

# EuroPhenome: a repository for high-throughput mouse phenotyping data

Hugh Morgan<sup>1</sup>, Tim Beck<sup>1</sup>, Andrew Blake<sup>1</sup>, Hilary Gates<sup>1</sup>, Niels Adams<sup>2</sup>, Guillaume Debouzy<sup>3</sup>, Sophie Leblanc<sup>3</sup>, Christoph Lengger<sup>4</sup>, Holger Maier<sup>4</sup>, David Melvin<sup>5</sup>, Hamid Meziane<sup>3</sup>, Dave Richardson<sup>5</sup>, Sara Wells<sup>2</sup>, Jacqui White<sup>5</sup>, Joe Wood<sup>2</sup>, The EUMODIC Consortium<sup>†</sup>, Martin Hrabé de Angelis<sup>4,6</sup>, Steve D. M. Brown<sup>1</sup>, John M. Hancock<sup>1</sup> and Ann-Marie Mallon<sup>1,\*</sup>

<sup>1</sup>MRC Harwell, Mammalian Genetics Unit, <sup>2</sup>MRC Harwell, Mary Lyon Centre, Harwell Science and Innovation Campus, Oxfordshire, OX11 0RD, UK, <sup>3</sup>Institut Clinique de la Souris (ICS), 1 Rue Laurent Fries, BP10142, 67404 Illkirch Cedex, France, <sup>4</sup>Helmholtz Zentrum Munich German Research Center for Environmental Health, Institute of Experimental Genetics and German Mouse Clinic, Ingolstaedter Landstrasse 1, D-85764 Neuherberg, Germany, <sup>5</sup>The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK and <sup>6</sup>Technische Universität München, Center of Life and Food Sciences Weihenstephan, Alte Akademie 8, 85354 Freising, Germany

Received September 22, 2009; Revised and Accepted October 16, 2009

## ABSTRACT

The broad aim of biomedical science in the postgenomic era is to link genomic and phenotype information to allow deeper understanding of the processes leading from genomic changes to altered phenotype and disease. The EuroPhenome project (<http://www.EuroPhenome.org>) is a comprehensive resource for raw and annotated high-throughput phenotyping data arising from projects such as EUMODIC. EUMODIC is gathering data from the EMPReSSslim pipeline (<http://www.empress.har.mrc.ac.uk/>) which is performed on inbred mouse strains and knock-out lines arising from the EUComm project. The EuroPhenome interface allows the user to access the data via the phenotype or genotype. It also allows the user to access the data in a variety of ways, including graphical display, statistical analysis and access to the raw data via web services. The raw phenotyping data captured in EuroPhenome is annotated by an annotation pipeline which automatically identifies statistically different mutants from the appropriate baseline and assigns ontology terms for that specific test. Mutant phenotypes can be quickly identified using two EuroPhenome tools: PhenoMap, a graphical representation of statistically relevant phenotypes,

and mining for a mutant using ontology terms. To assist with data definition and cross-database comparisons, phenotype data is annotated using combinations of terms from biological ontologies.

## INTRODUCTION

The laboratory mouse is fundamental to the understanding of mammalian biology and the investigation of human disease (1,2). There is now a wealth of genotypic information available as a result of the sequencing the mouse genome, and a central challenge in the identification of mouse disease models is the attribution of function to individual genes. The classical genetics approach of relating genotype to phenotype—the introduction of mutations into the genome of an experimental organism and the identification of novel phenotypes—can be furthered on a high-throughput scale with the establishment of projects devoted to the manufacture of mouse gene knockout lines. The International Knockout Mouse Consortium (IKMC; <http://www.knockoutmouse.org>) (3) is made up of the four major projects: (i) EUComm (<http://www.eucomm.org>), (ii) KOMP (<http://www.nih.gov/science/models/mouse/knockout/>), (iii) TIGM (<http://www.tigm.org>) and (iv) NorCOMM (<http://www.norcomm.org>) and is in the process of generating mutations in ES cells for all known protein coding genes.

\*To whom correspondence should be addressed. Tel: 01235 841077; Fax: 01235 841172; Email: a.mallon@har.mrc.ac.uk

<sup>†</sup>The members of the EUMODIC Consortium are listed in the Appendix 1.

The European Mouse Disease Clinic programme (EUMODIC; <http://www.eumodic.org>) in collaboration with EUCOMM (4) is in the process of generating mouse mutant lines from ES cell resource and aims to provide a phenotype assessment of 500 mouse mutant lines. The Eumorphia programme (<http://www.eumorphia.org>) (5) was the predecessor to EUMODIC and generated baseline phenotype data for four inbred mouse lines besides developing the European Mouse Phenotyping Resource for Standardised Screens (EMPreSS) (6,7), which can be utilized for mouse phenotyping. The EuroPhenome Resource (<http://www.EuroPhenome.org>) originally contained and presented the baseline data derived during Eumorphia (8). Our previous article in NAR described this data and the web tools we had developed to analyse and display them. Since this initial effort, EuroPhenome has undergone significant development; most notable is that EuroPhenome now captures and annotates mutant mouse phenotyping data from EUMODIC, presenting putative phenovariants to users.

Three critical steps have been achieved within the EUMODIC programme, which allow the comprehensive analysis and interpretation of the generated phenotype data. First, the phenotype data is generated using community-agreed phenotyping platforms that provide standardized methods, so that the results can be compared between laboratories and over time. All of the phenotyping procedures executed during the EUMODIC screens are part of the publicly available EMPReSS resource (<http://empres.har.mrc.ac.uk>) and organized into two phenotyping pipelines, collectively termed EMPReSSslim (<http://www.eumodic.org/EMPreSSslim.html>). Second, the raw phenotype data, which is generated by the participating research institutions, is collected within EuroPhenome which serves as a community database with open access. And third, ontologies provide the structured description of the phenotypes within EuroPhenome to allow the data to be interpreted in a consistent manner.

