Mouse Phenotyping Informatics Infrastructure (MPI2)

Vivek Iyer, Hugh Morgan, Henrik Westerberg, Terry Meehan and Helen Parkinson
Centers all around the world collaborating in IMPC

- Each center producing data from 25 different phenotyping procedures
- Organizing, comparing and integrating
MPI2 consortium

Data Flow

Tracking
SOPS
Data Validation
Analyze
Integrate
Disseminate

KOMP2
other
IMPC partners

JAX
BaSH
DTCC

IMPC
Achieving IMPC informatics goals

- Coordinate
- Standardize Protocols & Data Wrangling
- Validate
- Analyze & Archive
- Integrate & Disseminate
**DATA TRACKING**

Basic tracking system available to production centers

Advanced tracking system

Tracking portal

**SOPDB**

Review SOPS proposed by centers and complete version 1

Version 1 of SOPDB

Version 2 of SOPDB based on new guidelines

Extend SOPDB web portal with use case defined functions

Refine SOP definitions and manage versioning

**Pheno-DCC**

Review data export and upload process

Set up LIMS export with centers

Design and implement Pheno-DCC database schema

Develop export module

Validation and QC tool development

Image upload tools

Image annotation specification and tool development

Data management system complete and exported to portals

**Statistical Analysis and Data Annotation**

Review experimental design of SOPs

Modular Annotation Pipeline V1

Incorporation of EBI R analysis infrastructure

Incorporation of added value data sets

Modular Annotation Pipeline V2

**CDA and IT infrastructure**

Data warehouse

**KOMP2 Web Portal**

High priority portal use cases (D4.2.1)

Medium priority portal use cases (D4.2.2)

Low priority portal use cases (D4.2.3)

Further development incorporating new use cases

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DATA TRACKING</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Basic tracking system available to production centers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced tracking system</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tracking portal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOPDB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review SOPS proposed by centers and complete version 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Version 1 of SOPDB</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Version 2 of SOPDB based on new guidelines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Extend SOPDB web portal with use case defined functions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Refine SOP definitions and manage versioning</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pheno-DCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review data export and upload process</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Set up LIMS export with centers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Design and implement Pheno-DCC database schema</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Develop export module</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Validation and QC tool development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Image upload tools</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Image annotation specification and tool development</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data management system complete and exported to portals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical Analysis and Data Annotation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Review experimental design of SOPs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modular Annotation Pipeline V1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incorporation of EBI R analysis infrastructure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Incorporation of added value data sets</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modular Annotation Pipeline V2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CDA and IT infrastructure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data warehouse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KOMP2 Web Portal</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High priority portal use cases (D4.2.1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medium priority portal use cases (D4.2.2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low priority portal use cases (D4.2.3)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Further development incorporating new use cases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Detailed Overview

- Coordinate
  - Tracking: Vivek Iyer

- Cross center integration
  - Phenotype procedures and data wrangling: Henrik Westerberg

- Validate
  - Data upload and QC: Hugh Morgan

- Analyze & Archive
  - Central data archive: Terry Meehan

- Integrate & Disseminate
  - Plans for the future: Helen Parkinson
iMITS Core function

IMPC Production & Phenotyping Centers

PLANS

STATUS

iMITS

KOMP2

other IMPC partners

Report ...

IMPC
**iMITS Users**

Production centers *input* data

*Other* production centers *compare* data

NIH *Monitors* progress

IMPC Portal *displays* progress to *community*

---

Production & Phenotyping centers

iMITS

NIH

IMPC Portal

Production & Phenotyping centers
iMITS Value

1. **CONSISTENT STATUS INPUT**
   - ALL centers PROVIDE SAME FORMAT

2. **AVOID** accidental **Duplication Of Effort**

3. **CONSISTENT REPORT OUTPUT** (comparison)

Production & Phenotyping centers

NIH

IMPC Portal
1. PLANS and pre-production ES QC Statuses

2. Production Status, Strains Genotyping Assays

3. Phenotyping Statuses, Strains

Rederivation Of Mutant Mouse

Cre-Excision (Generation of null)

Phenotype Data to DCC
iMITS Planning reports

GOAL:
To avoid unintentional duplication of effort

SUCCESS this year
Consortia could avoid duplication before production

Small number of visible duplications

CHALLENGES
Known duplication to be flagged as intentional

Actively pushing duplication to producers
iMITS Avoiding duplication

Possible Double Production Reported

Discussion ...

GENE X

BaSH Interest

JAX Interest

GENE X

BaSH Conflict

JAX Conflict

GENE X

BaSH Withdraw

JAX Assigned

JAX ES QC starts
iMITS Reporting duplication

DOUBLE-PLANNING REPORTS via iMITS REPORTS page

“MATRIX” collisions reported with details of each plan

EACH collision must be inspected
iMITS Reporting duplication

**iMITS shows this report**

<table>
<thead>
<tr>
<th>Consortium</th>
<th>Double-plans or production inside KOMP2</th>
</tr>
</thead>
<tbody>
<tr>
<td>BaSH</td>
<td>5</td>
</tr>
<tr>
<td>DTCC</td>
<td>5</td>
</tr>
<tr>
<td>JAX</td>
<td>0</td>
</tr>
</tbody>
</table>

**DETAILS can be inspected ...**

<table>
<thead>
<tr>
<th>Celf6</th>
<th>DTCC</th>
<th>Assigned - ES Cell QC Complete</th>
<th>Chimeras obtained</th>
<th>UCD</th>
<th>2011-08-31</th>
</tr>
</thead>
<tbody>
<tr>
<td>Celf6</td>
<td>BaSH</td>
<td>Assigned</td>
<td>Genotype confirmed</td>
<td>WTSI</td>
<td>2011-10-31</td>
</tr>
</tbody>
</table>

**CHALLENGE: We need to actively push to producers**
SUCCESS this year
Consistent, Timely Production Reporting
Standard formats for Totals and Monthly Activity
Computer updates from all consortia

