of the examined strains. More than 1330 strains were examined, providing annotations for over 3500 phenotype and 1.8 billion genotypes. The MPD database is more detailed and comprehensive than the MGI, so that the phenotypes found in MGI can be traced back to MPD. However, the MPD strictly associates individual phenotypes to their corresponding strain, posing difficulties in the mapping process between the mouse nuclear receptors and the MGI data.

1.3 Outcast: HMDB, EHMN, Recon X, KEGG

The nuclear receptor variation has been widely studied not only in the mouse, but also in the human genome. The research hereby refers to two published genome-wide association studies in this field of interest, namely *Kora*[3] - a GWAS study initiated at the Helmholtz Zentrum in Munich involving approximately 18,000 adults from Southern Germany - and *TwinsUK*[4], being the biggest adult twin genetic registry in the UK.

Using the information provided by these studies regarding the correlations between different SNPs and metabolites to metabolic pathways, this research aimed to identify the similarities between the nuclear receptors and their associated phenotypes in human and in the mouse, respectively. Therefore, it makes use of the information about metabolite names, classes and associated diseases, pathways and tissues for metabolites found in the HMDB (Human Metabolome Database) [4], [5], EHMN (Edinburgh Human Metabolic Network) [6], Recon X (Reconstruction of Human Metabolism) [7] and KEGG (Kyoto Encyclopedia of Genes and Genomes) [8], [9].

2. Database

Since annotation from multiple public databases and internal data were used, there was a need to design and implement a database supporting the research. The database structure thereby presents 10 tables and aims to facilitate the processes of gathering, parsing and statistical analysis of the nuclear receptors data, as well as to allow for a clear mapping between the nuclear receptor genes and the phenotypes and SNPs associated with them. The main challenge was to establish a connection between the MGI, MPD and HMDB tables. Therefore, there are 2 tables containing information from the MGI, 4 tables for MPD, 3 HMDB tables and the core table:

---

[1]`http://www.informatics.jax.org/`, March 23, 2015

[2]`http://phenome.jax.org/`, March 23, 2015

[3]`http://epi.helmholtz-muenchen.de/kora-gen`, March 23, 2015

[4]`http://www.twinsuk.ac.uk/`, March 23, 2015

- **nr_mapping** - core table, which contains the associations between the 49 nuclear receptors in the mouse and their corresponding gene names.

- **mgi** - contains the MGI gene annotations (MGI gene id, type, attributes, transmission etc.).

- **mgi_phenotypes** - stores phenotypic information in association with MGI genes; more than one phenotype can be associated with a specific gene.

- **mpd_snp** - contains the MPD SNP-annotations in correlation with the nuclear receptor genes.

- **ucsc** - contains a more detailed overview on the MPD mutations from *mpd_snp*

- **mpd_strains** - contains the MPD strain annotations (MPD strain id, sex, number of mice, Z-score etc.)

- **mpd_phenotypes** - stores phenotypic information in association with MPD strains; more than one phenotype can be associated with a specific strain.

- **kora** - stores metabolic information from the Kora study.

- **twins** - stores metabolic information from the Twins study in the UK.