## DATA IN EUROPHENOME

EMPreSSslim consists of 20 phenotyping screens described by standard operating procedures (SOPs). These examine the broad biological domains of body morphology; metabolism, heart morphology and blood pressure, body fat and bone composition and skeletal morphology, behaviour and sensory testing, haematology; clinical chemistry, FACs analysis of leukocytes and immunoglobulin concentration. Data measured in each SOP are recorded as individual parameters. The parameters for each screen are given a unique identifier. These are listed in EMPReSS. The parameters are of two types: phenotype parameters and metadata parameters. The phenotype parameter defines a characteristic of the mouse which can be measured and so is analogous to a mouse trait. The data obtained for each phenotype parameter during the execution of the SOP is captured and analysed within EuroPhenome. Metadata parameters

capture technical aspects of the SOP which could make an impact in the comparison of phenotype parameters between executions of the same SOP, for example equipment models and settings. Currently, 622 phenotype parameters are recorded in the EMPReSSslim screen. EuroPhenome contains data for 76 mutant mouse strains (in some cases representing alleles in both homozygous and heterozygous states), making a total of 2.1 million data points.

The mouse clinics (Institut Clinique de la Souris, France; Helmholtz Zentrum Munich, Germany; MRC Harwell; and Wellcome Trust Sanger Institute, UK) which are part of EUMODIC process mutant mouse lines through EMPReSSslim, capturing the phenotype data in their local Laboratory Information Management Systems (LIMS). The database architectures of the various LIMS differ markedly, so a common data format is needed for the exchange of data from the LIMS to EuroPhenome. An XML data format is used for this purpose whereby the phenotype data for each individual mouse making up a mutant strain cohort, which for the EUMODIC project is defined as seven or more individuals, can be submitted for each procedure carried out. Each procedure is defined using the Minimal Information for Mouse Phenotyping Procedures (MIMPP) standard, which was developed as part of the larger community-wide effort for minimum information for biological and biomedical investigations (MIBBI) (9), and is captured using its associated XML data format. The phenotype data XML files are uploaded to the institution's FTP site which EuroPhenome scans daily for new data. Each uploaded XML file is validated to ensure all required data points are present and the data for each parameter is of the correct type.

## ANNOTATION PIPELINE

An SOP measures parameters which produce direct primary phenotype measurements. In combination with other primary phenotypes, these can be used to infer qualitative or descriptive phenotypes. For example, data for a mutant line for the hotplate test might include a time in seconds before the first response and the observation of a recognized response to heat. Where the time in seconds is less than the comparative background strain, the phenotype of 'decreased thermal nociceptive threshold' could be inferred. EuroPhenome describes the traits and phenotypes using ontologies, where both the pre-composed Mammalian Phenotype (MP) ontology (10) and the post-composed Entity and Quality (EQ) PATO (11) model are employed. The phenotype parameter (or mouse trait) is defined using EQ; for example, 'coat hair' + 'texture'. The resulting phenotype is also defined using EQ where the phenotype quality is a child term to parameter quality within the PATO ontology; for example 'coat hair' + 'greasy'. The quality terms 'texture' and 'greasy' share a parent-to-child relationship within PATO. Currently EuroPhenome annotates phenotypes using MP terms, where each MP term is the closest equivalent to the EQ phenotype description. The EQ

annotations of parameters and the EQ and MP annotations of phenotypes are available at <http://www.EuroPhenome.org/parameters.html>.

EuroPhenome performs statistical comparisons of mutant line phenotype parameter data to baseline inbred strain phenotype parameter data to identify phenovariants. The statistical comparison takes place between related datasets only when critical metadata parameters are identical. Pearson's Chi-square test of independence is applied to categorical data and the Student's *t*-test and the Mann-Whitney *U* test are applied to quantitative data. These tests were chosen because they are familiar to the community and computationally tractable for the size of data set. Phenovariants are mutant lines which show a phenotype to be present to a greater or lesser extent compared to the baseline inbred strains. EuroPhenome stores all possible annotations and automatically applies the relevant MP term to the mutant line depending on the direction (increased or decreased) of the phenovariant. For a more detailed description of the use of ontologies in the annotation of EuroPhenome data, refer Beck *et al.* (12).

## EUROPHENOME FUNCTIONALITY

The majority of users access EuroPhenome through the web interface and so a critical aim of the project has been to develop visual methods enabling quick and easy data access. The complexity and size of data generated from high-throughput phenotyping projects is a new challenge for bioinformatics and hence new methods and approaches to displaying this data are required.

EuroPhenome allows users to access the raw data and the annotated data described above through three new integrated web tools; the 'Phenome Data Viewer', the 'PhenoMap' and the 'Ontology Tree' Tool. In our previous NAR paper, we described the data browser for the visualization of the Eumorphia data; this browser is still available through the EuroPhenome portal by clicking on the 'Eumorphia data browser' link.

### Phenome Data Viewer

The Phenome Data Viewer is the primary tool enabling users navigate through the raw phenotyping data from both the inbred mouse strains and the mutant strains to view dynamically generated graphs and overview statistics built from user entered selections. The type of graph displayed is preset depending on the parameter type: single quantitative parameters are displayed as bar charts, time series quantitative parameters are shown as line graphs and qualitative data is displayed in a stacked bar graph.

The left-hand panel of the viewer displays a number of selection boxes which the user can click to build the specific query of interest. The inbred mouse data option presents users with a number of choices (centre, strain, sex and procedure) allowing them to build the specific query they are interested in; for example, to compare data from an EMPReSSlim procedure between two inbred strains, the user would choose the strains of interest in

the Strains selection box (e.g. C57BL/6NTac and C57BL/6J), the procedure and parameter/s from the relevant selection boxes (e.g. IPGTT and Blood glucose concentration) and click view data. The data view returned would be a line graph showing how blood glucose concentration varied over time (Figure 1a). This data view enables the user to drill further into the data and view the specific metadata parameters used when the data was collected, download the raw data used in generating the graph as well as viewing overview statistics.