CHALLENGES: Linking to ESCell Pre-production QC, Speed
Current production reports BaSH

IMPC Graph Report Display

<table>
<thead>
<tr>
<th>State</th>
<th>Current</th>
<th>Last Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assigned genes</td>
<td>1157</td>
<td>185</td>
</tr>
<tr>
<td>ES QC</td>
<td>502</td>
<td>35</td>
</tr>
<tr>
<td>ES QC Confirmed</td>
<td>301</td>
<td>75</td>
</tr>
<tr>
<td>ES QC Failed</td>
<td>43</td>
<td>4</td>
</tr>
<tr>
<td>Mouse Production</td>
<td>267</td>
<td>53</td>
</tr>
<tr>
<td>Confirmed Mice</td>
<td>77</td>
<td>8</td>
</tr>
<tr>
<td>Indent to Phenotype</td>
<td>93</td>
<td>14</td>
</tr>
<tr>
<td>Cre Excision Complete</td>
<td>27</td>
<td>3</td>
</tr>
<tr>
<td>Phenotyping Complete</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

BaSH MI Performance

BaSH GC Performance

Individual reports for Bash
Total production
Last month’s activity
Monthly progress graphs
Current production reports DTCC

<table>
<thead>
<tr>
<th>State</th>
<th>Current</th>
<th>Last Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assigned genes</td>
<td>352</td>
<td>0</td>
</tr>
<tr>
<td>ES QC</td>
<td>72</td>
<td>0</td>
</tr>
<tr>
<td>ES QC Confirmed</td>
<td>345</td>
<td>5</td>
</tr>
<tr>
<td>ES QC Failed</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mouse Production</td>
<td>242</td>
<td>26</td>
</tr>
<tr>
<td>Confirmed Mice</td>
<td>83</td>
<td>8</td>
</tr>
<tr>
<td>Indent to Phenotype</td>
<td>138</td>
<td>0</td>
</tr>
<tr>
<td>Cre Excision Complete</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Phenotyping Complete</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

DTCC MI Performance

DTCC GC Performance

Individual reports for DTCC
Total production
Last month’s activity
Monthly progress graphs

IMPC
Current production reports Jax

<table>
<thead>
<tr>
<th>State</th>
<th>Current</th>
<th>Last Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assigned genes</td>
<td>555</td>
<td>6</td>
</tr>
<tr>
<td>ES QC</td>
<td>124</td>
<td>7</td>
</tr>
<tr>
<td>ES QC Confirmed</td>
<td>127</td>
<td>53</td>
</tr>
<tr>
<td>ES QC Failed</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Mouse Production</td>
<td>96</td>
<td>14</td>
</tr>
<tr>
<td>Confirmed Mice</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Indent to Phenotype</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Cre Excision Complete</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Phenotyping Complete</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

**JAX MI Performance**
- mi Cumulative
- mi Goals
- Below
- Above

**JAX GC Performance**
- gc Cumulative
- gc Goals
- Below
- Above
iMITS Exploring new reports

SUCCESS this year

We have explored different reporting strategies, converging on what works and is informative

CHALLENGES:

Closer integration / feedback from senior users

<table>
<thead>
<tr>
<th>Month</th>
<th>Started</th>
<th>Chim</th>
<th>GLT</th>
<th>Cre S</th>
<th>Cre X</th>
<th>Ph complete</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dec 11</td>
<td>4</td>
<td>4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Nov 11</td>
<td>18</td>
<td>16</td>
<td>9</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Oct 11</td>
<td>32</td>
<td>30</td>
<td>16</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sep 11</td>
<td>42</td>
<td>40</td>
<td>29</td>
<td>15</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Example for BASH:
How many mouse production attempts were started in Dec 11?
How many of those have now produced GLT Mice?
How many of those have now had Cre Excision completed?
iMITS Mutant Strains from MGI

1. Status
   Genotype QC
   WT Strains

2. Genotypes

3. Allele strain names

Production centers report sufficient data to iMITS.

MGI reads that, and turns around names.

IMPC Portal: Consistent Mutant Strain Names for all IMPC Mutants.
Detailed Overview

• Coordinate
  - Tracking: Vivek Iyer

• Cross Center Integration
  - Phenotype procedures and data wrangling: Henrik Westerberg

• Validate
  - Data upload and QC: Hugh Morgan

• Analyze & Archive
  - Central data archive: Terry Meehan

• Integrate & Disseminate
  - Plans for the future: Helen Parkinson
• Centers all around the world collaborating in IMPC
• Each center producing data from 25 different phenotyping procedures
• Organizing, comparing and integrating
Phenotype Procedure Defined

http://www.mousephenotype.org/impress

• IMPC Phenotype Procedure Definition:
  
  • Generic procedure for collecting data for a phenotypic test, consisting of agreed data and metadata parameters.

  • Also contains ontological associations from Mammalian Phenotype (MP), e.g. MP:0000188: abnormal circulating glucose level.
## Phenotype Procedure Structure (Simplified)

<table>
<thead>
<tr>
<th>Parameter Name</th>
<th>Parameter Type</th>
<th>Unit</th>
<th>Data Type</th>
<th>MP Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole arena average speed</td>
<td>Measured</td>
<td>cm/s</td>
<td>Float</td>
<td>MP:0003313 abnormal locomotor activation</td>
</tr>
<tr>
<td>LacZ Images</td>
<td>Measured</td>
<td>-</td>
<td>Image</td>
<td>MP:0000689 abnormal spleen morphology</td>
</tr>
<tr>
<td>Equipment manufacturer</td>
<td>MetaData</td>
<td>-</td>
<td>Text</td>
<td>-</td>
</tr>
</tbody>
</table>
Phenotype Procedure Signoff Process
Phenotype Procedure Signoff Process

