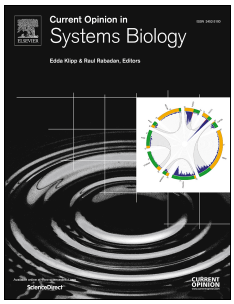
Accepted Manuscript



Big Data in Large-Scale Systemic Mouse Phenotyping

Holger Maier, Stefanie Leuchtenberger, Helmut Fuchs, Valerie Gailus-Durner, Martin Hrabe de Angelis

PII: S2452-3100(17)30052-5

DOI: 10.1016/j.coisb.2017.07.012

Reference: COISB 96

To appear in: *Current Opinion in Systems Biology*

Received Date: 17 March 2017

Revised Date: 14 July 2017

Accepted Date: 24 July 2017

Please cite this article as: Maier H, Leuchtenberger S, Fuchs H, Gailus-Durner V, de Angelis MH, Big Data in Large-Scale Systemic Mouse Phenotyping, *Current Opinion in Systems Biology* (2017), doi: 10.1016/j.coisb.2017.07.012.

Big Data in Large-Scale Systemic Mouse Phenotyping

Holger Maier[1], Stefanie Leuchtenberger[1], Helmut Fuchs[1], Valerie Gailus-Durner[1], Martin Hrabe de Angelis[†,1,2,3]

[1] Helmholtz Zentrum München - German Research Center for Environmental Health (GmbH), German Mouse Clinic, Institute of Experimental Genetics, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

[2] Chair for Experimental Genetics, Life and Food Science Center Weihenstephan, Technische Universität Munich, D-85354 Freising-Weihenstephan, Germany

[3] Member of German Center for Diabetes Research (DZD e.V.), Neuherberg, Germany

[†] corresponding author (correspondence to hrabe@helmholtz-muenchen.de, phone +49 89 31873502, fax +49 89 31873500)

Abstract

Systemic phenotyping of mutant mice has been established at large scale in the last decade as a new tool to uncover the relations between genotype, phenotype and environment. Recent advances in that field led to the generation of a valuable open access data resource that can be used to better understanding the underlying causes for human diseases. From an ethical perspective, systemic phenotyping significantly contributes to the reduction of experimental animals and the refinement of animal experiments by enforcing standardisation efforts. There are particular logistical, experimental and analytical challenges of systemic large-scale mouse phenotyping. On all levels, IT solutions are critical to implement and efficiently support breeding, phenotyping and data analysis processes that lead to the generation of high-quality systemic phenotyping data accessible for the scientific community.

Introduction

Phenotypic characterisation of mutant mouse lines has been used as a tool to study gene function since many decades. Usually, only a limited set of phenotypic traits is analysed in hypotheses-testing studies. Two decades ago, another paradigm of mouse phenotyping has emerged - the idea of systemic phenotyping. Large-scale random (ENU-driven) mutagenesis projects tried to identify novel phenotype-causing mutations by characterising every potential mutant for a large set of phenotypes in many organs [1-3]. Such genome-wide screening required a large panel of test procedures covering the whole physiological system of a mouse - thus the term *systemic phenotyping*. This led to the formation of so-called *mouse clinics,* with the German Mouse Clinic (GMC) as the first one established in 2001 [4-8]. Mouse clinics implement standardised breeding and

phenotyping pipelines for large-scale production of phenotype data to generate evidence-based hypotheses. Applying standardized systemic phenotyping on cohorts of mutant mouse strains is also known as "primary screening" [8], describing the fact of providing a first and unbiased, hypothesis-free look for phenotypic deviations in such mutants. As a next step, consortia of mouse clinics were formed in projects like Eumorphia and Eumodic [9-12] on the European level as well as the on-going International Mouse Phenotyping Consortium (IMPC) [13,14] on a global scale. Thus, in this review, the term *systemic phenotyping* also implies a large-scale, high-throughput component.

A rationale for large-scale systemic mouse phenotyping

Systemic phenotyping is no replacement for specific, hypothesis-driven mouse studies. Rather, primary screening of mutant mouse lines helps generating new hypotheses by providing a full picture of the system-wide effects of a specific genotype mutation. Pleiotropic effects (i.e. one gene influences more than one phenotypic trait), which often are not visible in hypothesis-driven, focussed studies, can be uncovered this way. This has been shown many times by mouse clinics [15-20]. For example, a recent joint study showed that spermidine treatment protects the heart from age associated deterioriatons and leads to life time extention. The cardioprotective effects of spermidine may be due to several underlying mechanisms, including both direct cardiac effects and extracardiac (systemic and renal) effects. Systemic effects by spermidine might involve anti-inflammatory processes, as well as a blood-pressure-lowering effect [17].

Systemic phenotyping can be understood as physiology-wide or alternatively, as a phenome-wide approach. On a genome scale, the IMPC aims to produce an encyclopaedia of gene function for all mouse genes [14].

Evidently, an open-access resource of well-structured systemic phenotype data can be subjected much better to data mining methods in order to identify biological mechanisms that cannot be uncovered otherwise. This is impressively supported by a series of first publications from the Eumodic and the IMPC projects [21-25]. For example, in order to study sexual dimorphism, IMPC scientists analysed up to 234 characteristics of more than 50,000 mice, including over 40,000 mutant mice. It was shown that sex influences the prevalence, course and severity of the majority of common diseases and disorders [24]. The IMPC provides full public access to the generated results, including data visualisation tools and machine-to-machine interfaces (APIs) [26]. Phenotypic similarities between inherited human diseases and knockout mouse lines are presented at the IMPC webpage and can be used to find suitable disease models for clinical researchers [27].