- **human_metabolites** - stores metabolic information from the HMDB, EHMN, Recon X and KEGG.
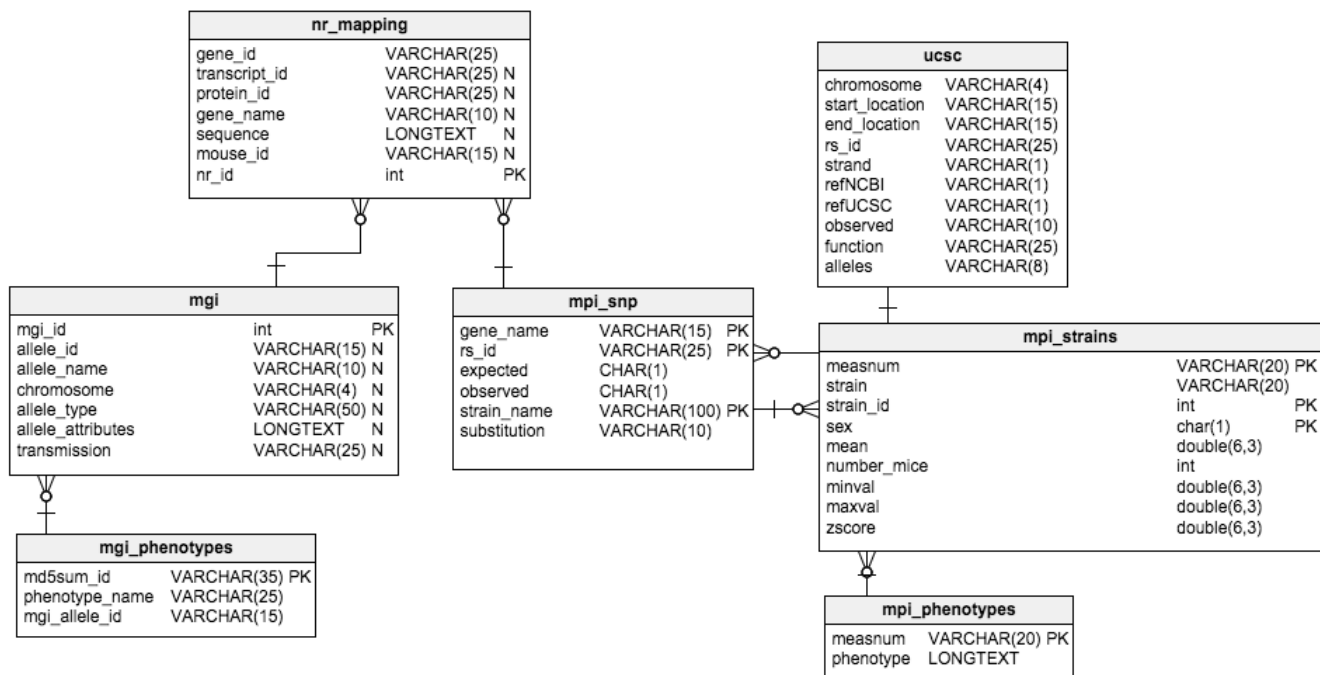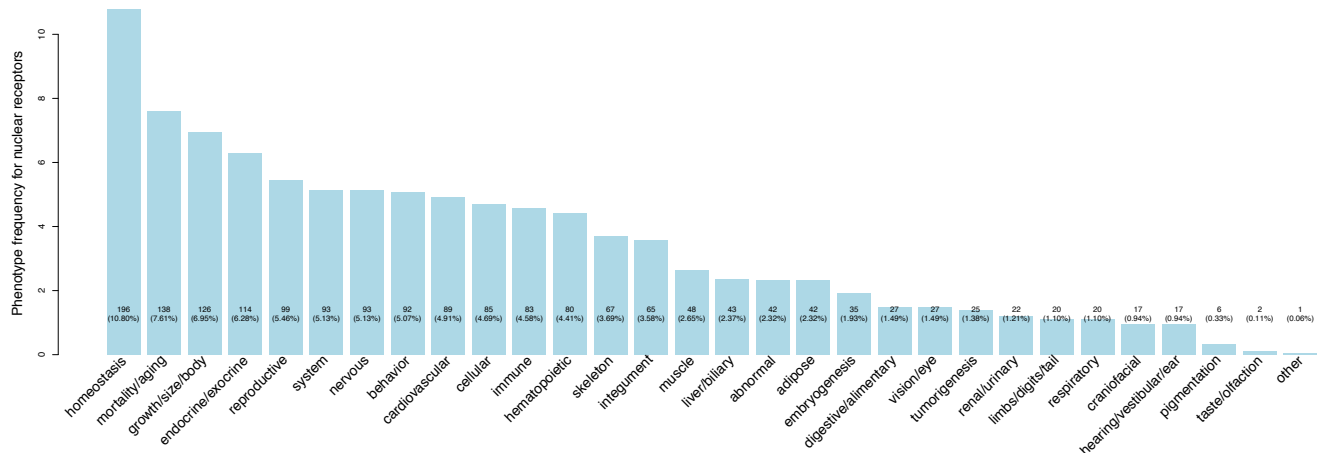

Figure 1. Database.
Database scheme for the 49 nuclear receptors with information about MGI, MPI and HMDB.

3. Results

The goal hereby was to visualise the most important phenotypes associated with the 49 nuclear receptors in the mouse. In this regard, the phenotypic and genomic data from the MGI and MPD was compared and analysed. All graphics were generated using the R package[5].

---

[5]www.r-project.org/, March 23, 2015

3.1 MGI Statistics

Figure 2. MGI phenotypes.
Mouse Genome Informatics phenotype distribution over the genes associated with the 49 nuclear receptors in the mouse.