1. A Center Produce Pilot Data
2. Phenotypers
3. Wranglers

Steps:
- Step 1: Communication
- Step 2: Pilot Data
- Step 3: Phenotypers

IMPC Centers

Researchers

Matrix

IMPRESS
IMPC Pipeline

Embryo

- Embryo LacZ
- Viability E9.5
- Viability E12.5
- Viability E15.5
- E9.5
- E12.5
- E15.5

Fertility
- Cohort Breeding

7M + 7F Mutant Adult Mice

Weight Curve - 4wk to 16wk

In life

- Open Field
- Modified SHIRPA & Dysmorphology
- Grip Strength
- Acoustic Startle/PPI
- Calorimetry
- ECG/Echo
- Challenge Whole Body Plethysmography
- Intraperitoneal Glucose Tolerance Test
- X-ray (5+5)
- Auditory Brain Stem Response
- Body Composition (lean/fat)
- Pain Test
- OCT
- Eye Morphology

Terminal

- Haematology
- Adult LacZ
- Clinical Blood Chemistry
- Insulin Blood Level
- FACS (spleen)
- Heart Weight
- Gross Pathology and Tissue Collection (2+2)
- Tissue embedding & block banking
- Histopathology (2+2)

KEY:
- Mandatory tests
- Non-Mandatory tests
- Tests in development or under consideration

IMPC
Currently, 25 protocols in total:

- 15 protocols approved
- 10 protocols in development

What is a protocol in development:

- Just started, generating pilot data
- Has parameters which need to be agreed upon by the centers
- Require sign off by call chair
- Requires draft protocol text
- Assignment of MP terms
Key Field: Ontology Associations

Ontology Associations

- Structured, controlled vocabulary used worldwide by scientists to describe phenotypes and beyond
- Mammalian Phenotype (MP) describes 40000 genotypes, 8000 genes
- Predefined option choices for high throughput ontology annotation

Breaking down MP terms to Entity Quantity Relationships is essential to compare annotations with other ontologies

- Implemented using PATO (NHGRI funded, Suzi Lewis)
- Enables comparison with human centric ontologies
Data Wrangler’s Role

- Center Phenotypers
- Data Managers
- IMPC Pipeline
- LIMS
- Export
- Pheno DCC
- Phenomap
- Phenotype Procedures
- Troubleshooting (IT, Data Export)
- Data Quality Control Statistical Analyses
- Future Work
Data Wrangling Example

Data QC: Developing tools and testing with legacy data.
Data Wrangling Example

Image Wrangling: Collecting examples from other centers, working with OPT and uCT.
Detailed Overview

- Coordinate
  - Tracking: Vivek Iyer

- Cross center integration
  - Phenotyping procedures and data wrangling: Henrik Westerberg

- Validate
  - Data upload and QC: Hugh Morgan

- Analyze & Archive
  - Central data archive: Terry Meehan

- Integrate & Disseminate
  - Plans for the future: Helen Parkinson
DCC Function

- Coordination with the centers about phenotyping pipeline
- Facilitate data exchange from the centers to the MPI2
- Ensure data integrity and accuracy
- Contribute to annotation pipeline design and implementation
- Present all data (pre and post QC) to project partners and the public
Data Flow through DCC

1. LIMS
2. Internet
3. PhenoDCC
4. Internet
5. CDA
6. Internet
7. Post-QC data
8. Researcher
9. Phenotype (pre or post-QC)
10. Data exporter / QC expert
Data Flow through DCC
Data Flow through DCC

- Diverse data stored in varied systems at each center, provides data integration challenge

- Common software solution provided to all phenotyping centers for data representation, validation and transfer

- Status: First stable version released
  - Successful export from Test Center
Data Flow through DCC

PhenoDCC

Internet

Phenotype (pre or post-QC)

Researcher

Internet

CDA

Post-QC data

Researcher

LIMS

Internet

Data exporter / QC expert
Data Import to DCC

- Immediate release data as cohorts progress through pipeline to ensure rapid data release
- Automated capture of data from centers ensures rapid data presentation
- Status: Sandbox released for test purposes
Data Flow through DCC

Phenotype
(pre or post-QC)

Researcher

PhenoDCC

Internet

LIMS

Data exporter / QC expert

Internet

CDA

Post-QC data

Researcher
Data Validation and QC

- Data from wide range of equipment and involving varying levels of human interpretation from 25 procedures

- Vital to ensure maximal level of data consistency and accuracy

- QC interface will allow visible validation across centers and DCC, manages all communication

- Status: Basic automated validation released. QC interface to be released in coming months
Select a background strain to get additional details.
Select a background strain to get additional details.
### Specimen Information

<table>
<thead>
<tr>
<th>Specimen ID</th>
<th>Issue</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specimen_1</td>
<td><code>cvc-complex-type.2.4.a: Invalid content was found starting with element 'sa...</code></td>
<td>failed</td>
</tr>
<tr>
<td>Specimen_2</td>
<td><code>cvc-complex-type.2.4.a: Invalid content was found starting with element 'sa...</code></td>
<td>failed</td>
</tr>
<tr>
<td>Specimen_3</td>
<td><code>cvc-complex-type.2.4.a: Invalid content was found starting with element 'sa...</code></td>
<td>failed</td>
</tr>
<tr>
<td>Specimen_4</td>
<td><code>cvc-complex-type.2.4.a: Invalid content was found starting with element 'sa...</code></td>
<td>failed</td>
</tr>
</tbody>
</table>