Due to the rather large panel of tests applied on a single animal, systemic phenotyping requires a well-defined *pipeline* - the composition as well as the order and time point of every test procedure. An example of such a primary screening pipeline from the GMC is shown in table 1.

Standardisation also applies on phenotyping methods. In the IMPC, IMPReSS (International Mouse Phenotyping Resource of Standardised Screens, https://www.mousephenotype.org/impress), is a database of pipeline definitions and Standard Operation Procedures (SOPs) for all tests, including lists and definitions for every measured test parameter (e.g. body weight) and metadata (e.g. experimenter).

Such standardisation efforts ensure good scientific practice, improve reproducibility, enhance data quality and make data more suitable for common analysis methods [28].

From an animal welfare perspective, systemic phenotyping directly means reduction of experimental animal use, as several hundred phenotypic parameters can be measured on an individual animal.

Systemic phenotyping as a process

Mouse clinic operations can best be described and handled in form of modular processes. A generic mouse clinic business process model has originally been developed at the GMC, but can also be applied to other mouse clinics, as shown in [29]. Large-scale mouse phenotyping requires a well-organised mutant generation and breeding pipeline to provide enough age-matched mice of desired genotype for phenotyping. Figure 1 shows a simplified version of such a process model as a flowchart.

The importance of electronic data management

Systemic phenotyping produces large amounts of data, which need to be highly structured to be suitable for subsequent software-assisted data processing. For instance, any measured parameter (e.g. blood glucose concentration) needs attributes like data type (float, integer, text) or unit (mmol/l, g). Also demographic data (sex, genotype, date of birth) is captured and stored for every mouse.

At the GMC, 695 test parameters and 410 metadata parameters (listed in IMPReSS, see above) are captured in 26 procedures in the IMPC phenotyping pipeline. These numbers show that spread sheet based data management is not possible here.

The IMPC data release 5.0, published 2 August 2016, includes data from 3532 mutant lines (80.781 mice) and 24.023 wild type control mice. Total data is composed of 8.107.737 categorical, 5.938.585 unidimensional and 111.319 text data points as well as 8.772.128 time series and 270.804 image records (http://ww.mousephenotype.org/data/release). These figures illustrate the scale and complexity of large-scale systemic phenotyping data.

The use of electronic Laboratory Information Management Systems (LIMS) is critical for capture and management of all phenotyping and demographic data. LIMS can support complex logistics processes with planning tools, so that every test is timely applied to a particular mouse according to the pipeline.

Customised LIMS have been developed at different phenotyping facilities, which is described and discussed in detail in [29].

Standardised quality control and data analysis

In large-scale systemic phenotyping, data needs to undergo standardised procedures for quality control (QC) due to the sheer amount of data and the possibility for errors. In the IMPC, a thorough QC process has been established at the Data Coordination Centre *PhenoDCC* [30]. The process involves automated and manual checking of data consistency and out of range data points as well as a ticket system to track possible data quality issues. Such massive data quality control and standardisation efforts are usually not possible in individual research labs.

Automated data analysis is crucial for working with large-scale phenotyping data. At the GMC, standardised R scripts [31] for data visualisation and statistics are developed for

every single phenotyping test and routinely applied to the data in order to determine genotype-related phenotype deviations. Another such toolkit, PhenStat, has been developed by the IMPC consortium for the same purpose [32].

Systemic phenotyping is also challenged with the n<<p problem, because a large number of parameters (p) is measured on a single mouse, while the sample size (n) is much lower. In the standard pipeline of the GMC, a cohort size of n=15 animals per sex and genotype is used (15+15 male/female mutants, 15+15 male/female controls), while in the IMPC pipeline, this number is lower (n=7). Sample size n directly affects the α and β error probabilities (false positives and false negatives). Keeping them low demands for increasing the sample size, which would result in higher statistical power, but also higher costs and animal use. Thus, the currently chosen numbers for n are a trade-off between these opposing requirements.

A further statistical challenge is the multiple testing problem. Simultaneously testing large numbers of parameters with inferential methods, requires α adjustment to avoid inflation of false positive detection rate.

Minimising metadata variability is another requirement. For instance, mutant and control mice should be measured by the same experimenter. Otherwise, a positive result could just indicate a possible experimenter influence rather than a genotype effect. Naturally, procedures that use a human scoring step are most prone to experimenter bias, although this can be reduced by experimenter training and procedure standardisation. While human scoring is not part of every test procedure, experimenter metadata is still routinely captured to allow retrospective studies of such influence on a growing data pool. Thus, experimenter bias can routinely be monitored and corrected for.

Standardisation of phenotyping results - the use of ontologies

Direct raw data comparison is not always possible, since metadata may differ between mouse clinics and accordingly, shifts in data ranges can be observed. In such cases, mutant and control mice usually exhibit a similar shift and statistical analysis still leads to comparable results between clinics. However, a qualitative results level ("phenotypic difference yes/no?") is required in order to facilitate data comparison.

Ontologies provide a perfect solution here. Ontologies are hierarchically organised structures of controlled vocabulary. The Mammalian Phenotype Ontology (MP) [33] currently includes almost 13.000 classes to describe any mammalian phenotype in a standardised way, e.g. "MP:0005559 increased circulating glucose level". Providing unique IDs, MP terms allow systematic and programmatic exploitation of phenotyping result databases [34]. The assignment of a distinct MP term to a mutant mouse line is based on statistical analysis of raw data to make a binary decision between "MP term assigned" and "MP term not assigned".