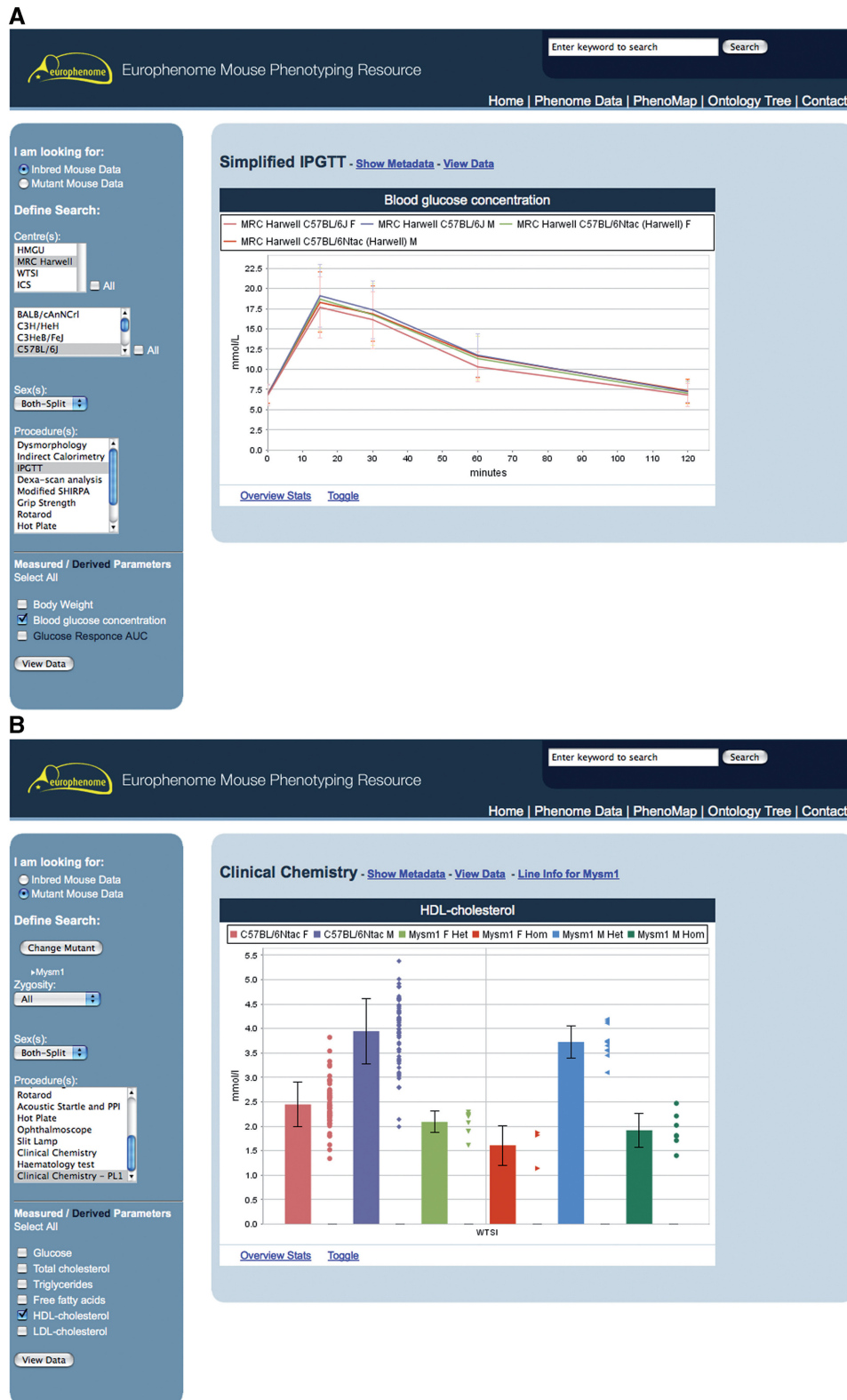
The mutant mouse data viewer enables the user to choose one or more mutants of interest from the 'Select Mutants' list, before selecting the zygosity of interest, sex and the procedure and parameter in the same way as before (e.g. Mysm1, Clinical Chemistry PL1 and HDL cholesterol). The data view returned to the user comprises a graph, which now shows the results for the relevant baseline strain as well as the mutant line split by zygosity and sex, if chosen. In addition to the graph, the user can access overview statistics, which include the mean, standard deviation, standard error, results from the statistical tests shown as 'Not significant, Significant, Highly Significant' as well as highlighting differences in metadata (Figure 1b). The levels for these annotations are: not significant,  $P > 0.05$ ; significant,  $P < 0.05$ ; and highly significant,  $P < 0.001$ .