### Experiment Information

<table>
<thead>
<tr>
<th>File Name</th>
<th>Specimen ID</th>
<th>Experiment</th>
<th>Issue</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>CENT.2012-03-27.01.experiment.e...</td>
<td>Specimen_1</td>
<td>Exp_1</td>
<td>Invalid content was found starting with element 'sampleID'</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.02.experiment.e...</td>
<td>Specimen_2</td>
<td>Exp_2</td>
<td>Procedure Open Field Missing required parameter Arena Size</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.03.experiment.e...</td>
<td>Specimen_3</td>
<td>Exp_3</td>
<td>Value 0.0 below range for Center average speed</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.04.experiment.e...</td>
<td>Specimen_4</td>
<td>Exp_4</td>
<td>Attribute 'fileType' must appear on element 'mediaParameter'.</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.04.experiment.e...</td>
<td>Specimen_5</td>
<td>Exp_5</td>
<td>Attribute 'name' is not allowed to appear in element 'dim'.</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.05.experiment.e...</td>
<td>Specimen_6</td>
<td>Exp_6</td>
<td>Specimen Specimen_6 not found for 1 procedure Exp_6</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.06.experiment.e...</td>
<td>Specimen_7</td>
<td>Exp_7</td>
<td>The element type</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.06.experiment.e...</td>
<td>Specimen_8</td>
<td>Exp_8</td>
<td>Failed to download Image: Dock10m-30184555-b-ho-110404.jpg</td>
<td>failed</td>
</tr>
</tbody>
</table>

---

**Id:** MGI:1333875  
**Symbol:** Bre  
**Name:** brain and reproductive organ-expressed protein  
**Allele:** tm1a(EUCOMM)Wtsi  
**Strain:** C57BL/6Dnk  
**Genotype:** EPD00654_D03
### Experiment Information

<table>
<thead>
<tr>
<th>File Name</th>
<th>Specimen ID</th>
<th>Experiment</th>
<th>Issue</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>CENT.2012-03-27.01.experiment.euph.xml</td>
<td>Specimen_1</td>
<td>Exp_1</td>
<td>Invalid content was found starting with element 'sampleID'</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.02.experiment.euph.xml</td>
<td>Specimen_2</td>
<td>Exp_2</td>
<td><strong>Procedure Open Field Missing required parameter Arena Size</strong></td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.03.experiment.euph.xml</td>
<td>Specimen_3</td>
<td>Exp_3</td>
<td>Value 0.0 below range for Center average speed</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.04.experiment.euph.xml</td>
<td>Specimen_4</td>
<td>Exp_4</td>
<td>Attribute 'fileType' must appear on element 'mediaParameter'.</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.04.experiment.euph.xml</td>
<td>Specimen_5</td>
<td>Exp_5</td>
<td>Attribute 'name' is not allowed to appear in element 'dim'.</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.05.experiment.euph.xml</td>
<td>Specimen_6</td>
<td>Exp_6</td>
<td>Specimen Specimen_6 not found for 1 procedure Exp_6</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.06.experiment.euph.xml</td>
<td>Specimen_7</td>
<td>Exp_7</td>
<td>The element type</td>
<td>failed</td>
</tr>
<tr>
<td>CENT.2012-03-27.06.experiment.euph.xml</td>
<td>Specimen_8</td>
<td>Exp_8</td>
<td>Failed to download Image: Dock10m-30184555-b-ho-110404....</td>
<td>failed</td>
</tr>
</tbody>
</table>
### Specimen and experiment details

<table>
<thead>
<tr>
<th>Name</th>
<th>DOB</th>
<th>Sex</th>
<th>Zygorosity</th>
<th>Cohort</th>
<th>Experimenter</th>
<th>Start Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRE-B6N/15.3e_439...</td>
<td>2009/11/17</td>
<td>Male</td>
<td>Homzygous</td>
<td>BRE-B6N-1</td>
<td></td>
<td>2010/02/19</td>
</tr>
<tr>
<td>BRE-B6N/15.3c_439...</td>
<td>2009/11/17</td>
<td>Male</td>
<td>Homzygous</td>
<td>BRE-B6N-1</td>
<td></td>
<td>2010/02/19</td>
</tr>
<tr>
<td>BRE-B6N/15.3f_439...</td>
<td>2009/11/17</td>
<td>Male</td>
<td>Homzygous</td>
<td>BRE-B6N-1</td>
<td></td>
<td>2010/02/19</td>
</tr>
<tr>
<td>BRE-B6N/15.3d_439...</td>
<td>2009/11/17</td>
<td>Male</td>
<td>Homzygous</td>
<td>BRE-B6N-1</td>
<td></td>
<td>2010/02/19</td>
</tr>
<tr>
<td>BRE-B6N/13.3f_437...</td>
<td>2009/10/22</td>
<td>Male</td>
<td>Homzygous</td>
<td>BRE-B6N-1</td>
<td></td>
<td>2010/01/21</td>
</tr>
<tr>
<td>BRE-B6N/13.3d_437...</td>
<td>2009/10/22</td>
<td>Male</td>
<td>Homzygous</td>
<td>BRE-B6N-1</td>
<td></td>
<td>2010/01/21</td>
</tr>
<tr>
<td>BRE-B6N/13.3g_437...</td>
<td>2009/10/22</td>
<td>Male</td>
<td>Homzygous</td>
<td>BRE-B6N-1</td>
<td></td>
<td>2010/01/21</td>
</tr>
<tr>
<td>BRE-B6N/13.3e_437...</td>
<td>2009/10/22</td>
<td>Male</td>
<td>Homzygous</td>
<td>BRE-B6N-1</td>
<td></td>
<td>2010/01/21</td>
</tr>
<tr>
<td>BRE-B6N/14.1c_435...</td>
<td>2010/05/05</td>
<td>Female</td>
<td>Homzygous</td>
<td>BRE-B6N-1</td>
<td></td>
<td>2010/08/01</td>
</tr>
<tr>
<td>BRE-B6N/14.2g_436...</td>
<td>2009/08/05</td>
<td>Male</td>
<td>Homzygous</td>
<td>BRE-B6N-1</td>
<td></td>
<td>2010/08/01</td>
</tr>
<tr>
<td>BRE-B6N/14.1h_437...</td>
<td>2010/05/05</td>
<td>Male</td>
<td>Homzygous</td>
<td>BRE-B6N-1</td>
<td></td>
<td>2010/08/01</td>
</tr>
</tbody>
</table>