The possibility of cross-linking phenotyping results with other public databases allows mapping of mouse phenotyping results to phenotyping results of other species, e.g. the Human Phenotype Ontology (HPO) [35,36]. Building such "data bridges" from biology to medicine have been subject of the recent EU-funded BioMedBridges project (http://www.biomedbridges.eu) and is followed up in the CORBEL project (http://www.corbel-project.eu).

Working with systemic phenotyping data - interactive, data mining and multivariate approaches

While the use of ontologies requires bioinformatics expertise, a much simpler and even more intuitive approach is to apply so-called *phenomaps*, an adaptation of the heatmap visualisation well-known from the transcriptomics field. In this case, the phenotyping results are drastically reduced to a qualitative yes/no statement. As shown in figure 2, a simple matrix of mutant lines vs. physiological category allows intuitive identification of mutant mouse lines of interest for the non-expert user. An application of phenomaps is the use of clustering methods to identify mutant mouse lines that show a similar overall or partial phenotype profile.

A still very intuitive, however quantitative approach is using the full raw data set. It can be applied for interval-scaled phenotype parameters. For a given parameter, the mean value of mutants is divided by the mean value of control animals to form a mutant/control ratio. Mutant/control ratios from many lines can be plotted as a histogram, as shown in figure 3. In the resulting distribution, mutant/control ratios near 1.0 correspond to "no genotype-related phenotype deviation", whereas mutant/control ratios at the left and right margins of the histogram mean "decreased/increased parameter phenotype". Being quantitative, this allows applying an individual threshold to factor in biological relevance. For example, this method can be used to rapidly identify mutant mouse lines showing an extreme deviation from blood glucose levels compared to control animals by selecting lower and upper 5% from a distribution of several thousand mutant lines - these can be considered candidate genes for a "low/high glucose" phenotype and put in a gene set enrichment analysis.

More advanced data mining methods include data integration approaches to link phenotype information with other public databases (e.g. Gene Ontology (GO) [37] or KEGG [38,39]) and then apply data mining algorithms or gene set enrichment analysis (GSEA). For instance, PhenoDigm (Phenotype comparisons for DIsease Genes and Models) [40] uses an association rule mining approach to automatically integrate data from a variety of model organisms using several scoring methods to identify only strongly data- supported gene candidates for human genetic diseases. The PhenoDigm automated pipeline and manual curation lead to the analysis of the frequency of IMPC models that correspond to Mendelian Disease-Genes in OMIM or Orphanet. 650 rare disease-gene associations were identified leading to valuable mouse models for these genetic diseases, published in Meehan et al., 2017 [25].

Phenotypic readout for a given disease or syndrome typically involves several phenotypic parameters. A well-known example is the metabolic syndrome, which involves abdominal obesity, elevated blood pressure, plasma glucose, serum triglycerides and low HDL levels [41,42]. Multivariate methods are therefore needed to address the large-scale analysis of such phenotypic patterns and comorbidity. Clustering and heatmap displays of phenotype data, as performed in [22], support the visual identification of such patterns. However, this approach has limitations, e.g. the handling of missing data.

In general, missing data is frequently observed in large-scale mouse phenotyping. Reasons are: phenotyping may not have been performed due to animal welfare procedures. Far more frequent are scheduling issues: as every test is scheduled for a particular age of mice, no data can be taken if a test is skipped due to a broken phenotyping device. However, a certain missing data portion is counterbalanced by the

advantage of having a large, consistent data set that cannot be obtained by collecting data from individual labs in a meta study approach.

Conclusions

The complex logistics to organise high-throughput phenotyping, to manage large data sets and to ensure standardized and transparent QC and analysis of the data, is a major challenge in large-scale systemic phenotyping performed by mouse clinics. Electronic data management solutions support these processes. Large-scale data sets obtained in well-structured, quality controlled and standardized format enable further and comprehensive analysis of systemic phenotyping data, even across different centres. Employing data analysis tools linking phenotypic traits to known biological pathways and other information in public databases, aims for the discovery of disease-associated network of genes that can be investigated as a next step in more depth. Importantly, these data sets are the prerequisite to reach a new level in combining data sets from different species and disciplines to unravel the complexity of health and disease. Currently, about 1/4 of all mouse genes have phenotyping data in the IMPC project. However, phenotyping data collection continues while at the same time, automation (e.g. histological image analysis), data integration and analysis methods (e.g. machine learning) are further developed, investigated and improved. During the next years, this will result in a high quality and highly annotated comprehensive phenotyping data set for every mouse gene. The complete, publicly available data set and the associated methods and tools will provide a valuable ressource for big data projects involving mammalian gene function.

Acknowledgements

We would like to thank Manuela Östereicher for her critical comments on the manuscript and for preparing Figure 3. This work was funded by the German Federal Ministry of Education and Research (Infrafrontier grant 01KX1012).

References

Papers of particular interest have been highlighted as:

* of special interest

** of outstanding interest

1. Hrabe de Angelis MH, Flaswinkel H, Fuchs H, Rathkolb B, Soewarto D, Marschall S, Heffner S, Pargent W, Wuensch K, Jung M, et al.: **Genome-wide, large-scale production of mutant mice by ENU mutagenesis**. *Nat Genet* 2000, **25**:444-447.
2. Nolan PM, Peters J, Strivens M, Rogers D, Hagan J, Spurr N, Gray IC, Vizor L, Brooker D, Whitehill E, et al.: **A systematic, genome-wide, phenotype-driven mutagenesis programme for gene function studies in the mouse**. *Nat Genet* 2000, **25**:440-443.
3. Sabrautzki S, Rubio-Aliaga I, Hans W, Fuchs H, Rathkolb B, Calzada-Wack J, Cohrs CM, Klaften M, Seedorf H, Eck S, et al.: **New mouse models for metabolic bone diseases generated by genome-wide ENU mutagenesis**. *Mamm Genome* 2012, **23**:416-430.
4. Fuchs H, Gailus-Durner V, Adler T, Aguilar-Pimentel JA, Becker L, Calzada-Wack J, Da Silva-Buttkus P, Neff F, Gotz A, Hans W, et al.: **Mouse phenotyping**. *Methods* 2011, **53**:120-135.
5. Fuchs H, Gailus-Durner V, Adler T, Pimentel JA, Becker L, Bolle I, Brielmeier M, Calzada-Wack J, Dalke C, Ehrhardt N, et al.: **The German Mouse Clinic: a platform for systemic phenotype analysis of mouse models**. *Curr Pharm Biotechnol* 2009, **10**:236-243.
6. Fuchs H, Gailus-Durner V, Neschen S, Adler T, Afonso LC, Aguilar-Pimentel JA, Becker L, Bohla A, Calzada-Wack J, Cohrs C, et al.: **Innovations in phenotyping of mouse models in the German Mouse Clinic**. *Mamm Genome* 2012, **23**:611-622.
7. Gailus-Durner V, Fuchs H, Adler T, Aguilar Pimentel A, Becker L, Bolle I, Calzada-Wack J, Dalke C, Ehrhardt N, Ferwagner B, et al.: **Systemic first-line phenotyping**. *Methods Mol Biol* 2009, **530**:463-509.
8. Gailus-Durner V, Fuchs H, Becker L, Bolle I, Brielmeier M, Calzada-Wack J, Elvert R, Ehrhardt N, Dalke C, Franz TJ, et al.: **Introducing the German Mouse Clinic: open access platform for standardized phenotyping**. *Nat Methods* 2005, **2**:403-404.
9. Ayadi A, Birling MC, Bottomley J, Bussell J, Fuchs H, Fray M, Gailus-Durner V, Greenaway S, Houghton R, Karp N, et al.: **Mouse large-scale phenotyping initiatives: overview of the European Mouse Disease Clinic (EUMODIC) and of the Wellcome Trust Sanger Institute Mouse Genetics Project**. *Mamm Genome* 2012, **23**:600-610.
10. Brown SD, Chambon P, de Angelis MH, Eumorphia C: **EMPReSS: standardized phenotype screens for functional annotation of the mouse genome**. *Nat Genet* 2005, **37**:1155.
11. Mallon AM, Blake A, Hancock JM: **EuroPhenome and EMPReSS: online mouse phenotyping resource**. *Nucleic Acids Res* 2008, **36**:D715-718.

12. Morgan H, Beck T, Blake A, Gates H, Adams N, Debouzy G, Leblanc S, Lengger C, Maier H, Melvin D, et al.: **EuroPhenome: a repository for high-throughput mouse phenotyping data**. *Nucleic Acids Res* 2010, **38**:D577-585.
13. Brown SD, Moore MW: **The International Mouse Phenotyping Consortium: past and future perspectives on mouse phenotyping**. *Mamm Genome* 2012, **23**:632-640.
14. Brown SD, Moore MW: **Towards an encyclopaedia of mammalian gene function: the International Mouse Phenotyping Consortium**. *Dis Model Mech* 2012, **5**:289-292.
15. Brina D, Miluzio A, Ricciardi S, Clarke K, Davidsen PK, Viero G, Tebaldi T, Offenhauser N, Rozman J, Rathkolb B, et al.: **eIF6 coordinates insulin sensitivity and lipid metabolism by coupling translation to transcription**. *Nat Commun* 2015, **6**:8261.
16. Eden M, Meder B, Volkers M, Poomvanicha M, Domes K, Branchereau M, Marck P, Will R, Bernt A, Rangrez A, et al.: **Myoscape controls cardiac calcium cycling and contractility via regulation of L-type calcium channel surface expression**. *Nat Commun* 2016, **7**:11317.
17. Eisenberg T, Abdellatif M, Schroeder S, Primessnig U, Stekovic S, Pendl T, Harger A, Schipke J, Zimmermann A, Schmidt A, et al.: **Cardioprotection and lifespan extension by the natural polyamine spermidine**. *Nat Med* 2016, **22**:1428-1438.
18. Fuchs H, Sabrautzki S, Przemeck GK, Leuchtenberger S, Lorenz-Depiereux B, Becker L, Rathkolb B, Horsch M, Garrett L, Ostereicher MA, et al.: **The First Scube3 Mutant Mouse Line with Pleiotropic Phenotypic Alterations**. *G3 (Bethesda)* 2016, **6**:4035-4046.
19. Saarikangas J, Kourdougli N, Senju Y, Chazal G, Segerstrale M, Minkeviciene R, Kuurne J, Mattila PK, Garrett L, Holter SM, et al.: **MIM-Induced Membrane Bending Promotes Dendritic Spine Initiation**. *Dev Cell* 2015, **33**:644-659.
20. Thor T, Kunkele A, Pajtler KW, Wefers AK, Stephan H, Mestdagh P, Heukamp L, Hartmann W, Vandesompele J, Sadowski N, et al.: **MiR-34a deficiency accelerates medulloblastoma formation in vivo**. *Int J Cancer* 2015, **136**:2293-2303.
*21. Dickinson ME, Flenniken AM, Ji X, Teboul L, Wong MD, White JK, Meehan TF, Weninger WJ, Westerberg H, Adissu H, et al.: **High-throughput discovery of novel developmental phenotypes**. *Nature* 2016, **537**:508-514.