The MGI database consists of generalised definitions of the phenotypes associated with the nuclear receptors found in the mouse. Figure 2 illustrates the most significant phenotypes and their occurrence frequency across the 242 genes corresponding to the nuclear receptors in the mouse. With a count of 196 matches across the dataset, the most prominent phenotype correlates with the homeostatic metabolic processes. Moreover, Figure 3 provides an insight into the exact association of each phenotype with the corresponding nuclear receptor genes, such that homeostatis, for instance, is prominently associated with mutations in the following genes: *Esr1 - estrogen receptor 1 alpha*, *Pparg - peroxisome proliferator-activated receptor gamma*, *Thrb - thyroid hormone receptor, beta*, *Thra - thyroid hormone receptor, alpha*, *Vdr - vitamin D receptor gene* etc. Other phenotypes describe body size and growth features and are representative for the *Pparg* genes.
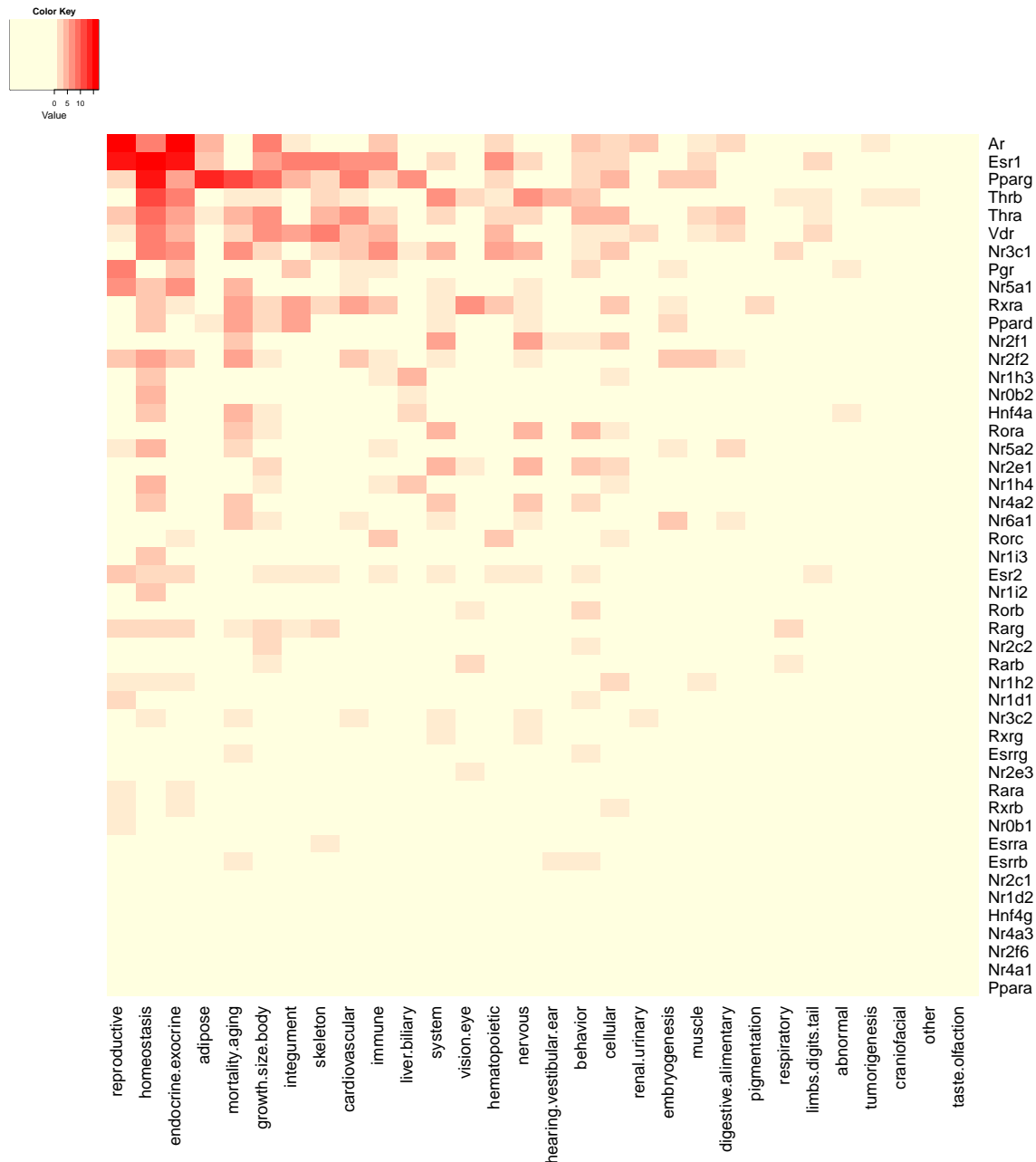
Figure 3. MGI phenotype - nuclear receptor gene associations.
Mouse Genome Informatics phenotype occurrence frequency among the nuclear receptor genes.