#### Measured values for parameter

<table>
<thead>
<tr>
<th>Increment</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>5.8</td>
</tr>
<tr>
<td>15.0</td>
<td>16.1</td>
</tr>
<tr>
<td>30.0</td>
<td>17.2</td>
</tr>
<tr>
<td>60.0</td>
<td>13.1</td>
</tr>
<tr>
<td>120.0</td>
<td>7.4</td>
</tr>
</tbody>
</table>
Data Flow through DCC
Transfer to CDA

- Lines presented for manual sign-off when QC completed
- Automatic transfer to CDA of data
- Status: Full test dataset sent by this method
Detailed Overview

- Coordinate
  - Tracking: Vivek Iyer
- Cross center integration
  - Phenotyping procedures and data wrangling: Henrik Westerberg
- Validate
  - Data upload and QC: Hugh Morgan
- Analyze & Archive
  - Central data archive: Terry Meehan
- Integrate & Disseminate
  - Plans for the future: Helen Parkinson
CDA Data sources and integration

- **Genome**
  - Ensembl
  - MGI

- **Alleles**
  - MGI
  - IKMC

- **Strains**
  - MGI
  - IMSR

- **IMPReSS**
  - Ontologies

Legacy Data ➔ Central Data Archive
Sources and new data types

- Genome
  - Ensembl
  - MGI

- Alleles
  - MGI
  - IKMC

- Strains
  - MGI
  - IMSR

- IMPReSS
  - Ontologies

- Human disease links

IMPC Data → Central Data Archive

IMPC

- Variation

- Statistical R Package

- Gene Expression
Statistics require transparency

- Experimental work flow capture module
  - Module refined by Statistics Working Group
  - Developed standardised vocabulary
  - Cross institute review
  - Good practice

- Housing and husbandry capture module
  - Draft module
  - Teleconference organised to confirm content

- Next steps
  - Finish modules
  - Update by centers
Statistics: Making phenotypic calls

- Review current methodologies

- Evaluate new mixed model approach
  - Pilots completed
  - Value added by collaborating
  - Compatible with all institutes workflow
  - Developed pilot script for automated multi-step process

- Next steps
  - Implement first version of production mixed model method
  - Visualization of different stats approaches
  - Design a decision tree for phenotypic call confidence
  - Build R package for dissemination and download
Image data capture

- LacZ image capability is KOMP2 priority
- Imaging data exchange protocols defined
- Agreement on ontologies (MA, EMAPA)
- Prototype image annotation/viewing tools built (both 2D and 3D).
  - Image type agnostic
  - Ontology based annotation
  - Provided to the community as open source
  - Being modularized to use on IMPC web portal
- New imaging techniques ‘coming of age’ for high throughput
  - OCT, OPT, HREM, microCT, MRI
  - High data volume anticipated
Transition to a data portal

Welcome to the International Mouse Phenotyping Consortium

SEARCH OUR GENE LIST

Use a gene symbol e.g. per

SEARCH

To find your favorite gene or browse/download the complete gene list.

The International Mouse Phenotyping Consortium (IMPC)
Comprises a group of major mouse genetics research institutions along with national funding organisations formed to address the challenge of developing an encyclopedia of mammalian gene function.

Project Status

[Bar chart showing project status]

- Grey: All Projects
- Blue: Project started
- Green: Microarray in progress
- Orange: Genotype confirmed
- Pink: Phenotype data available
Register interest in a knockout

- Email indicating interest
- Subsequent emails when data is available
- Build up user profile
- Tailored home page
Intraperitoneal glucose tolerance test (IPGTT) (Under Review) [IMPC_IPG_001]

Purpose

The glucose tolerance test measures the clearance of an intraperitoneally injected glucose load from the body. It is used to detect disturbances in glucose metabolism that can be linked to human conditions such as diabetes or metabolic syndrome. Animals are fasted for approximately 16 hours, fasted blood glucose levels are determined before a solution of glucose is administered by intra-peritoneal (IP) injection. Subsequently, the blood glucose level is measured at different time points during the following 2 hours.

Ontological description: MP:000188 - abnormal circulating glucose level.

Experimental Design

Minimum number of mutant animals: 7 mice for each sex.

Age at test: 13 weeks.

Sex: Both (sexually dimorphic)

Equipment

Equipment:

1. Mouse restraining device (optional)
2. Glucose meter
3. Scalpel blade
4. Balance
5. Timer
6. Clean cages

Supplies:

1. Glucose solution 20% (0.9 NaCl)
2. Gauge needle (25 G 5/8)
3. Syringe 1 ml
4. Glucose test-strips
5. Microvette tube
6. Tissues
7. Food pellets
8. Topical anaesthetic cream (optional) e.g. Lignocaine
Gene: Akt2

- **Gene symbol:** Akt2
- **Gene name:** Thymoma viral proto-oncogene 2
- **Gene Type:** protein coding gene
- **Synonyms:** pkb
- **Gene Location:** Chr7:2876571 - 28423845
- **Ensembl Links:** Gene View, Location View, Compare View
- **HGNC Id:** HGNC:12081

**Show/Hide More Identifiers**

**Phenotyping Panel**

- Phenotype data available in OmniPheno
- Phenotype data available in the Sanger Mouse Bank

**Expression Information**

data here for gene IN51638716

**ES Cell and Mouse Alleles**

<table>
<thead>
<tr>
<th>Product</th>
<th>Allele Type</th>
<th>Strain of Origin</th>
<th>HGNC Allele Name</th>
<th>Allele Map</th>
<th>Allele Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES Cell</td>
<td>Conditional Knock</td>
<td>C57BL/6N</td>
<td>IN51638716</td>
<td><img src="image" alt="Allele Map" /></td>
<td><img src="image" alt="Allele Sequence" /></td>
</tr>
</tbody>
</table>
IMPC Beta- Gene details page

