



An International Centre for Mouse Genetics



MIMPP – **M**inimal **I**nformation for **M**ouse **P**henotyping **P**rocedures

Ann-Marie Mallon

Bioinformatics Group, MRC Harwell

InterPhenome

<http://www.interphenome.org>



- **Began in Feb 2006 (Barcelona) – International Collaboration**

- 1 MRC Mammalian Genetics Unit, Harwell, Oxfordshire OX11 0RD, UK
- 2 Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA, UK.
- 3 BSRC Fleming, 34 Fleming Street, 16672, Vari, Athens, Greece
- 4 The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609, USA
- 5 Oak Ridge National Laboratory, Mammalian Genetics & Genomics Group, Biosciences Division, Oak Ridge, TN 37831-6445
- 6 MRC Human Genetics Unit, Edinburgh, EH4 2XU, UK
- 7 GSF – National Research Center for Environment and Health, Ingolstaedter Landstrasse 1, Institute of Experimental Genetics, D-85764, Neuherberg, Germany
- 8 Department of Genetics, University of Cambridge, Cambridge CB2 3EH, UK
- 9 Institut Clinique de la Souris, BP10142, 67404 Illkirch Cedex, France
- 10 RIKEN Genomic Sciences Center, 3-1-1 Koyadai, Tsukuba, Ibaraki, Japan
- 11 Helmholtz-Zentrum für Infektionsforschung GmbH, Inhoffenstraße 7, D-38124, Braunschweig, Germany
- 12 Faculty of Life Sciences, University of Manchester, Manchester, M13 9PT, UK
- 13 European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SD, UK
- 14 Mouse Biology Unit, EMBL Monterotondo Outstation, via Ramarini 32, 00016 Monterotondo, Rome, Italy
- 15 Department of Physiology, Development and Neuroscience, University of Cambridge, Downing Street, Cambridge CB2 3DY UK
- 16 Oxford University Computing Laboratory, Wolfson Building, Parks Road, Oxford, OX1 3QD, UK.
- 17 MRC Mary Lyon Centre, Harwell, Oxfordshire OX11 0RD, UK



An International Centre for Mouse Genetics



InterPhenome

<http://www.interphenome.org>



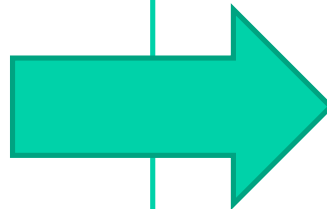
- Outcomes/Requirements
 - Standards for the description of phenotypes
 - Minimum information required for the description of a phenotyping experiment/protocol
 - Data exchange file formats
 - Integration of mouse phenome databases
- Paper (**Mouse Phenotype Database Integration Consortium** (2007) Integration of mouse phenome data resources. *Mamm Genome* **18**: 157-163.)
- Next Meeting - IMGIC Prague



An International Centre for Mouse Genetics

MI checklist and data format for reporting a phenotyping standard operating procedure

ID
Edit history
Type
Title
Source Information
Contact
Documentation
Purpose
Associated ontology terms
Emergency Procedures
(Investigator) Notes
Quality control
Technical Requirements
Equipment
Consumables
Procedure.
Workflow (Order of Testing)
Animal Husbandry
Time and Capacity
Experimental Procedure
Procedural parameters
Appropriate experimental controls
Parameters & Units of Measurement.



Data Format
PPML (Phenotyping Procedure
XML Schema)

(Hiroshi Masuya, RIKEN))



- Database of SOPs, developed by the EUMORPHIA consortium, that can be used to describe the phenotype of a mouse
- 86 SOPs
 - These cover all of the main body systems including; clinical chemistry, hormonal and metabolic systems, cardiovascular, allergy and infection, renal function, sensory function, neurological and behavioral function, cancer, bone and cartilage, and respiratory function.
- Requirement for standards to allow new submissions by users and database interoperability.
- Currently implementing the PPML data format in EMPreSS for the EMPreSSlim Pipeline