3.2 MPD Statistics

The MPDatabase contains annotations for extreme phenotypes associated with different mouse strains. The determining factor in *MPD* for associating a variation with a phenotypes is the Z-score, which describes the standard deviation of the standard distribution function. The range of the Z-score which indicates a significant deviation is below -2 (extremely low) and above 2 (extremely high). Between the range of -2 and 2, the Z-score does not deviate significantly enough from the mean to indicate any major changes and is therefore not relevant for this association.
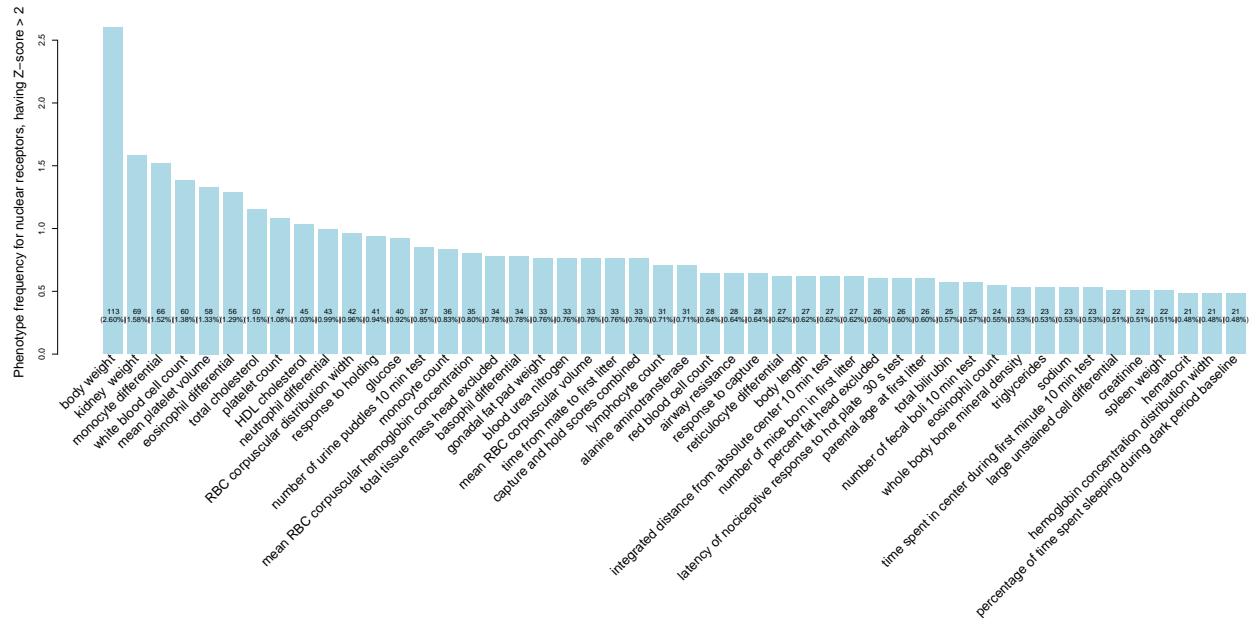
Figure 4. MPD extreme high phenotypes, having Z-score > 2.
Mouse Phenotype Database extreme high phenotype distribution over the genes associated with the 49 nuclear receptors in the mouse, having Z-score > 2
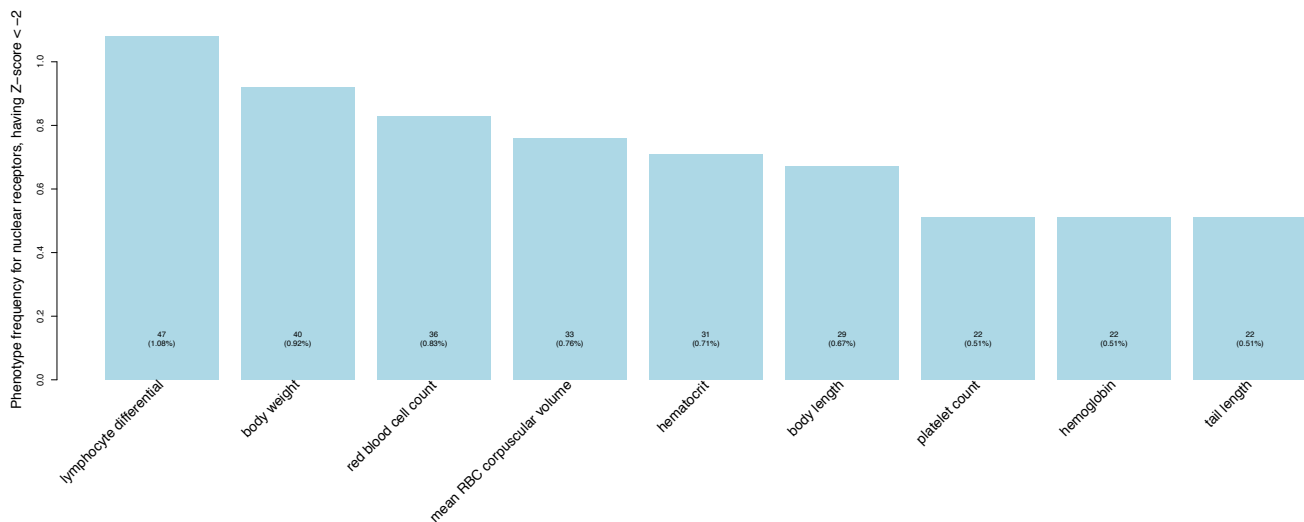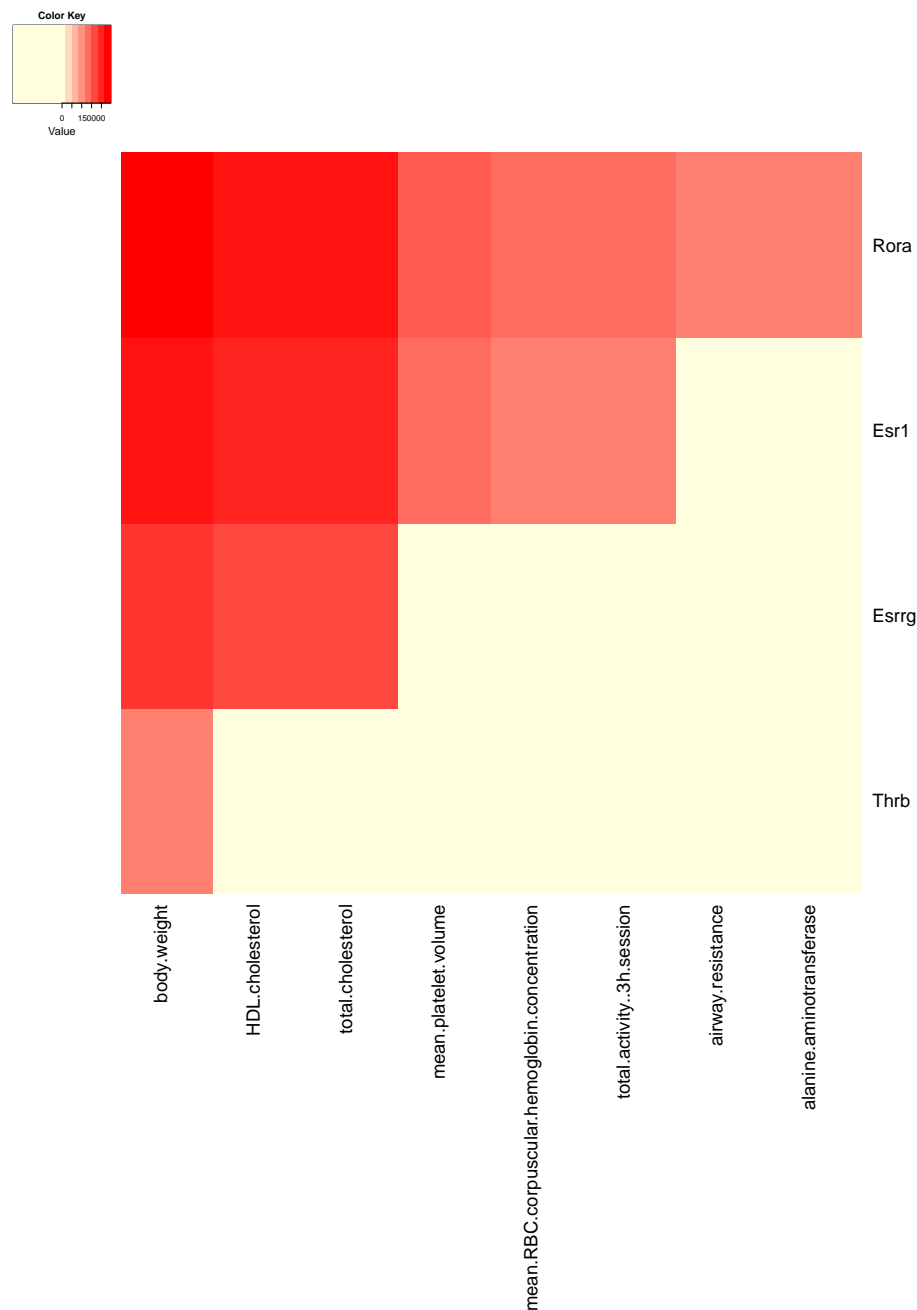