- Scrollable, zoomable browser
- Isoforms
- Can add CpG islands, other tracks
ENSEMBL Compara viewer
• one click access
• compares mouse and human gene structure
IMPC Beta- Gene details page

Phenotyping Panel

Expression Information
data here soon for gene MGI:104874

ES Cell and Mouse Alleles

<table>
<thead>
<tr>
<th>Product</th>
<th>Allele Type</th>
<th>Strain of Origin</th>
<th>MGI Allele Name</th>
<th>Allele Map</th>
<th>Allele Sequence</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>ES Cell</td>
<td>Conditional Ready</td>
<td>C57BL/6N</td>
<td>Akt2 tm1a(KOMP)Wtsi</td>
<td><img src="image" alt="Allele Map" /></td>
<td>Genbank file</td>
<td>KOMP</td>
</tr>
</tbody>
</table>
Summary

• Robust
• Modular
• Portable
• Extensible
Detailed Overview

- **Coordinate**
  - Tracking - Vivek Iyer

- **Cross center integration**
  - Phenotyping procedures and data wrangling - Henrik Westerberg

- **Validate**
  - Data upload and QC - Hugh Morgan

- **Phenotype calls**
  - Central data archive - Terry Meehan

- **Integrate and Disseminate**
  - Plans for the future - Helen Parkinson
Targeting, disseminating, integrating

Targeting Users

Targeting Resources
  • Pulling data in, pushing data out

Adding value
  • Mining, presentation, integration
    • DiseaseFinder
    • GWAS integration

Challenges
Users

Larry the LIMS guy
Andy the biologist
Chris the clinician
Barbara the bioinformatician
User eXperience Activities

Participation in translational meetings
IMPC IT meetings

Questionnaire submitted to Infrafrontier /EUMODIC Symposium gathering Mouse/Human clinicians (January 2012)

Questionnaire submitted to MRC mouse network (June 2012)

How biologists and clinicians use phenotype web resources?
Mock-ups design
Ensembl integration

2011-2012 15,000 unique users, 150,000 page views per day
90% of Ensembl searches are for human and mouse
Top 10 visitor countries USA, UK, Germany, China, France, Spain, Japan, Canada, India, Italy
Top 10 searches are disease based

<table>
<thead>
<tr>
<th>Gene</th>
<th>Variation</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>BRAF</td>
<td>rs12979860 response to Hep C</td>
<td>Autism</td>
</tr>
<tr>
<td>BRCA1</td>
<td>rs1333049 coronary artery disease</td>
<td>Breast cancer</td>
</tr>
<tr>
<td>BRCA2</td>
<td>rs1738074 coeliac disease</td>
<td>Cancer</td>
</tr>
<tr>
<td>CFTR</td>
<td>rs1801133 many linked diseases</td>
<td>Coronary heart disease</td>
</tr>
<tr>
<td>DMD</td>
<td>rs334 sickle cell anemia</td>
<td>Cystic Fibrosis</td>
</tr>
<tr>
<td>EGFR</td>
<td>rs429358 Alzheimer’s disease</td>
<td>Diabetes</td>
</tr>
<tr>
<td>GAPDH</td>
<td>rs4988235 lactose intolerance</td>
<td>Heart disease</td>
</tr>
<tr>
<td>HBB</td>
<td>rs5743618 chlamydia</td>
<td>Osteoarthritis</td>
</tr>
<tr>
<td>KRAS</td>
<td>rs80358450 Breast cancer</td>
<td>Schizophrenia</td>
</tr>
<tr>
<td>TP53/P53</td>
<td>rs9376173 schizophrenia</td>
<td>Tay-Saches</td>
</tr>
</tbody>
</table>
Transcript: Krt76-001  ENSMUST0000100179

Description  keratin 76 [Source:MGI Symbol,Acc:MGI:1924305]
Gene  This transcript is a product of gene ENSMUSG00000075402 - This gene has 1 transcript

<table>
<thead>
<tr>
<th>Name</th>
<th>Transcript ID</th>
<th>Length (bp)</th>
<th>Protein ID</th>
<th>Length (aa)</th>
<th>Biotype</th>
<th>CCDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Krt76-001</td>
<td>ENSMUST0000100179</td>
<td>2312</td>
<td>ENSMUSP00000097754</td>
<td>594</td>
<td>Protein coding</td>
<td>CCDS27866</td>
</tr>
</tbody>
</table>

Compara Phenotypes
Description  Phenotypes linked through homologous and protein families
Homepage  http://www.ebi.ac.uk/das-srv/gene artículo/phenotypes
Ensembl Gene Accession ENSMUSG00000075402  [view DAS response]
Phenotype Driven Data Mining

**Human Data**
OMIM annotations
Orphanet
Exome sequencing projects
NHGRI Centers for Mendelian Genomics
DDD - WTSI
GWAS catalog
CNVs (Decipher)

**Model organism phenotype data**
HTP mouse data
MGI
ZFIN/ZMP
RGD

**Ontologies**
HPO
MP
PATO

**DiseaseFinder**

Candidates based on shared attributes and phenotypic distance

**MouseFinder: Candidate Disease Genes from Mouse Phenotype Data**
OMIM
http://omim.org/entry/600231
PALMOPPLANTAR KERATODERMA, BOTHNIAN TYPE; PPKB

Cytogenetic location: 12q11-q13  Genomic coordinates (GRCh37): 12:35,800,000 - 58,100,000 (from NCBI)

Gene Phenotype Relationships

<table>
<thead>
<tr>
<th>Location</th>
<th>Phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>12q11-q13</td>
<td>Palmoplantar keratoderma, Bottnia type</td>
</tr>
</tbody>
</table>

MGP
Phenotyping Data for Krt76 (MADM) Associated with 'integument' (MP:0010771)

- Tail Epidermis Wholemount: Standard Protocol - significant parameters
- Skin Histopathology: Standard Protocol - significant parameters
- Dysmorphology: Standard Protocol
- Dysmorphology: Nail Analysis - significant parameters
- Dysmorphology: Hair Follicle Cycling