### PhenoMap

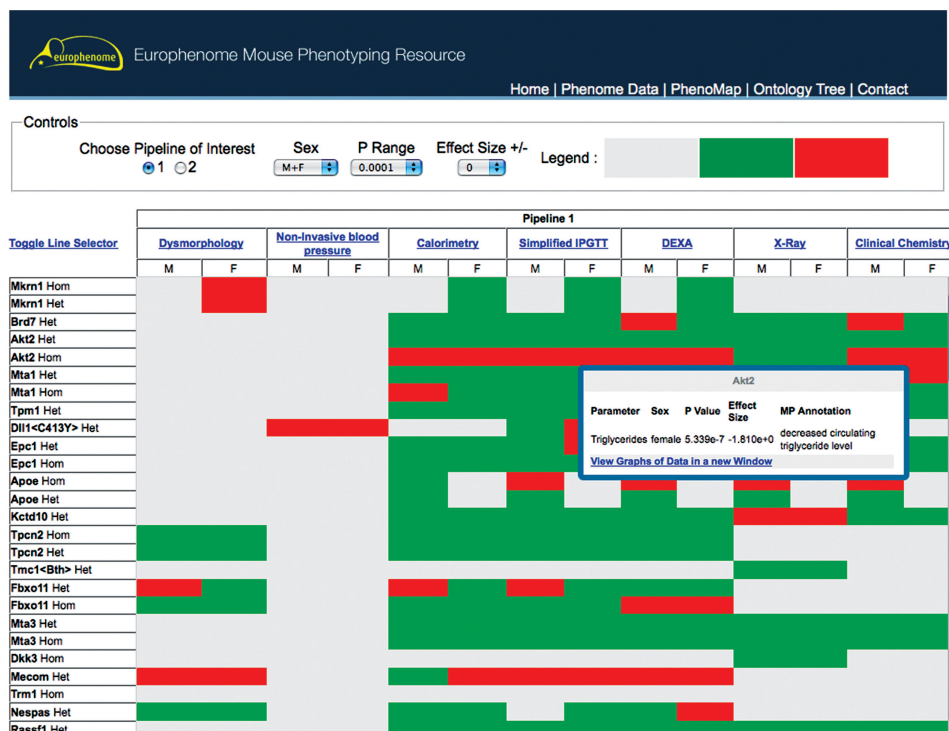
Interaction with the scientific user community in EUMODIC identified that the majority of users of EuroPhenome would want to ask 'Can I quickly find a phenovariant/s with the phenotype/s of interest?'. The PhenoMap (Figure 2) is a graphical representation of statistically significant phenovariants produced from the annotation pipeline described above. The left-hand panel of the PhenoMap allows user to view all mutant lines in EuroPhenome or to choose those of interest. The top panel has functionality for the user to choose the phenotyping pipeline (relating to EMPReSSlim) of interest, the sex, the *P* value range and the effect size of interest. Grey boxes in the map depict a procedure for which we currently do not hold data, the green boxes depict when data is available but no significant annotations were found at the current *P*-value and effect size. The red boxes show a significant annotation at the current *P*-value and effect size. Further information about the significant annotation can be easily accessed by hovering over the red box. At this point, the user can jump directly to the relevant graph in the phenome viewer. A user can also drill down into a specific procedure to view the annotations at a parameter level by clicking on the procedure name.

### Ontology Tree

The 'Ontology Tree' tool enables the user to move through the MP ontology tree to view ontology annotations produced from the annotation pipeline for each line (Figure 3). Selecting a MP term of interest displays a table in the right-hand panel showing the mutant line, the procedure, the parameter and the



**Figure 1.** This interface shows the results of two queries using the Phenome Data Viewer. **(A)** The query builder menu on the left-hand side and the results panel on the right which displays a line graph of Blood glucose concentration from the Intraperitoneal glucose tolerance test (IPGTT) procedure performed on inbred mouse strains C57BL/6Ntac and C57BL/6J. **(B)** The same query builder for mutant data on the left-hand side and on the right, the bar chart graph of HDL-cholesterol for the Mym-1 mutant.



**Figure 2.** This interface shows the EuroPhenome PhenoMap tool, which is a heatmap representation of statistically significant phenovariants produced from the annotation pipeline. The top panel is an interactive tool to change the phenotyping pipeline (relating to EMPReSS) of interest, the sex, the *P* range and the effect size. Grey boxes in the map depict a procedure for which we currently do not hold data, the green boxes depict when data is available but no significant annotations were found at the current *P*-value and effect size. The red boxes show a significant annotation at the current *P*-value and effect size. Further details on red boxes can be obtained by hovering over the box.

ontology term it was annotated with. Again, the user can quickly jump to the phenome data viewer graph of interest by clicking on the graph link.

## IMPLEMENTATION

The XML generation, validation and export is automated using a Java library available under the GPL at <http://sourceforge.net/projects/EuroPhenome/>. This is implemented at the mouse clinics and allows a common process of validation and XML generation as well as minimizing the work required to export data to EuroPhenome.

When new data is entered into EuroPhenome, a number of processes are automatically performed. The annotation pipeline is implemented as a Java application that uses the same code that generates the statistical information on the website. If a statistically significant phenovariant is discovered, the data is inserted into a database called the Ontology Annotation system at Harwell (OATH). This stores information about the mutant (such as gene affected and zygosity of group showing difference), the parameter and detected trait along with the results of the statistical test. This allows the ontology visualization tools to tailor the result to the user-specified values.

If new mutant lines are uploaded, a number of resources are used to collect the genotype information. These include the external web services of EUCOMM, EMMA

and KERMITs for EUCOMM lines and an internal copy of the MGI database for non-EUCOMM lines.

The EuroPhenome database is implemented in MySQL (<http://www.mysql.com/>) running on Solaris, and interacts with the OATH and EMPReSS databases running on the same platform. The Phenome Data Viewer is implemented as a Java Servlet and the PhenoMap and Ontology Tree are implemented in PHP. The website is styled using css. The Ontology Tree uses DHTML and AJAX to provide the relevant information as the tree is traversed, and an auto suggest function to find a term of interest. The PhenoMap uses JavaScript to provide user functionality, including the Prototip library to provide the hover-over information. The Phenome Data Viewer uses JavaScript to provide the user filter selection, including providing the user with only relevant options depending on their selection. The images are generated by JFreeChart (<http://www.jfree.org/jfreechart/>). Memcached (<http://www.danga.com/memcached/>) is used to improve performance.

The statistics are calculated using the Apache Commons Math library for the Pearson's chi-square test, Student's *t*-test and the descriptive statistics, and the Java Statistical Classes library for the Mann-Whitney *U*-test.