Figure 5. MPD phenotypes, having Z-score < -2.
Mouse Phenotype Database extreme low phenotype distribution over the genes associated with the 49 nuclear receptors in the mouse, having Z-score < -2

Based on these specifications, Figure 4 illustrates the most significant phenotypes and their occurrence frequency across the 242 genes corresponding to the nuclear receptors in the mouse, with a Z-score > 2 (extreme high phenotypes). Here, the most significant phenotypes include *body weight*, *kidney weight* and *cholesterol*. Similarly, Figure 5 presents the most significant phenotypes and their occurrence frequency across the 242 genes corresponding to the nuclear receptors in the mouse, having a Z-score < -2 (extreme low phenotypes). In this case, the most significant phenotypes include *body weight* as well, but the focus lies on several blood phenotypes (e.g. lymphocyte

differential, red blood cell count, haemoglobin).

Figure 6. MPD extreme phenotypes, having Z-score > 2.
Mouse Phenome Database extreme phenotype occurrence frequency among the 49 nuclear receptors in the mouse, having Z-score > 2

Moreover, Figures 6 and 7 provide an insight into the exact association of each phenotype with mutation in the corresponding nuclear receptor genes. Therefore, there are 4 genes associated with the extreme high and low phenotypes respectively, as following: *Rora - RAR-related orphan receptor alpha*, *Esr1 - estrogen receptor 1 alpha*, *Esrrg - estrogen-related receptor gamma*, *Thrb - thyroid hormone receptor, beta* and a fifth gene associated only with extreme low phenotypes - *Nr3c2*.
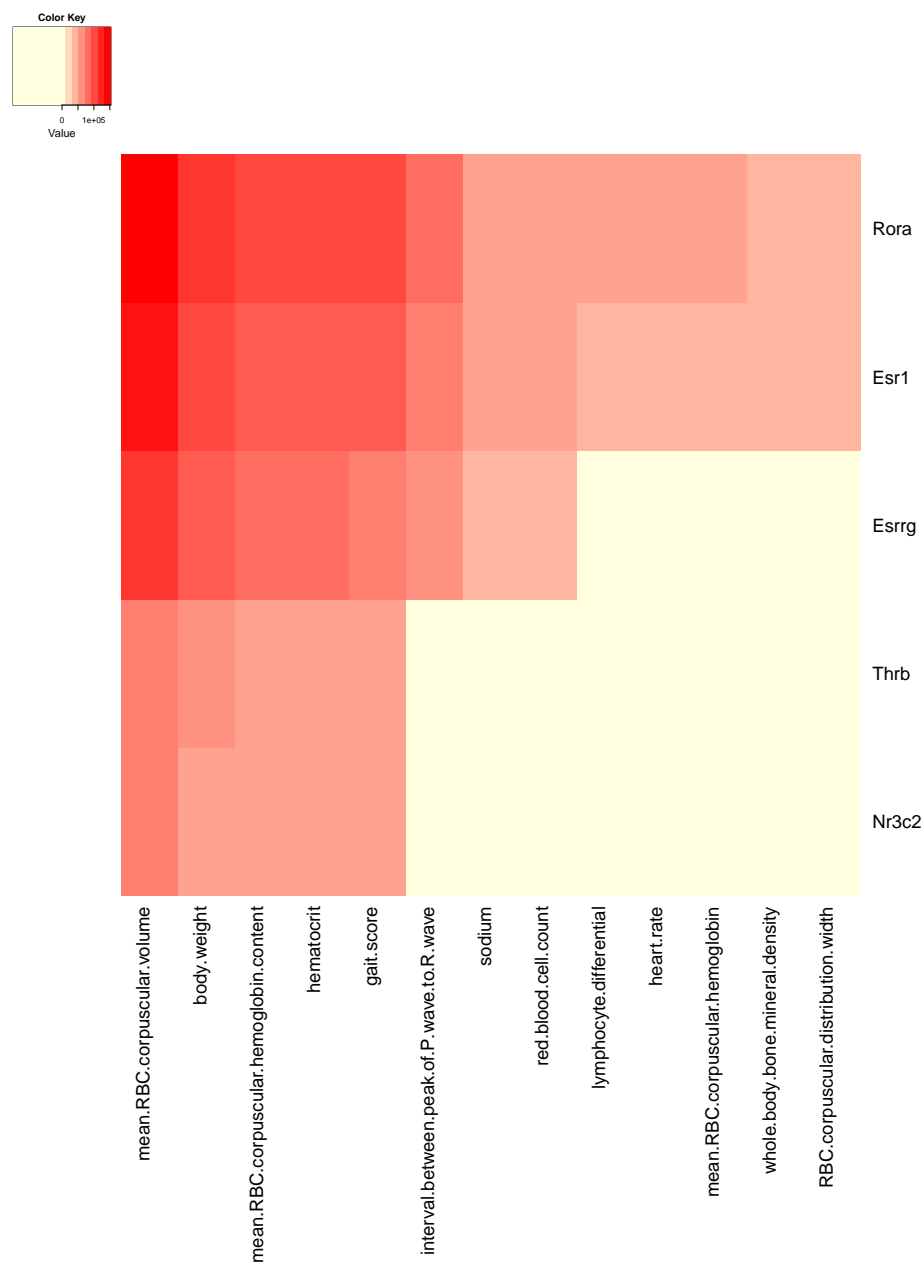
Figure 7. MPD extreme phenotypes, having Z-score $<$ -2.
Mouse Phenome Database extreme phenotype occurrence frequency among the 49 nuclear receptors in the mouse, having Z-score $<$ -2

The Z-score in *MPD* derives from measurements of metabolites or observations in specific mouse strains. In this regard, Table 1 shows a snippet of strains and their gene names associated based on their Z-score. This shows that the phenotypes are not directly associated with the expressed proteins, but rather with the observed abnormal condition in the mouse strains.

Strain	Gene	Phenotype	Z-score
CAST/EiJ	Rora (RAR-related orphan receptor alpha)	trigonelline relative abundance	8.8
C58/J	Rorc (RAR-related orphan receptor gamma)	N-acetylglutamate relative abundance	8.4
CAST/EiJ	Rara (retinoic acid receptor, alpha)	N2-acetyllysine (16h fast), relative abundance	4.6
NZB/BlNJ	Nr1i3 (nuclear receptor subfamily 1, group I, member 3)	ECG parameters interval between peak of P-wave to R-wave (PR)	4.6
CE/J	Nr0b1 (nuclear receptor subfamily 0, group B, member 1)	kidney total weight	4.4
NZB/BlNJ	Nr5a1 (nuclear receptor subfamily 5, group A, member 1)	succinylcarnitine relative abundance	3.8
C57L/J	Esr2 (estrogen receptor 2 beta)	relative size of perivascular immune cell clusters	3.5
CE/J	Rorb (RAR-related orphan receptor beta)	tigloylglycine (16h fast), relative abundance	3.3
NOD/ ShiLtJ	Ppard (peroxisome proliferator activator receptor delta)	percentage of parasites in brain relative to all organs tested	3.1
AKR/J	Esrrg (estrogen-related receptor gamma)	allantoin (16h fast), relative abundance	2.8
BUB/BnJ	Esr1 (estrogen receptor 1 alpha)	3-dehydrocarnitine relative abundance	2.7

Table 1. A snippet of extreme phenotypes.
Mouse strains associated with MPD extreme phenotypes, based on their Z-score.