MP Annotation
- MP:0001242 - hyperkeratosis
- MP:0003427 - parakeratosis

DiseaseFinder

Novel disease prediction - Krt76

<table>
<thead>
<tr>
<th>OMIM disease ID</th>
<th>OMIM disease</th>
<th>MGI gene Symbol</th>
<th>Rank</th>
<th>Score</th>
<th>Source of model organism phenotypes</th>
<th>Ensembl Gene ID</th>
<th>Associated Gene Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>600231</td>
<td>PALMOPPLANTAR KERATODERMA, BOTHNIAN TYPE</td>
<td>Krt76</td>
<td>1</td>
<td>82.36</td>
<td>MGP</td>
<td>ENSG00000195069</td>
<td>KRT76</td>
</tr>
</tbody>
</table>

IMPC
GWAS Phenotype integration

- Abdominal aortic aneurysm
- Acute lymphoblastic leukemia
- Adhesion molecules
- Adiponectin levels
- Age-related macular degeneration
- AIDS progression
- Alcohol dependence
- Alopecia areata
- Alzheimer disease
- Amyloid A levels
- Amyotrophic lateral sclerosis
- Angiotensin-converting enzyme activity
- Ankylosing spondylitis
- Arterial stiffness
- Asparagus anemia
- Asthma
- Atherosclerosis in HIV
- Atrial fibrillation
- Attention deficit hyperactivity disorder
- Autism
- Basal cell cancer
- Behcet's disease
- Bipolar disorder
- Bilirubin
- Bitter taste response
- Birth weight
- Bladder cancer
- Bleomycin sensitivity
- Blond or brown hair
- Blood pressure
- Blue or green eyes
- BMI, waist circumference
- Bone density
- Breast cancer
- C-reactive protein
- Calcium levels
- Cardiac structure/function
- Cardiovascular risk factors
- Carminic level
- Carotenoid/tocopherol levels
- Celiac disease
- Celiac disease and rheumatoid arthritis
- Cerebral atrophy measures
- Chronic lymphocytic leukemia
- Chronic myeloid leukemia
- Cleft lip/palate
- Coffee consumption
- Cognitive function
- Conduct disorder
- Colorectal cancer
- Corneal thickness
- Coronary disease
- Creutzfeldt-Jakob disease
- Crohn's disease
- Crohn's disease and celiac disease
- Cutaneous nevi
- Cystic fibrosis severity
- Dermatitis
- DHEA-s levels
- Diabetic retinopathy
- Dilated cardiomyopathy
- Drug-induced liver injury
- Drug-induced liver injury
- Endometrial cancer
- Endometriosis
- Esosiphil count
- Esosiphilic esophagitis
- Erectile dysfunction and prostate cancer treatment
- Ethylene parameters
- Esophageal cancer
- Essential tremor
- Exfoliation glaucoma
- Eye color traits
- F cell distribution
- Fibrinogen levels
- Folate pathway vitamins
- Follicular lymphoma
- Fuch's corneal dystrophy
- Freckles and burning
- Gallstones
- Gastric cancer
- Gloma
- Glycemic traits
- Hair color
- Hair morphology
- Handledness in dyslexia
- HDL cholesterol
- Heart failure
- Heart rate
- Height
- Hemostasis parameters
- Hepatic steatosis
- Hepatitis
- Hepatocellular carcinoma
- Hirschsprung's disease
- HIV-1 control
- Hodgkin's lymphoma
- Homocystinemia levels
- Hypoplasia
- Idiopathic pulmonary fibrosis
- IFN-related cytokines
- IgA levels
- IgG levels
- Inflammatory bowel disease
- Insulin-like growth factors
- Intracranial aneurysm
- Iris color
- Iron status markers
- Ischemic stroke
- Juvenile idiopathic arthritis
- Keloid
- Kidney stones
- LDL cholesterol
- Leprosy
- Leptin receptor levels
- Liver enzymes
- Longevity
- LP (a) levels
- LpPLA2 activity and mass
- Lung cancer
- Magnesium levels
- Major mood disorders
- Malaria
- Male pattern baldness
- Mamnographic density
- Matrix metalloproteinase levels
- MCP-1
- Melanoma
- Menarche & menopause
- Meningococcal disease
- Metabolic syndrome
- Migraine
- Moyamoya disease
- Multiple sclerosis
- Myeloproliferative neoplasms
- Myopia (pathological)
- N-glycan levels
- Narcolepsy
- Nasopharyngeal cancer
- Natriuretic peptide levels
- Neuroblastoma
- Nicotine dependence
- Obesity
- Open angle glaucoma
- Open personality
- Optic disc parameters
- Osteoarthrosis
- Osteoporosis
- Otochlosis
- Other metabolic traits
- Ovarian cancer
- Pancreatic cancer
- Pain
- Paget's disease
- Panic disorder
- Parkinson's disease
- Periostitis
- Peripheral arterial disease
- Personality dimensions
- Phosphatidylcholine levels
- Phosphorus levels
- Photic sneeze
- Phytosterol levels
- Platelet count
- Polygenic ovary syndrome
- Primary biliary cirrhosis
- Primary sclerosing cholangitis
- PR interval
- Progranulin levels
- Progressive supranuclear palsy
- Prostate cancer
- Protein levels
- PSA levels
- Psoriasis
- Psoriatic arthritis
- Pulmonary funct. COPD
- QRS interval
- QT interval
- Quantitative traits
- Recombination rate
- Red vs. non-red hair
- Refractive error
- Renal cell carcinoma
- Renal function
- Response to antidepressants
- Response to antipsychotic therapy
- Response to carbamazepine
- Response to clopidogrel therapy
- Response to hepatitis C treatment
- Response to interferon beta therapy
- Response to metformin
- Response to statin therapy
- Restless legs syndrome
- Retinal vascular caliber
- Rheumatoid arthritis
- Ribavirin-induced anemia
- Schizophrenia
- Serum metabolites
- Skin pigmentation
- Smoking behavior
- Speech perception
- Sphingolipid levels
- Statin-induced myopathy
- Stroke
- Sudden cardiac arrest
- Suicide attempts
- Systemic lupus erythematosus
- Systemic sclerosis
- Tau levels
- Tau A81-42 levels
- Telomere length
- Testicular germ cell tumor
- Thyroid cancer
- Thyroid volume
- Tooth development
- Total cholesterol
- Triglycerides
- Tuberculosis
- Type 1 diabetes
- Type 2 diabetes
- Ulcerative colitis
- Urate
- Urinary albumin excretion
- Urinary metabolites
- Uterine fibroids
- Venous thromboembolism
- Venous thrombosis
- Ventricular conduction
- Vertical cup-disc ratio
- Vitamin B12 levels
- Vitamin D insufficiency
- Vitiligo
- Warfarin dose
- Weight
- White cell count
- White matter hyperintensity
- YKL-40 levels
Cross species integration challenges

- 90% queries GWAS are about diseases
- GWAS data have measurement traits with implicit links to disease
- Connecting anatomy, phenotype, disease, measurements is essential
- Cross species phenotypic queries implemented in existing resources e.g. Ensembl becomes feasible
- Human curated data for human and mouse is essential
- GWAS is a good use case to consider in dissemination of mouse data
  - Measuring similar phenotypes
  - Can be consumed by DiseaseFinder
Where do our users go?

C-reactive protein

From Wikipedia, the free encyclopedia
(Redirected from C-reactive protein)

C-reactive protein (CRP) is a protein found in the blood, the levels of which inflammation (i.e., C-reactive protein is an acute-phase protein). Its physiological function is expressed on the surface of dead or dying cells (and some order to activate the complement system via the C1q complex.\[^1\]

CRP is synthesized by the liver\[^2\] in response to infection (adipocytes).\[^3\] It is a member of the pentraxin protein family. C-reactive protein was the first acute-phase protein to be identified.

Contents [hide]
1 History and nomenclature
2 Genetics and biochemistry
3 Function
4 Clinical significance
4.1 Role in cardiovascular disease
4.2 Role in cancer
5 Diagnostic use
5.1 Cardiology diagnostic test
6 See also
7 Additional images
8 References
9 External links

Palmoplantar keratodermas

From Wikipedia, the free encyclopedia

Palmoplantar keratodermas are a heterogeneous group of disorders characterized by abnormal thickening of the palms and soles.\[^4\]

Autosomal recessive and dominant, X-linked, and acquired forms have all been described.\[^5\]\[^6\]\[^7\]\[^8\]

There are also acquired forms of the condition.\[^9\]\[^10\]

Knockout mouse

A knockout mouse is a genetically engineered mouse in which researchers have inactivated, or "knocked out," an existing gene by replacing it or disrupting it with an artificial piece of DNA. The loss of gene activity often causes changes in a mouse's phenotype, which includes appearance, behavior, and other observable physical and biochemical characteristics.

Knockout mice are important animal models for studying the role of genes which have been sequenced but whose functions have not been determined. By expressing a specific gene to be inactive in the mouse, and observing any differences from normal behavior or physiology, researchers can infer its probable function.

Knockout mice are currently the most closely related laboratory animal species to humans for which the knockout technique can easily be applied. They are widely used in knockout experiments, especially those investigating genetic questions that relate to human physiology. Gene knockout in rats is much harder and has only been possible since 2003.\[^11\]\[^12\]\[^13\]

The first described knockout mouse was created by Mario R. Capocchi, Martin Evans and Oliver Smithies in 1983, for which they were awarded the Nobel Prize for Medicine in 2007. Aspects of the technology for generating knockout mice, and the mice themselves, have been patented in many countries by private companies.

Contents [hide]
1 Use
2 Strains
3 Procedure
4 Limitations
5 See also
6 References
7 External links
Challenges

Intuitive presentation of data in context

New imaging modalities, volumes

Usable ontologies and annotation

Shifting user focus from gene to phenotype

Secondary phenotyping data integration
Meeting the challenges

- Developing context specific views
- Pushing data to relevant resources and exposing services from CDA
- Leverage collaborative projects, existing resources
  - GWAS, DiseaseFinder, HPO etc
- Collaboration with ontologists
  - Ontology workshops
  - Cross species query use cases
- Training, outreach, user experience
  - Mouse users
  - Translational users
DATA TRACKING
Basic tracking system available to production centers
Advanced tracking system
Tracking portal

SOPDB
Review SOPS proposed by centers and complete version 1
Version 1 of SOPDB
Version 2 of SOPDB based on new guidelines
Extend SOPDB web portal with use case defined functions
Refine SOP definitions and manage versioning

Pheno-DCC
Review data export and upload process
Set up LIMS export with centers
Design and implement Pheno-DCC database schema
Develop export module
Validation and QC tool development
Image upload tools
Image annotation specification and tool development
Data management system complete and exported to portals

Statistical Analysis and Data Annotation
Review experimental design of SOPs
Modular Annotation Pipeline V1
Incorporation of EBI R analysis infrastructure
Incorporation of added value data sets
Modular Annotation Pipeline V2

CDA and IT infrastructure
Data warehouse

KOMP2 Web Portal
High priority portal use cases (D4.2.1)

Medium priority portal use cases (D4.2.2)

Low priority portal use cases (D4.2.3)
Further development incorporating new use cases
Funding

U54 HG006370-02 - KOMP2

U41 HG006104-01S1 – GWAS Collaboration
Tell us what you think

IMPC site
http://mousephenotype.org/

IMPC beta site
http://beta.mousephenotype.org/

GWAS Catalog
http://wwwdev.ebi.ac.uk/fgpt/gwas/