## FUTURE DIRECTIONS

The systematic acquisition of phenotype data on mutant mouse lines will be central to the functional annotation of

Ontology : (MP) Mammalian Phenotype

Term Search =>

- adipose tissue phenotype(0)
- behavior/neurological phenotype(35)
- cardiovascular system phenotype(1)
- cellular phenotype(0)
- craniofacial phenotype(3)
- digestive/alimentary phenotype(1)
- embryogenesis phenotype(0)
- endocrine/exocrine gland phenotype(0)
- growth/size phenotype(23)
- hearing/vestibular/ear phenotype(7)
- hematopoietic system phenotype(2)
- homeostasis/metabolism phenotype(19)
- immune system phenotype(2)
- lethality-postnatal(0)
- lethality-prenatal/perinatal(0)
- life span-post-weaning/aging(0)
- limbs/digits/tail phenotype(6)
- liver/biliary system phenotype(0)
- muscle phenotype(0)
- nervous system phenotype(0)
- no phenotypic analysis(0)
- normal phenotype(0)
- other phenotype(38)
- pigmentation phenotype(0)
- renal/urinary system phenotype(0)
- reproductive system phenotype(0)
- respiratory system phenotype(0)
- skeleton phenotype(8)
- skin/coat/nails phenotype(0)
- taste/olfaction phenotype(0)
- touch/vibrissae phenotype(5)
- tumorigenesis(0)
- vision/eye phenotype(1)

Annotations					
Gene	Internal Id	Procedure	Parameter	Term	Data
Mir96	ENSMUSG00000065586Dmd0	Modified SHIRPA	Startle Response	abnormal startle reflex	Graph
Ptk7	C3H.Cg-Ptk7 <sup>chz</sup> /H	Modified SHIRPA	Pinna Reflex	absent pinna reflex	Graph
Fbxo11	C3H.Cg-Fbxo11 <sup>fl</sup> /H	Modified SHIRPA	Startle Response	abnormal startle reflex	Graph
Mecom	C3H.Cg-Evi1 <sup>bb</sup> /H	Modified SHIRPA	Startle Response	abnormal startle reflex	Graph
Kcnj16	BB6J.129-Kcnj16 <sup>tm1Sru</sup> /H	Modified SHIRPA	Startle Response	abnormal startle reflex	Graph
Trpc3	C3H.Cg-Trpc3 <sup>Mw</sup> /H	Modified SHIRPA	Pinna Reflex	absent pinna reflex	Graph
Tg(Thy1-MAPT <sup>P301S</sup> )2541Godt	B6J-Tg(Thy1-MAPT <sup>P301S</sup> )	Modified SHIRPA	Startle Response	abnormal startle reflex	Graph

**Figure 3.** This interface shows the EuroPhenome Ontology Tree tool, with the MP ontology tree on the left-hand panel and six mutants on the right-hand panel, that have been identified from the annotation pipeline as having 'hearing/vestibular/ear phenotypes'. A direct link to the graphical display for the procedures can be accessed from the data column.

the mouse genome over the coming decades (13,14). These data will be equivalent to that produced by genome projects in providing the fundamental information upon which detailed characterization of mammalian gene function can be built. Because of the fundamental role of these resources there is a strong move to make both material mouse resources and data about them freely accessible (15). The EUMODIC consortium is laying the foundations of this phenotyping effort but wider plans are being developed to build an International Mouse Phenotype Consortium (IMPC) to carry out this systematic phenotyping.

Both physical and bioinformatics infrastructure will be needed to support such a major project. EuroPhenome represents the first step towards developing the bioinformatics infrastructure. Major developments can be expected in future in ways of collecting, representing and visualizing phenotype data in conjunction with other data, such as genomic data, to make them more readily interpretable by the bench biologist. Novel modes of data analysis, including a more refined approach to statistical analysis, will also be needed to provide the maximum value from phenotyping data. EuroPhenome aims to extend its scope beyond EUMODIC to allow systematically acquired phenotype data from different sources to be

presented within the same framework, adding value to all of these data sets by integrating them with others.

## ACKNOWLEDGEMENTS

The authors thank the members of the EUMODIC Bioinformatics workpackage, the EUMODIC consortium and the members of the mouse clinics who worked very hard to produce and export the data. They also thank their IT colleagues at EMMA and EUCOMM for feedback on EuroPhenome and help with data integration and Karen Pickford, the scientific curator at Harwell, who helped to ensure the nomenclature of the mutant lines is correct.

## FUNDING

European Commission within its FP6 Programme (contract number LSHG-CT-2006-037188 to EUMODIC; UK Medical Research Council and the NGFN (National Genome Research Network) (grant number 01GS0850). Funding for open access charge: EUMODIC FP6 Programme.

*Conflict of interest statement.* None declared.