3.3 Nuclear receptors in the human genome

This research shows that there are only a few phenotypes shared between the mouse and human. The reason therefore is that the phenotypes for the nuclear receptors in both human and mouse are involved in more than one metabolic pathway. For each nuclear receptor there were roughly 8 to 10 phenotypic annotations available, but only the ones which share the same function were mentioned in Tables 2 and 3.

Due to issues concerning the mapping of the mouse and human phenotypes relative to nuclear receptors, only two genes were found to have the same biological function in the mouse and in human - see Table 2. On the other hand, Table 3 shows a list of nuclear receptors from publications found in both the mouse and human and which present a similar phenotype or share the same biological function.

Nuclear Receptor Gene	Mouse phenotype	Human phenotype
Nr2e1	corticosterone, *steroid hormone*	lathosterol, *steroid hormone*
Essrg	carnitine, *lysine metabolism*	glutaroyl carnitine, *lysine metabolism*

Table 2. Nuclear receptors in human and in the mouse
Nuclear receptors in human and in the mouse which share a similar phenotype or biological function.

Nuclear Receptor Gene	Mouse phenotype	Human phenotype
Thrb	*thyroid hormone resistance*	*thyroid hormone resistance*
Nr0b1	*adrenal hypoplasia*	*adrenal hypoplasia*
Vdr	*osteoporosis*	*osteoporosis*
Ar	*spinal and bulbar muscular atrophy*	*spinal and bulbar muscular atrophy*

Table 3. Nuclear receptors in human and in the mouse
Nuclear receptors in human and in the mouse which share a similar phenotype or biological function.

4. Discussion and Outlook

Over the last year, the number of studies and publications involving the nuclear receptor variation has increased and scientist have been focusing on exploring the effects of nuclear receptors in different organisms, on various levels: in the cells, tissues, and even full-body impact. The goal hereby is to link their metabolic activity to possible human diseases and facilitate the drug development processes. In this regard, extensive experiments and studies have been performed on humans, as well as on laboratory mice, which are known to present a striking genomic resemblance to humans (up to 95%) and and time and money-wise more advantageous. As a result, it has been proved that nuclear receptors have a great influence over the extreme phenotypes showcased in the mouse, as well as in human.

4.1 MGI - MPD correlation

In spite of the ascending progress in this field of research, there annotation data provided in the mouse public databases - *Mouse Genome Informatics* and *Mouse Phenome Database* on nuclear receptors is still rather limited. However, the two databases provide complementary information: the phenotypes annotated in *MGI* describe roughly the biological systems in which they occur rather than the description of differences between non-variants, as of *MPD*. On the other hand, the nuclear receptor database for the human genome is far more detailed, comprising also information regarding the metabolic pathways associated with various phenotypes.

According to the MGI database, as shown in Figures 2 and 3, the most prominent phenotype correlates with the homeostatic metabolic processes, such as temperature regulation, pH-balance as well as lipid and glucose homeostasis and is associated with mutations in the following genes: *Esr1 - estrogen receptor 1 alpha*, *Pparg - peroxisome proliferator-activated receptor gamma*, *Thrb - thyroid hormone receptor, beta*, *Thra - thyroid hormone receptor, alpha*, *Vdr - vitamin D receptor gene*. The Esr1 - Estrogen receptor alpha gene is known to mediates effects of estrogen on glucose homeostasis [11]. Also the thyroid hormone, acting through thyroid hormone receptors has been proven to be a key regulator of metabolic homeostasis [12].

Comparably, the MPD database presents rather different phenotypes, associated with an extreme Z-score. Therefore, Figures 4 and Figure 5 illustrate extreme high and low phenotypes, respectively. Phenotypes presented with an extreme high effect include *body weight*, *kidney weight* and *cholesterol*, while those characterised by an extreme low effect on the organism depict blood phenotypes (e.g. *lymphocyte differential*, *red blood cell count*, *haemoglobin*). Nevertheless, there are 4 genes associated with the extreme high and low phenotypes respectively, as following: *Rora - RAR-related orphan receptor alpha*, *Esr1 - estrogen receptor 1 alpha*, *Esrrg - estrogen-related receptor gamma*, *Thrb - thyroid hormone receptor, beta* and a fifth gene associated only with extreme low phenotypes - *Nr3c2*, as supported by a recent review [13]. Studies [14], [15] have shown that the oestrogen receptor genes Esr1 and Esrrg are indeed involved in body growth processes, by playing an essential role in cell division and DNA recombination.

4.2 Mouse - Human correlation

This research presents the most significant phenotypes found in the mouse, based on the 49 nuclear receptor variation, as well as a mapping of these phenotypes on the human phenotypic expression, providing insight into possible human disease conditions. Nevertheless, the mouse public databases are far less comprehensive than the HMDB, for instance, allowing for only a generalised comparison between the two species, relative to the phenotypic expression in case of a variation in one of the same nuclear receptors.

Tables 2 and 3 present genes which were found to have the same biological function in the mouse and in human. On the one hand, gene *Nr2e1* regulates the expression for the same kind of molecule, a steroid hormone; this is consistent to the conclusions of several published studies[6], [7] and leads to the assumption that the nuclear receptors have somehow a similar influence on the individual in each species, although there are only few annotated by this

---

[6]http://www.phosphosite.org/proteinAction.do?id=2204525&showAllSites=true, March 23, 2015
[7]http://www.phosphosite.org/proteinAction.do?id=2204526&showAllSites=true, March 23, 2015

time. On the other hand, the overly researched gene Essrg has also been found, [16] to play an important role in the lysine metabolism in both human and the mouse.

The mapping of phenotype information between mouse and human showed that regardless if the genes are homolog, their function can differ between these two species, at least with the information we have by this time on the human nuclear receptors. As long as the human nuclear receptors are not enough researched, mapping the nuclear receptor variations from mouse on the human can only result in an indicator to get information about what protein the gene could code for or in which biological process is it likely to be involved.

5. Acknowledgments

6. Appendix

List of gene names associated with the 49 mouse nuclear receptors:

Esrrg, Rorc, Nr1i2, Hnf4g, Nr0b1, Nr3c1, Hnf4a, Rxrg, Esrra, Thrb, Nr5a1, Pparg, Nr4a1, Nr5a2, Nr1h3, Nr4a3, Nr1i3, Nr2f2, Nr4a2, Nr2f6, Esrrb, Ppard, Nr2e1, Vdr, Nr6a1, Nr1h5, Rarb, Nr1h2, Esr1, Rora, Rarg, Nr1h4, Rara, Nr3c2, Pgr, Ar, Rxrb, Thra, Esr2, Rorb, Ppara, Nr1d1, Nr0b2, Nr2e3, Nr2f1, Nr1d2, Nr2c1, Rxra, Nr2c2.

References

[1] Ali J. Marian. Modeling human disease phenotype in model organisms. 2011.

[2] Judith A. Blake, Carol J. Bult, Janan T. Eppig, James A. Kadin, and Joel E. Richardson. The mouse genome database: integration of and access to knowledge about the laboratory mouse. 2013.

[3] Terry P. Maddatu, Stephen C. Grubb, Carol J. Bult, and Molly A. Bogue. Mouse phenome database (mpd). 2011.

[4] Wishart DS, Tzur D, Knox C, Eisner R, and Guo AC et al. Hmdb: the human metabolome database. 2007.

[5] Wishart DS, Jewison T, Guo AC, Wilson M, and Knox C et al. Hmdb 3.0: the human metabolome database. 2013.

[6] Hao T, Ma HW, Zhao XM, and Goryanin I. Compartmentalization of the edinburgh human metabolic network. 2010.

[7] Thiele I, Swainston N, Fleming RM, Hoppe A, and Sahoo S et al. A community-driven global reconstruction of human metabolism. 2013.

[8] Kanehisa M and Goto S. Kegg: kyoto encyclopedia of genes and genomes. 2000.

[9] Kanehisa M, Goto S, Sato Y, Kawashima M, and Furumichi M et al. Data, information, knowledge and principle: back to metabolism in kegg. 2014.

[10] Qiongming Liu, Chen Ding, Wanlin Liu, Lei Song, Mingwei Liu, Liang Qi, Tianyi Fu, Anna Malovannaya, Yi Wang, Jun Qin, and Bei Zhen. In-depth proteomic characterization of endogenous nuclear receptors in mouse liver. 2013.

[11] Dahlman I, Vaxillaire M, Nilsson M, Lecoeur C, Gu HF, Cavalcanti-Proença C, Efendic S, Ostenson CG, Brismar K, Charpentier G, Gustafsson JA, Froguel P, Dahlman-Wright K, and Steffensen KR. Estrogen receptor alpha gene variants associate with type 2 diabetes and fasting plasma glucose. 2008.

[12] Michael L Goodson and Brenda J Mengeling. The role of corepressors in thyroid hormone-regulated metabolic homeostasis. 2014.

[13] Luo JQ, Wang LY, He FZ, Sun NL, Tang GF, Wen JG, Luo ZY, Liu ZQ, Zhou HH, Chen XP, and Zhang W. Effect of nr3c2 genetic polymorphisms on the blood pressure response to enalapril treatment. 2014.

[14] R Kato and H Ogawa. An essential gene, esr1, is required for mitotic cell growth, dna repair and meiotic recombination in saccharomyces cerevisiae. 1994.

[15] Michelle Percharde, Fabrice Lavial, Jia-Hui Ng, Vibhor Kumar, Rute A Tomaz, Nadine Martin, Jia-Chi Yeo, Jesús Gil, Shyam Prabhakar, Huck-Hui Ng, Malcolm G Parker, and Véronique Azuara. Ncoa3 functions as essential esrrb coactivator to sustain embryonic stem cell self-renewal and reprogramming. 2002.

[16] Fellows FC and Lewis MH. Lysine metabolism in mammals. 1973.