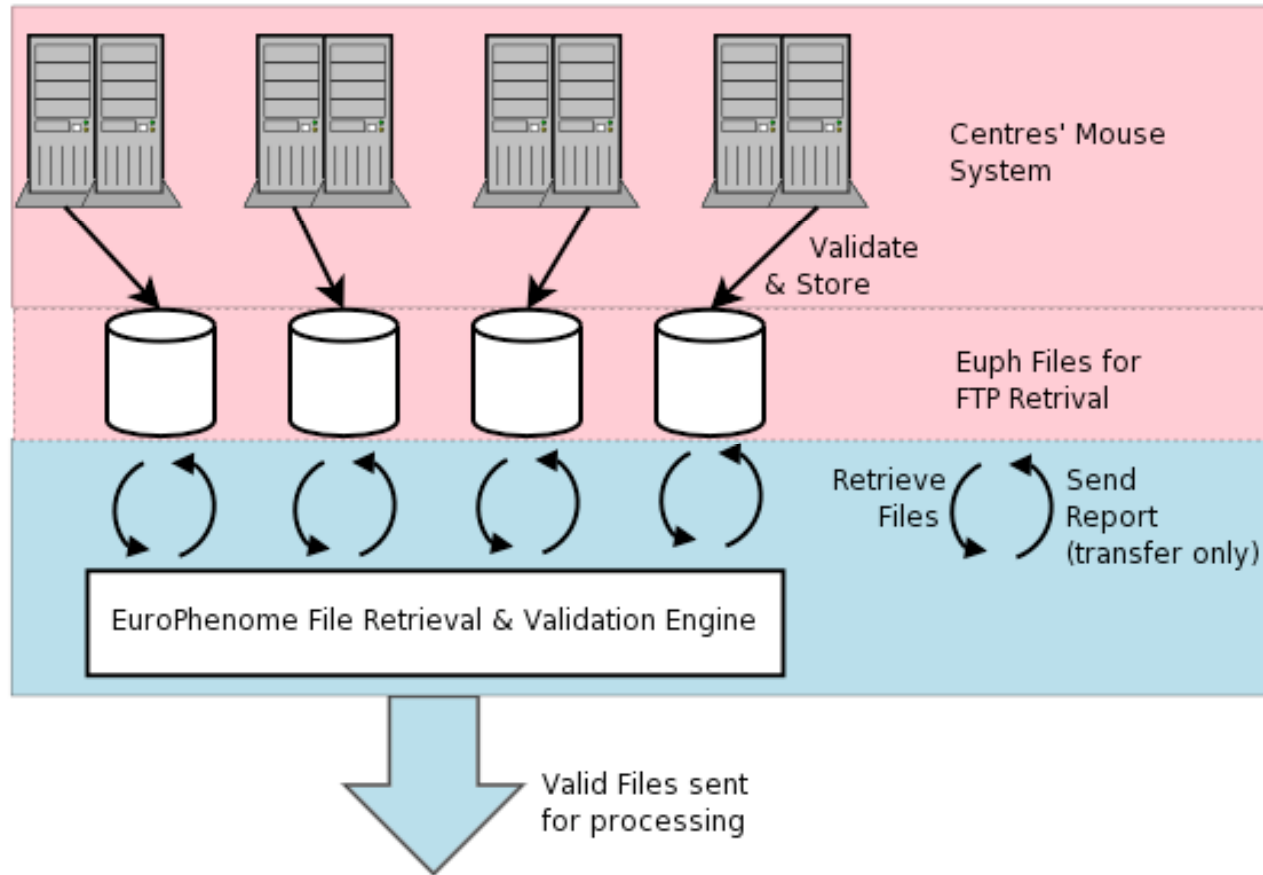
EuroPhenome



- EuroPhenome is a project to develop a repository for capturing and storing raw mouse phenotyping data as well as developing user interfaces and visualisation tools for querying and viewing this data.
- Data captured in EuroPhenome **if the SOP** is defined in EMPReSS
- EUMODIC
 - Primary high-throughput phenotype assessment using the EMPReSSlim Pipeline of up to 650 mouse mutant lines from EUCOMM.
 - Data generated in four mouse clinics across Europe