## REFERENCES

- Hancock, J.M., Mallon, A.M., Beck, T., Gkoutos, G.V., Mungall, C. and Schofield, P.N. (2009) Mouse, man, and meaning: bridging the semantics of mouse phenotype and human disease. *Mamm Genome*, **8**, 457–461.
  - Rosenthal, N. and Brown, S. (2007) The mouse ascending: perspectives for human-disease models. *Nat. Cell Biol.*, **9**, 993–999.
  - International Mouse Knockout Consortium. Collins, F.S., Rossant, J. and Wurst, W. (2007) A mouse for all reasons. *Cell*, **128**, 9–13.
  - Friedel, R.H., Seisenberger, C., Kaloff, C. and Wurst, W. (2007) EUCOMM—the European conditional mouse mutagenesis program. *Brief Funct. Genomic Proteomic*, **6**, 180–185.
  - Brown, S.D., Lad, H.V., Green, E.D., Gkoutos, G., Gates, H., Hrabe de Angelis, H. and Consortium, T.E. (2006) EUMORPHIA and the European mouse phenotyping resource for standardised screens (EMPreSS). In Hrabe de Angelis, M.M., Chambon, P. and Brown, S. (eds), *Standards of Mouse Model Phenotyping*. Wiley, pp. 311–320.
  - Brown, S.D., Chambon, P. and de Angelis, M.H. (2005) EMPreSS: standardized phenotype screens for functional annotation of the mouse genome. *Nat. Genet.*, **37**, 1155.
  - Green, E.C.J., Gkoutos, G.V., Lad, H.V., Blake, A., Weekes, J. and Hancock, J.M. (2005) EMPreSS: European mouse phenotyping resource for standardised screens. *Bioinformatics*, **21**, 2930–2931.
  - Mallon, A.M., Blake, A. and Hancock, J.M. (2008) EuroPhenome and EMPreSS: online mouse phenotyping resource. *Nucleic Acids Res.*, **36**, D715–D718.
  - Taylor, C.F., Field, D., Sansone, S.A., Aerts, J., Apweiler, R., Ashburner, M., Ball, C.A., Binz, P.A., Bogue, M., Booth, T. *et al.* (2008) Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. *Nat. Biotechnol.*, **26**, 889–896.
  - Smith, C.L., Goldsmith, C.A. and Eppig, J.T. (2005) The Mammalian Phenotype Ontology as a tool for annotating, analyzing and comparing phenotypic information. *Genome Biol.*, **6**, R7.
  - Gkoutos, G.V., Green, E.C., Mallon, A.M., Hancock, J.M. and Davidson, D. (2005) Using ontologies to describe mouse phenotypes. *Genome Biol.*, **6**, R8.
  - Beck, T., Morgan, H., Blake, A., Wells, S., Hancock, J.M. and Mallon, A.M. (2009) Practical application of ontologies to annotate and analyse large scale raw mouse phenotype data. *BMC Bioinformatics*, **10**(Suppl. 5), S2.
  - Brown, S.D., Hancock, J.M. and Gates, H. (2006) Understanding mammalian genetic systems: the challenge of phenotyping in the mouse. *PLoS Genet.*, **2**, e118.
  - Brown, S.D., Wurst, W., Kuhn, R. and Hancock, J. (2009) The functional annotation of mammalian genomes: the challenge of phenotyping. *Annu. Rev. Genet.*, **43**, 305–333.
  - Schofield, P.N., Bubela, T., Weaver, T., Portilla, L., Brown, S.D., Hancock, J.M., Einhorn, D., Tocchini-Valentini, G., Hrabe de Angelis, M., Rosenthal, N. *et al.* (2009) Post-publication sharing of data and tools. *Nature*, **461**, 171–173.
- APPENDIX 1: THE EUMODIC CONSORTIUM**
- Autonomous University of Barcelona (UAB), Spain**  
Fatima Bosch<sup>1</sup>, Jesús Ruberte<sup>1</sup>, Tura Ferre<sup>1</sup>, Anna Pujol<sup>1</sup>, Pedro Otaegui<sup>1</sup>, Sylvie Franckhauser<sup>1</sup>
- Ani.Rhône-Alpes Lyon, France**  
Jacqueline Marvel<sup>2</sup>, Veronique Queste<sup>2</sup>, Romain Dacquin<sup>2</sup>, Sophia Djebali<sup>2</sup>, Pierre Jurdic<sup>2</sup>
- Alexander Fleming Biomedical Sciences Research Center, Athens, Greece**  
George Kollias<sup>3</sup>, Christina Chandras<sup>3</sup>, Eleni Douni<sup>3</sup>, Vassilis Aidinis<sup>3</sup>
- CNIO, Madrid, Spain**
- Mariano Barbacid<sup>4</sup>, Marta Cañamero<sup>4</sup>, Pierre Dubus<sup>5</sup>
- Helmholtz Centre for Infection Research (HZI), Braunschweig, Germany**  
Rudi Balling<sup>9</sup>, Hansjörg Hauser<sup>9</sup>, Bastian Pasche<sup>9</sup>, Werner Müller<sup>10</sup>, Andreas Lengeling<sup>11</sup>
- CNR, IBC Monterotondo, Italy**  
Glauco Tocchini-Valentini<sup>6</sup>, Silvia Mandillo<sup>6</sup>, Elisabetta Golini<sup>6</sup>, Daniela Marazziti<sup>6</sup>, Giancarlo Deidda<sup>6</sup>, Nicoletta Rossi<sup>6</sup>, Rafaele Matteoni<sup>6</sup>, Marcello Raspa<sup>6</sup>, Ferdinando Scavizzi<sup>6</sup>, Richard Hugh Butler<sup>6</sup>, Gianfranco Di Segni<sup>6</sup>, Paolo Fruscoloni<sup>6</sup>, Patrizia Calandra<sup>6</sup>, Cecilia Mannironi<sup>6</sup>, Daniela Scarabino<sup>6</sup>, Giuseppe D. Tocchini-Valentini<sup>6</sup>, Michela Zamboni<sup>6</sup>, Sabrina Putti<sup>6</sup>, Chiara Di Pietro<sup>6</sup>, Serena Gastaldi<sup>6</sup>
- CNRS, Institut de Transgenose, Orleans, France**  
Yann Herault<sup>7</sup>, Bernard Ryffel<sup>7</sup>, Marc le Bert<sup>7</sup>, Patricia Lopes Pereira<sup>7</sup>, Véronique Brault<sup>7</sup>, Emilie Dalloneau<sup>7</sup>, Stéphanie Pothion<sup>7</sup>
- EMBL Mouse Biology Monterotondo, Italy**  
Nadia Rosenthal<sup>8</sup>, Mumna Al Banchaabouchi<sup>8</sup>, Raffaele Migliozzi<sup>8</sup>, Ekaterina Salimova<sup>8</sup>
- Helmholtz Zentrum München, German Research Center for Environmental Health, Munich, Germany**  
Martin Hrabé de Angelis<sup>12,a</sup>, Valérie Gailus-Durner<sup>12</sup>, Helmut Fuchs<sup>12</sup>, Thure Adler<sup>12,14</sup>, Antonio Aguilar-Pimentel<sup>12,b</sup>, Lore Becker<sup>12,c</sup>, Raffi Bekerredjian<sup>12,d</sup>, Dirk H. Busch<sup>14</sup>, Julia Calzada-Wack<sup>e</sup>, Patricia Da Silva-Buttkus<sup>e</sup>, Irene Esposito<sup>e</sup>, Jack Favor<sup>f</sup>, Wolfgang Fleischmann<sup>12</sup>, Lilian Garrett<sup>13</sup>, Lisa Glasl<sup>13</sup>, Alexander Götz<sup>g</sup>, Jochen Graw<sup>13</sup>, Wolfgang Hans<sup>12</sup>, Sabine M. Hölter-Koch<sup>13</sup>, Boris Ivandic<sup>12,d</sup>, Hugo A. Katus<sup>12,d</sup>, Martin Klingenspor<sup>12,h</sup>, Thomas Klopstock<sup>12,c</sup>, Christoph Lengger<sup>12</sup>, Tonia Ludwig<sup>12</sup>, Holger Maier<sup>12</sup>, Kateryna Micklich<sup>12</sup>, Sun Minxuan<sup>13</sup>, Beatrix Naton<sup>12</sup>, Frauke Neff<sup>e</sup>, Markus Ollert<sup>12,b</sup>, Oliver Puk<sup>13</sup>, Leticia Quintanilla-Fend<sup>e</sup>, Ildiko Racz<sup>12,i</sup>, Birgit Rathkolb<sup>12,j</sup>, Jan Rozman<sup>12,h</sup>, Karl-Heinz Schäble<sup>12</sup>, Evelyn Schiller<sup>12</sup>, Anja Schrewe<sup>12</sup>, Ralf Steinkamp<sup>12</sup>, Claudia Stöger<sup>12</sup>, Tobias Stöger<sup>g</sup>, Holger Schulz<sup>g</sup>, Monica Tost<sup>e</sup>, Irina Treise<sup>12</sup>, Daniela Vogt-Weisenhorn<sup>13</sup>, Monja Willershäuser<sup>12</sup>, Eckhard Wolf<sup>12</sup>, Annemarie Wolff-Muscate<sup>13</sup>, Wolfgang Wurst<sup>13,k</sup>, Ramona Zeh<sup>12</sup>, Andreas Zimmer<sup>12,i</sup>
- Institut Clinique de la Souris, Strasbourg, France**  
Jean-Louis Mandel<sup>15</sup>, Yann Héroult<sup>15</sup>, Marie-France Champy<sup>15</sup>, Mohammed Selloum<sup>15</sup>, Hamid Meziane<sup>15</sup>, Laurent Monassier<sup>15</sup>, Roy Combe<sup>15</sup>, Michel Roux<sup>15</sup>, Abdel Ayadi<sup>15</sup>, Stéphanie Muller<sup>15</sup>, Elodie Bedu<sup>15</sup>, Julien Becker<sup>15</sup>, Sophie Leblanc<sup>15</sup>, Tania Sorg<sup>15</sup>, Laurent Vasseur<sup>15</sup>
- MRC Human Genetics Unit, Edinburgh, UK**  
Ian Jackson<sup>18</sup>, Sally Cross<sup>18</sup>, Russell Joynson<sup>18</sup>, Shalini Jadeja<sup>18</sup>
- Tel Aviv University, Israel**