A third of all genes in the mammalian genome are important for survival. The characterization of the first 1751 mouse lines analyzed by the International Mouse Phenotyping consortium (IMPC) led to the discovery of 410 genes whose genetic deactivation impaired the development of embryos so strongly that they were not viable. In addition mutations in further 198 genes led to fewer offspring. Interestingly, many of the essential genes found here also play a key role in human diseases.

22. Hrabe de Angelis M, Nicholson G, Selloum M, White JK, Morgan H, Ramirez-Solis R, Sorg T, Wells S, Fuchs H, Fray M, et al.: **Analysis of mammalian gene function through broad-based phenotypic screens across a consortium of mouse clinics. *Nat Genet* 2015, **47**:969-978.

Studying gene function is the major goal of the European EUMODIC consortium. Knock-out mouse lines, that were deficient for one gene, were produced and analyzed in different mouse clinics. Disease-relevant organ-systems were analyzed for each line, allowing to gain insight in the function of the missing genes. The role of over 300 genes has been uncovered and the findings contribute to the understanding in the part they play in disease and biology.

23. West DB, Pasumarthi RK, Baridon B, Djan E, Trainor A, Griffey SM, Engelhard EK, Rapp J, Li B, de Jong PJ, et al.: **A lacZ reporter gene expression atlas for 313 adult KOMP mutant mouse lines**. *Genome Res* 2015, **25**:598-607.

24. Karp NA, Mason J, Beaudet AL, Benjamini Y, Bower L, Braun RE, Brown SDM, Chesler EJ, Dickinson ME, Flenniken AM, et al.: **Prevalence of sexual dimorphism in mammalian phenotypic traits. *Nat Commun* 2017, **8**:15475.

The IMPC has an invaluable amount of phenotyping data produced under exact standardized conditions. In this new study, researchers have quantified the difference between male and female mice, looking across multiple experiments and institutes. In the largest study of its kind, scientists analysed up to 234 characteristics of more than 50,000 mice. The data show that sex influences the prevalence, course and severity of the majority of common diseases and disorders.

25. Meehan TF, Conte N, West DB, Jacobsen JO, Mason J, Warren J, Chen CK, Tudose I, Relac M, Matthews P, et al.: **Disease model discovery from 3,328 gene knockouts by The International Mouse Phenotyping Consortium. *Nat Genet* 2017.

Mouse models contribute substantially to a better understanding of human diseases. The International Mouse Phenotyping Consortium has characterized the function of 3,328 genes and identified 360 new disease models. Many new disease models reflect characteristics of human diseases and can facilitate the investigation of molecular

mechanisms and the development of new therapies. In addition the scientists find unknown genes mimicking symptoms of human disorders. For many genes new information in different organ systems was uncovered, showing that the systemic approach is crucial to observe the pleiotropy of genes.

26. Koscielny G, Yaikhom G, Iyer V, Meehan TF, Morgan H, Atienza-Herrero J, Blake A, Chen CK, Easty R, Di Fenza A, et al.: **The International Mouse Phenotyping Consortium Web Portal, a unified point of access for knockout mice and related phenotyping data**. *Nucleic Acids Res* 2014, **42**:D802-809.
27. Ring N, Meehan TF, Blake A, Brown J, Chen CK, Conte N, Di Fenza A, Fiegel T, Horner N, Jacobsen JO, et al.: **A mouse informatics platform for phenotypic and translational discovery**. *Mamm Genome* 2015, **26**:413-421.
28. Karp NA, Meehan TF, Morgan H, Mason JC, Blake A, Kurbatova N, Smedley D, Jacobsen J, Mott RF, Iyer V, et al.: **Applying the ARRIVE Guidelines to an In Vivo Database**. *PLoS Biol* 2015, **13**:e1002151.
29. Maier H, Schutt C, Steinkamp R, Hurt A, Schneltzer E, Gormanns P, Lengger C, Griffiths M, Melvin D, Agrawal N, et al.: **Principles and application of LIMS in mouse clinics**. *Mamm Genome* 2015, **26**:467-481.
30. Mallon AM, Iyer V, Melvin D, Morgan H, Parkinson H, Brown SD, Flicek P, Skarnes WC: **Accessing data from the International Mouse Phenotyping Consortium: state of the art and future plans**. *Mamm Genome* 2012, **23**:641-652.
31. R Core Team: **R: A language and environment for statistical computing**. Edited by: R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. 2013.
32. Kurbatova N, Mason JC, Morgan H, Meehan TF, Karp NA: **PhenStat: A Tool Kit for Standardized Analysis of High Throughput Phenotypic Data**. *PLoS One* 2015, **10**:e0131274.
33. Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R, Eilbeck K, Lewis S, Marshall B, Mungall C, et al.: **The Gene Ontology (GO) database and informatics resource**. *Nucleic Acids Res* 2004, **32**:D258-261.
34. Smith CL, Eppig JT: **The Mammalian Phenotype Ontology as a unifying standard for experimental and high-throughput phenotyping data**. *Mamm Genome* 2012, **23**:653-668.
35. Kohler S, Doelken SC, Mungall CJ, Bauer S, Firth HV, Bailleul-Forestier I, Black GC, Brown DL, Brudno M, Campbell J, et al.: **The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data**. *Nucleic Acids Res* 2014, **42**:D966-974.
36. Kohler S, Vasilevsky NA, Engelstad M, Foster E, McMurry J, Ayme S, Baynam G, Bello SM, Boerkoel CF, Boycott KM, et al.: **The Human Phenotype Ontology in 2017**. *Nucleic Acids Res* 2017, **45**:D865-D876.
37. The Gene Ontology C: **Expansion of the Gene Ontology knowledgebase and resources**. *Nucleic Acids Res* 2017, **45**:D331-D338.
38. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K: **KEGG: new perspectives on genomes, pathways, diseases and drugs**. *Nucleic Acids Res* 2017, **45**:D353-D361.

39. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M: **KEGG as a reference resource for gene and protein annotation**. *Nucleic Acids Res* 2016, **44**:D457-462.
40. Smedley D, Oellrich A, Kohler S, Ruef B, Sanger Mouse Genetics P, Westerfield M, Robinson P, Lewis S, Mungall C: **PhenoDigm: analyzing curated annotations to associate animal models with human diseases**. *Database (Oxford)* 2013, **2013**:bat025.
41. Forouhi NG, Wareham NJ: **Epidemiology of diabetes**. *Medicine (Abingdon)* 2014, **42**:698-702.
42. Despres JP, Lemieux I: **Abdominal obesity and metabolic syndrome**. *Nature* 2006, **444**:881-887.

Figure and Table Legends

Table 1

Overview of the primary phenotyping pipeline performed by the German Mouse Clinic

The tests of the primary phenotyping pipeline cover all relevant organ systems in order to have a full picture of disease-associated alterations a deficient gene might cause in the organism. The first column ("Screens") names the principal physiological field respectively the organ system, the second column ("Methods") specifies the applied phenotyping procedure, e.g. IpGTT (Intraperitoneal glucose tolerance test). Columns 3-11 indicate the age (in weeks) in which mice are subjected to the particular procedure (marked by "x"). For example, IpGTT is applied at age 14 weeks. Optional tests are presented in light grey.

Figure 1

Systemic phenotyping as a process

Systemic phenotyping is not an isolated task, but a complex process embedded in a whole process workflow. Mouse clinic operations can be described and handled in form of modular processes to cope with the complex logistics needed for the different areas. In the top row, basic processes of systemic phenotypic are depicted in boxes, connected by arrows. In the middle row, processes are described in more details, starting from timely production of age-matched mutant and control cohorts, performing the actual phenotyping procedures according to the pipeline and finally the assignment of MP terms after data quality control and analysis. Throughout the whole process, tight scheduling and tracking of activities and resources is necessary (bottom row) in order to identify workflow problems and to ensure continuous data flow.

Figure 2

Phenomap of the German Mouse Clinic

The phenotyping results of mutant mice analysed at the German Mouse Clinic are provided to the scientific community by presenting them in the form of a so-called phenomap. A tabular phenomap provides a very condensed qualitative summary of phenotypes for a set of mutant mouse lines, details of which are given in the first three columns. The other columns represent the investigated physiological fields respectively organ systems. The colour fields in the matrix depict the qualitative findings of the tests performed (clear/no/subtle phenotype) according to the legend (top). A graph symbol indicates a link to more detailed graphs and data. For example, clear genotype-related differences could be identified in Cap2-deficient mice in the behaviour and neurology fields (red cells, bottom row). Filters (table header row) can be used to identify mutant lines that show a particular phenotype of interest in one or more physiological fields.