Data Capture



Data Formats / XML Schemas



- XML Schemas
 - Centre / Environmental Data
 - Cohort / Individual Animal Data
 - Procedure / Parameter Data
 - Experimenter, Date Stamp, Start Time
 - Parameter types [simple, series, media, metadata]
 - Schemas include overall metadata relevant to all the procedures in the pipeline



EMPreSSlim procedures

- For all 21 procedures in EMPReSSlim
 - Defined the content required to report each procedure - Excel!
 - Measured Parameters, Derived Parameters, Metadata, **Ontology Annotations**
 - Stored in EMPReSS
 - Used to validate the data entry into EuroPhenome from the procedure schema
 - Allows us to analyse data across the clinics
 - Developed with experts from 4 mouse clinics in EUMODIC



EMPReSSlim Pipeline

Assay Name	ESLIM-ID	Parameter measured		Increment Value	Units	Datatype	Bounds/Options	Required		
Pipeline 1 Morphology Metabolism Dysmorphology Cardiovascular	IPGTT	ESLIM_004_001	Name	Abbreviation						
		ESLIM_004_001_001	Body Weight		g	float		yes		
		ESLIM_004_001_002	Blood glucose concentration	Glu-conc		0mmol/L	float	yes		
						15mmol/L	float			
						30mmol/L	float			
						60mmol/L	float			
						120mmol/L	float			
Specific Assay Metadata										
Assay Name	ESLIM-ID	Parameter measured		Abbreviation	Increment Value	Units	Datatype	Bounds/Options	Required	Derived Parameters Calculated
Indirect Calorimetry	ESLIM_003_001	Name								Name
	ESLIM_003_001_001	Body mass before experiment				g	float		yes	
	ESLIM_003_001_002	Body mass after experiment				g	float		yes	
	ESLIM_003_001_003	Oxygen consumption	VO2			date/time ml/h/animal			yes	
	ESLIM_003_001_004	Carbon dioxide production	VCO2			date/time ml/h/animal			yes	
	ESLIM_003_001_006	Heat Production (Metabolic Rate)	HP			date/time kcal/h/animal or kJ/h/animal			yes	
	ESLIM_003_001_007	Ambulatory activity (no of beam cuts)	xamb			date/time count/interval			no	
	ESLIM_003_001_008	Total activity (no of fine movement + no of beam cuts)	xtot			date/time count/interval			no	
	ESLIM_003_001_009	Ambulatory activity (no of beam cuts)	yamb			date/time Number (count/interval)			no	
	ESLIM_003_001_010	Total activity (no of fine movement + no of beam cuts)	ytot			date/time Number (count/interval)			no	
	ESLIM_003_001_011	Total Food intake				g			no	
	ESLIM_003_001_012	Food intake				date/time g			no	
										Respiratory Exchange Ratio
Specific Assay Metadata										
	ESLIM_003_001_801	Equipment name		Text						
	ESLIM_003_001_802	Equipment manufacturer		Text						
	ESLIM_003_001_803	Equipment model		Text						
	ESLIM_003_001_804	Normal room temperature span								
	ESLIM_003_001_805	Room Temperature								



Summary MI to date



- MI Checklist for reporting a SOP (General to all procedures) - InterPhenome
- Defined MI/Core Information for 21 EMPReSSlim procedures – EUMODIC
- Checklists in many formats/files – requires extraction
- Move from data formats to guidelines so as any researcher can use them (independent of data format)



Discussion

- MI/Content (working with MIBBI)
 - How to present these
 - How to get expert advice
 - Process of review
 - How to convince users and data submitters these are important.
 - Who to bring on board – vendors/industry
- Data Formats
 - PPML
 - Fuge
- Ontologies
 - Assay/Procedure
 - OBI
 - Exact



Acknowledgements

- Interphenome consortium
- EUMODIC (WTSI, HMGU, ICS, MLC)
- Bioinformatics Group, MRC Harwell
 - Hugh Morgan
 - Tim Beck
 - Andrew Blake
- John Hancock
- Hilary Gates