Karen Avraham<sup>19</sup>, Amiel Dror<sup>19</sup>, Shaked Shivatzki<sup>19</sup>  
**Telethon Institute of Genetics and Medicine (TIGEM),  
Italy**

Andrea Ballabio<sup>20</sup>, Graciana Diez-Roux<sup>20</sup>

**MRC Harwell (MGU and MLC), UK**

Steve Brown<sup>16</sup>, Hilary Gates<sup>16</sup>, Niels Adams<sup>17</sup>, Sarah  
Atkins<sup>16</sup>, Tim Beck<sup>16</sup>, Andy Blake<sup>16</sup>, Kan Pai Chiev<sup>16</sup>,  
Andre Chouankam<sup>16</sup>, Roger Cox<sup>16</sup>, Paul Denny<sup>16</sup>, Irina  
Emelyanova<sup>16</sup>, Martin Fray<sup>17</sup>, Simon Greenaway<sup>16</sup>, John  
Hancock<sup>16</sup>, Tertius Hough<sup>17</sup>, Kelly Hunt<sup>16</sup>, Rachel  
Kendall<sup>16</sup>, Kirsty Lee<sup>16</sup>, Ann-Marie Mallon<sup>16</sup>, Hugh  
Morgan<sup>16</sup>, Pat Nolan<sup>16</sup>, Paul Potter<sup>17</sup>, Michelle Stewart<sup>17</sup>,  
Lydia Teboul<sup>17</sup>, Simon Vowell<sup>16</sup>, Sara Wells<sup>17</sup>, Joe  
Wood<sup>17</sup>, Rumana Zaman<sup>16</sup>

**University of Cambridge, UK**

Paul Schofield<sup>21</sup>, Michael Gruenberger<sup>21</sup>, Jules Griffin<sup>22</sup>

**University of Lausanne, Switzerland**

Walter Wahli<sup>23</sup>, Frederic Preitner<sup>23</sup>, Maria Belen  
Delgado<sup>23</sup>, Béatrice Desvergne<sup>23</sup>, Liliane Michalik<sup>23</sup>,  
Elodie Bedu<sup>23</sup>

**University of Manchester, UK**

Ludwig Neyses<sup>24</sup>, Elizabeth Cartwright<sup>25</sup>, Sukhpal  
Prehar<sup>25</sup>, Min Zi<sup>25</sup>

**The Wellcome Trust Sanger Institute, Cambridge, UK**

Ramiro Ramirez-Solis<sup>26</sup>, Karen Steel<sup>26</sup>, Jacqui White<sup>26</sup>,  
Dave Sunter<sup>26</sup>, David Melvin<sup>26</sup>, Anna Karin Gerdin<sup>26</sup>,  
Natalia Igosheva<sup>26</sup>, Jeanne Estabel<sup>26</sup>, Sophie Messenger<sup>26</sup>,  
Natasha Karp<sup>26</sup>, Carl Shannon<sup>26</sup>, Laura Roberson<sup>26</sup>,  
David Richardson<sup>26</sup>, Christine Podrini<sup>26</sup>, Mark Lucas<sup>26</sup>

<sup>1</sup>Centre of Animal Biotechnology and Gene Therapy,  
School of Veterinary Medicine, Autonomous University  
of Barcelona, 08193 Bellaterra, Spain