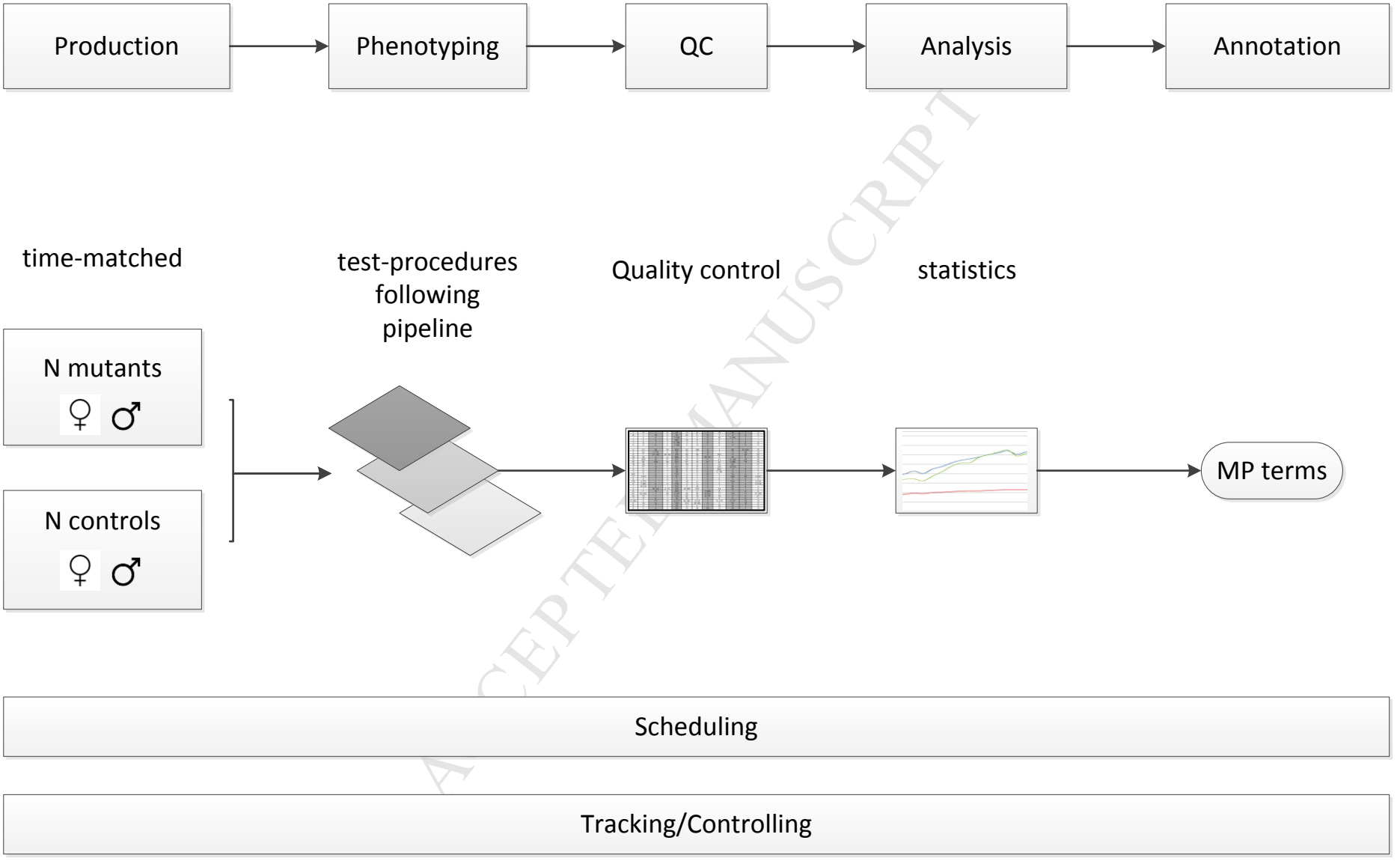
Figure 3

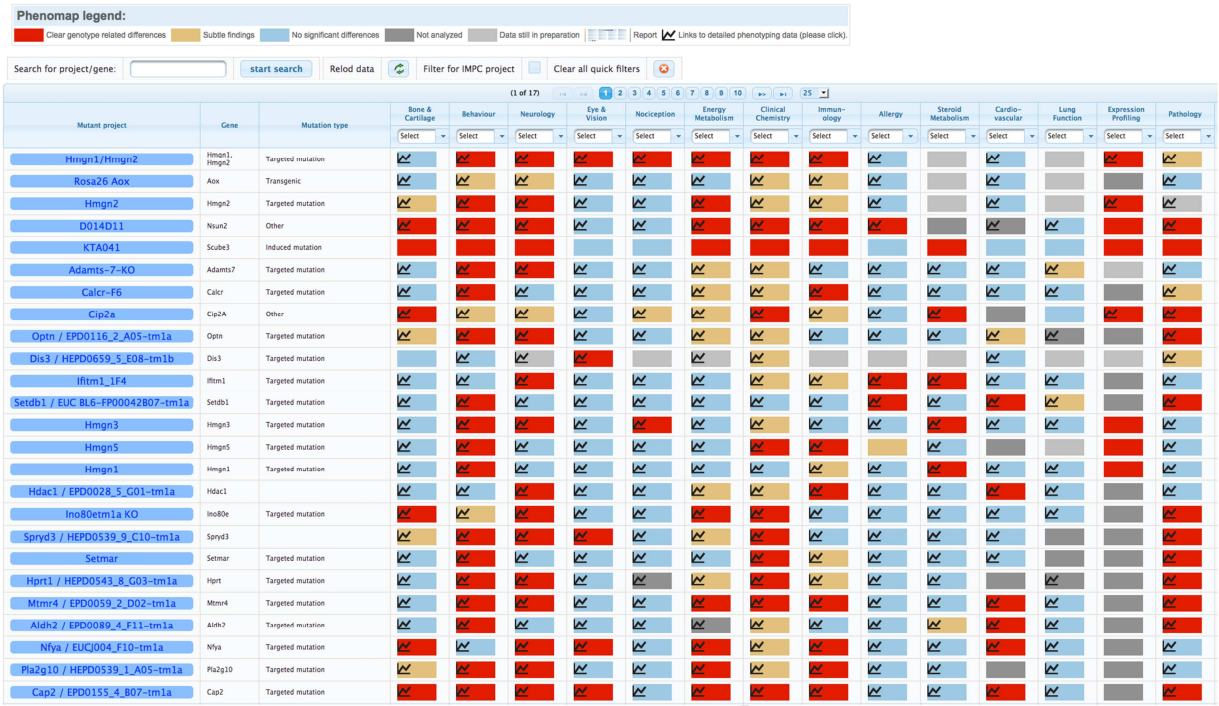
Mutant/control ratio histogram as a tool to identify phenotypic hits in big data sets

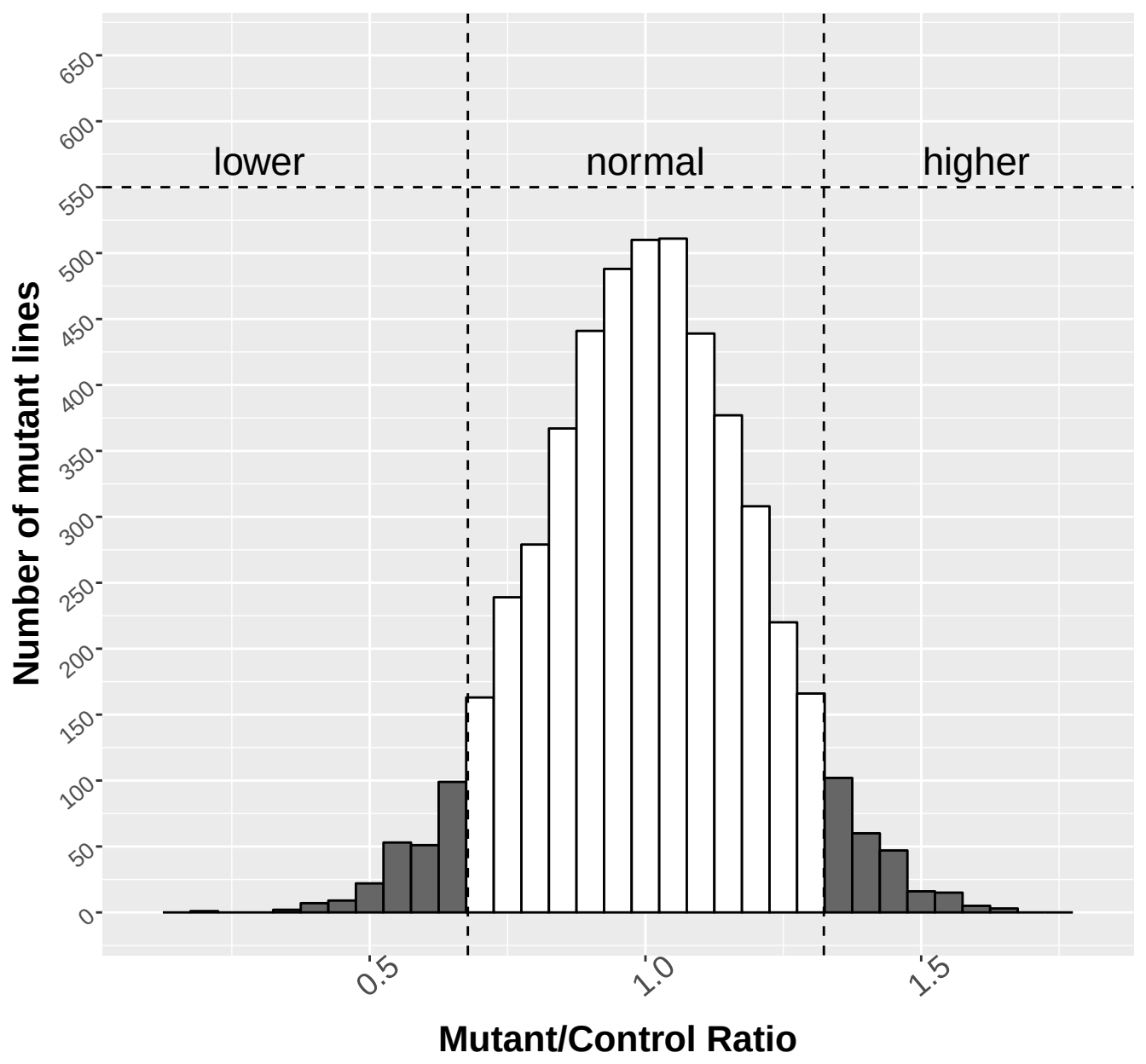
This quantitative approach can be applied for interval-scaled phenotype parameters. For each parameter, the ratio of the mean value for mutant and control mice is calculated. Performed on a large number of mutant lines, this can be used to produce a histogram with "Mutant/Control Ratio" on the x-axis and the respective "Number of mutant lines" on the y-axis. Please note: the shown histogram is based on simulated data. Ratios around 1.0 show no strong difference between mutants and controls and the respective mutant lines are therefore classified as "normal phenotype" (top middle). Defined thresholds (dashed vertical lines) are used to identify mutant lines showing an unusual high or low ratio (top left/top right) for a given parameter of interest. The definition of upper and lower thresholds can be based on different criteria, e.g. prior knowledge about biological relevance or mere statistical considerations (percentile-based thresholds).

	Age (weeks)	9	10	11	12	13	14	15	16	17	18	19
Screens	***Methods***											
Behavior	Open Field	X										
	Acoustic startle response & PPI		X									
Neurology	Modified SHIRPA, grip strength	X										
	Rotarod		X									
Clinical Chemistry	Clinical Chemistry after fasting			X								
Nociception	Hot Plate				X							
Dysmorphology	Anatomical observation				X							
Allergy	Transepidermal water loss (TEWL), body surface temperature				X							
Energy Metabolism	Indirect calorimetry, NMR					X						
Clinical Chemistry	IpGTT						X					
Cardiovascular	Awake ECG/Echocardiography							X				
Eye	Scheimpflug imaging, OCT, LIB, drum								X			
Neurology	ABR (Auditory brain stem response)									X		
Dysmorphology	X-ray, DEXA									X		
Energy Metabolism	NMR										X	
Clinical Chemistry	Clinical Chemical analysis, hematology											X
Immunology	FACS analysis of PBCs											X
Allergy	BIOPLEX ELISA (IG concentration)											X
Steroid Metabolism, optional	Corticost., Androst., Testosterone											X
Lung Function, optional	Lung function measurements											X
Molecular Phenotyping	Expression profiling											X
Pathology	Macro & microscopic analysis											X

Abbreviations: PPI – prepulse inhibition; SHIRPA - SmithKline Beecham, Harwell, Imperial College, Royal London Hospital, phenotype assessment; NMR – nuclear magnetic resonance; IpGTT - intraperitoneal glucose tolerance Test; ECG – electrocardiography; OCT - optical coherence tomography; LIB - laser interference biometry ; DXA - dual-energy X-ray absorptiometry; FACS - fluorescence- activated cell sorting; PBC – peripheral blood cell ; IG - immunoglobulin






lower
normal
higher
Number of mutant lines
0
50
100
150
200
250
300
350
400
450
500
550
600
650
0.5
1.0
1.5
Mutant/Control Ratio

Highlights

- Importance of standardized large-scale phenotyping of mice
- High-throughput phenotyping as a logistic challenge
- Data management solutions and standardised procedures for quality control are crucial
- Working with systemic phenotyping data
- Systemic standardized phenotyping data of mice as basis for translational approaches