<sup>2</sup>Ani.Rhone-Alpes, IFR128, 21 Avenue Tony Garnier,  
Lyon Cedex 07, France

<sup>3</sup>Alexander Fleming Biomedical Sciences Research  
Center, Athens, Greece

<sup>4</sup>Centro Nacional de Investigaciones Oncologicas  
(CNIO), Molecular Oncology & Comparative Pathology,  
Melchor Fernandez Almagro 3, 28029 Madrid, Spain

<sup>5</sup>Histology and Molecular Pathology Department,  
University of Bordeaux, EA2406, 146 Rue Leo Saignat,  
33076 Bordeaux, France

<sup>6</sup>IBC-CNR, Campus 'A.Buzzati-Traverso', via Ramarini  
32 00016 Monterotondo Scalo, Rome, Italy

<sup>7</sup>CNRS Institut de Transgenose, IEM, 3B rue de la  
ferrollerie 45071 Orleans cedex 2, France

<sup>8</sup>EMBL Monterotondo, Mouse Phenotyping Facility,  
'A.Buzzati-Traverso Campus', via Ramarini 32, 00015  
Monterotondo Scalo, Rome, Italy

<sup>9</sup>Helmholtz Centre for Infection Research (HZI),  
Mascheroder Weg1, 38124, Braunschweig, Germany

<sup>10</sup>The University of Manchester, AV Hill Building, Oxford  
Road, Manchester, M13 9PT, UK

<sup>11</sup>The Roslin Institute, University of Edinburgh, Easter  
Bush Veterinary Research Centre, Roslin, Midlothian,  
EH25 9RG, UK

<sup>12</sup>Helmholtz Zentrum München, German Research  
Center for Environmental Health, Institute of  
Experimental Genetics and German Mouse Clinic,  
Ingolstaedter Landstrasse 1, D-85764 Neuherberg,  
Germany

<sup>13</sup>Helmholtz Zentrum München, German Research  
Center for Environmental Health, Institute of  
Developmental Genetics, Ingolstaedter Landstrasse 1, D-  
85764 Neuherberg, Germany

<sup>14</sup>Institute for Medical Microbiology, Immunology and  
Hygiene, Technische Universität München, Trogerstrasse  
9, 81675 Munich, Germany

<sup>15</sup>Institut Clinique de la Souris (ICS), 1 Rue Laurent Fries,  
BP10142, 67404 Illkirch, Cedex, France

<sup>16</sup>Medical Research Council Mammalian Genetics Unit,  
Harwell, Oxfordshire, OX11 0RD, UK

<sup>17</sup>Medical Research Council Mary Lyon Centre, Harwell,  
Oxfordshire, OX11 0RD, UK

<sup>18</sup>Medical Research Council Human Genetics Unit,  
Comparative and Developmental Genetics Department,  
Western General Hospital, Crewe Road, Edinburgh, EH4  
2XU, UK

<sup>19</sup>Department of Human Molecular Genetics and  
Biochemistry, Sackler Faculty of Medicine, Tel Aviv  
University, Tel Aviv, Israel

<sup>20</sup>Telethon Institute of Genetics and Medicine, Via Pietro  
Castellino 111, 80131 Napoli, Italy

<sup>21</sup>Department of Physiology, Development and  
Neuroscience, University of Cambridge, Downing Street,  
Cambridge, CB2 3DY, UK

<sup>22</sup>Department of Clinical Biochemistry, University of  
Cambridge, Addenbrooke's Hospital, Box 232, Hills  
Road, Cambridge CB2 2QR, UK

<sup>23</sup>Center for Integrative Genomics, National Centre of  
Research 'Frontiers in Genetics', University of Lausanne,  
BEP, CH-1015 Lausanne, Switzerland

<sup>24</sup>Manchester Heart Centre, Manchester Royal Infirmary,  
University of Manchester, Oxford Road, Manchester M13  
9WL, UK

<sup>25</sup>Medical School, Rm 1.302, Stopford Building,  
University of Manchester, Oxford Road, Manchester,  
M13 9PT, UK

<sup>26</sup>The Wellcome Trust Sanger Institute, Wellcome Trust  
Genome Campus, Hinxton, Cambridge, CB10 1SA, UK

<sup>3</sup>Chair of Experimental Genetics, Center of Life and Food  
Sciences Weihenstephan, Technische Universität  
München, Ingolstädter Landstr. 1, 85764 Neuherberg,  
Germany

<sup>b</sup>Division of Environmental Dermatology and Allergy  
(UDA), Helmholtz Zentrum München/Technische  
Universität München, and Clinical Research Division of



Molecular and Clinical Allergotoxicology, Department of Dermatology and Allergy, Technische Universität München, Munich, Germany

<sup>c</sup>Neurological Clinic, Friedrich-Baur-Institute, Ludwig-Maximilians-Universität, Ziemssenstr. 1a, 80336 Munich, Germany

<sup>d</sup>Heidelberg University Hospital, Im Neuenheimer Feld 410, 69120 Heidelberg, Germany

<sup>e</sup>Institute of Pathology, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

<sup>f</sup>Institute of Human Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

<sup>g</sup>Institute of Lung Biology and Disease, Helmholtz Zentrum München, German Research Center for

Environmental Health GmbH, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany

<sup>h</sup>Technische Universität München, Molecular Nutritional Medicine, Else Kröner-Fresenius Center, Am Forum 8, 85354 Freising-Weihenstephan, Germany

<sup>i</sup>Institute of Molecular Psychiatry, Life & Brain Center, University of Bonn, Sigmund Freud Str. 25, 53127 Bonn, Germany

<sup>j</sup>Gene Center, Institute of Molecular Animal Breeding and Biotechnology, Ludwigs-Maximilians-Universität München, Feodor-Lynen Str. 25, 81377 Munich, Germany

<sup>k</sup>Chair of Developmental Genetics, Center of Life and Food Sciences Weihenstephan, Technische Universität München, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany