



IMPC

International Mouse Phenotyping Consortium



@impc



@GeneOfTheDay

The International Mouse Phenotyping Consortium (IMPC): a large-scale functional catalogue of mammalian genes

Tudor Groza

Phenomics Team Lead
EMBL-EBI

Open Science and FAIR Data for Neuroscience
Torino, 9 June 2022

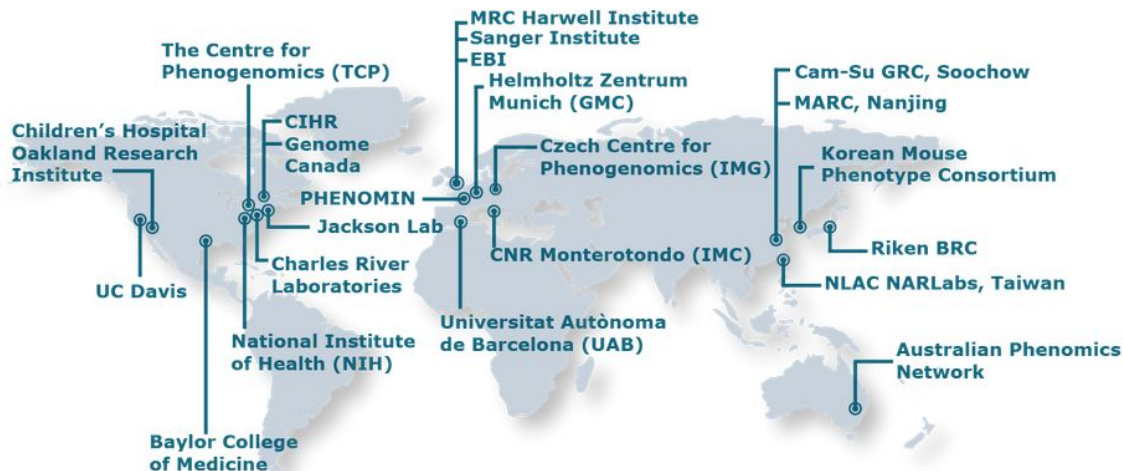


mousephenotype.org

EMBL-EBI



IMPC goals and organization

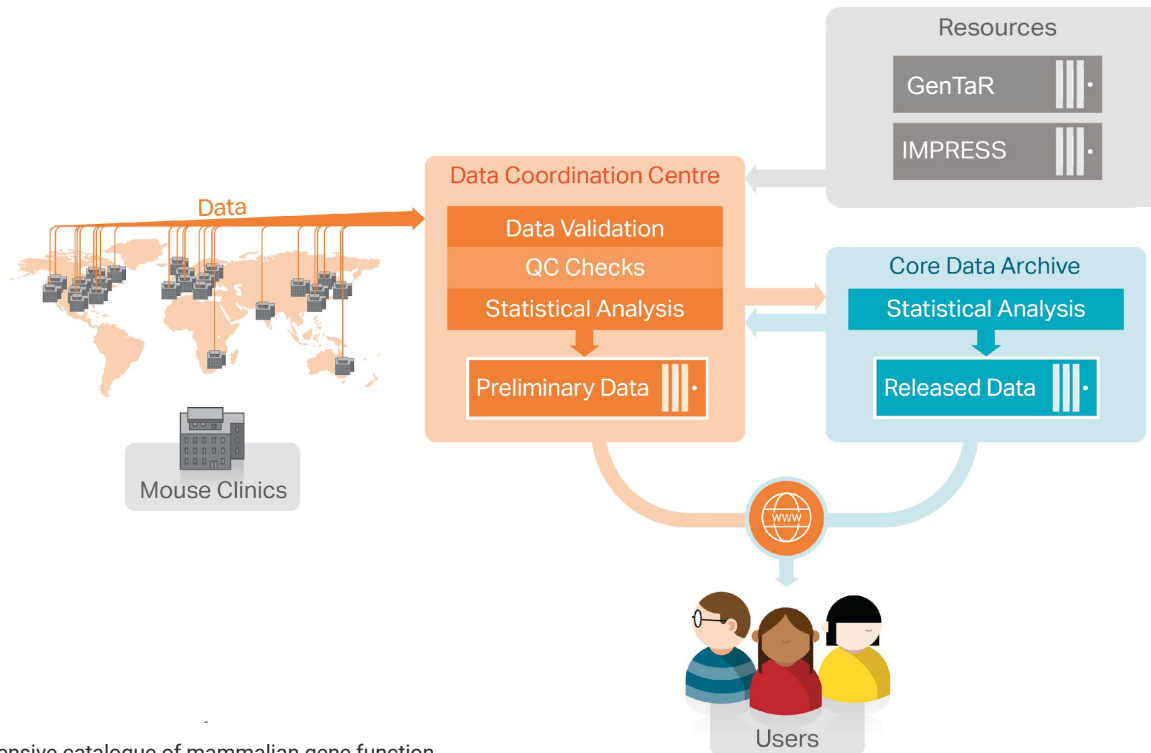


- Goal is to make **KO mice** of virtually every protein-coding gene (~18,000 genes)
- Focus on poorly characterized genes, **the ignorome**
- **8000 KO lines** produced and characterized

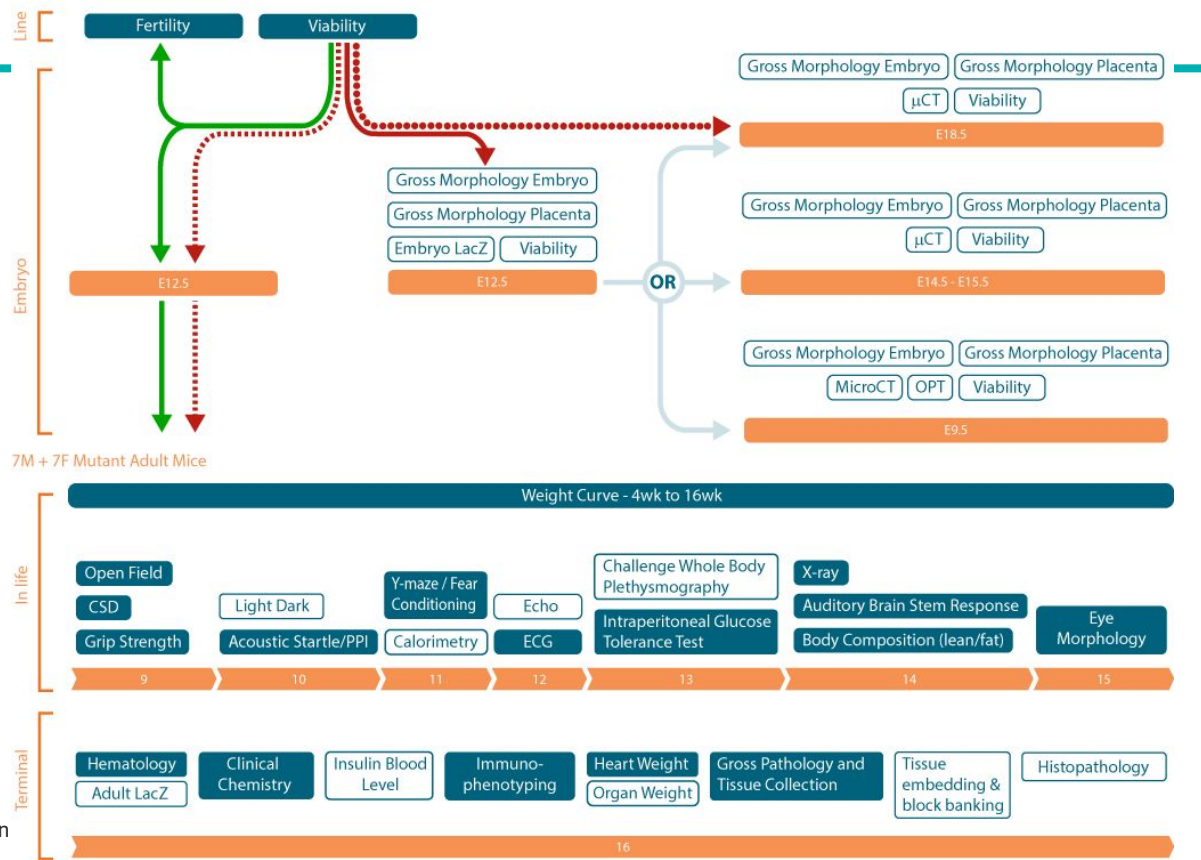


Data generation

IMPC Data flow



IMPRESS



7M + 7F Mutant Adult Mice



GenTaR



GenTaR

[Search](#)
[Reports](#)
[Help](#)


Show Filters

Projects search

Enter marker symbols/MGI accession ids or upload a file with the terms to search:

E.g: MGI:1888496, Atrx, Lbr

or upload a csv file (every search term in a new line)

 No file chosen

[Download as CSV](#)

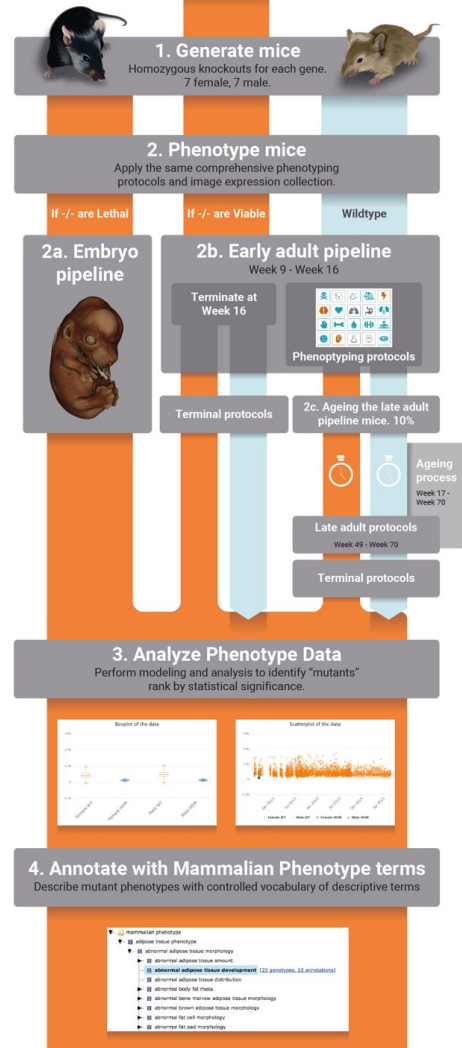
Project summary	Gene Symbol / Location	MGI	Intention	Best Human ortholog	Work Unit(s)	Work Group(s)	Project Assignment	Summary Status	Colony Name(s)	Phenotyping External Reference(s)	Privacy	Consortia	Access Status
TPN:000000001	Otog	MGI:1202064	Deletion	OTOG	JAX	JAX	Assigned	Phenotyping Started	jr34077	JR34077	public	IMPC CMG	✓
TPN:000000002	Dennd6b	MGI:1916690	Deletion	DENND6B	JAX	JAX	Assigned	Phenotyping Started	jr34293	JR34293	public	CMG IMPC	✓
TPN:000000003	Zmym2	MGI:1923257	Deletion	ZMYM2	JAX	JAX	Assigned	Phenotyping Started	jr34017	JR34017	public	CMG IMPC	✓
TPN:000000004	Sox9	MGI:98371	Deletion	SOX9	UCD	DTCC	Assigned	Founder Obtained			public	CMG IMPC	✓



Creating a comprehensive catalogue of mammalian gene function.

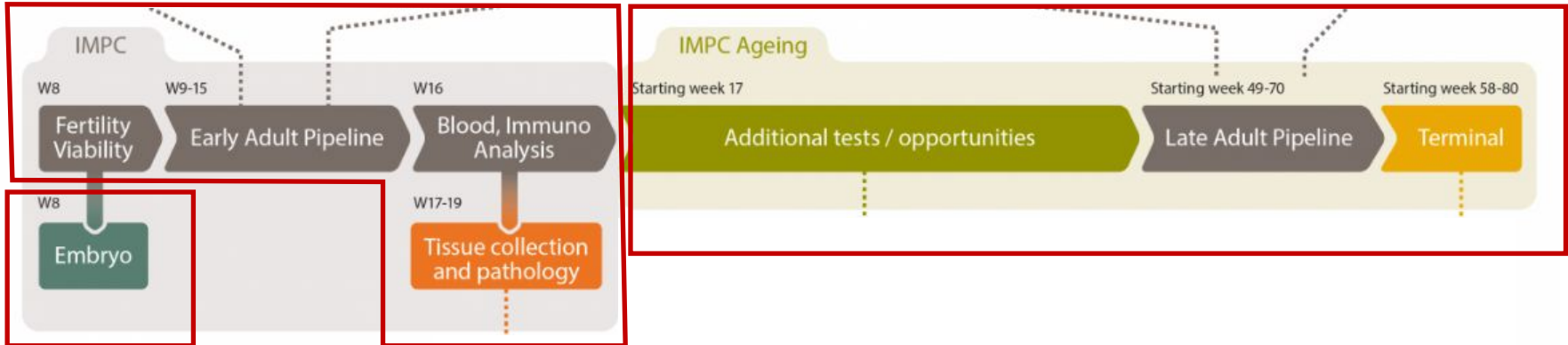
How IMPC data is generated

- **Production** of KO mice and validation
- **Phenotype** KO and wildtype mice
- Application of a **standardized phenotyping pipeline, at 3 stages**: embryonic, early adult, late adult
- Analyze data with our **custom statistical pipeline**, comparing KO with WT mice
- Annotate **significant phenotypes** using the **Mammalian Phenotype ontology**





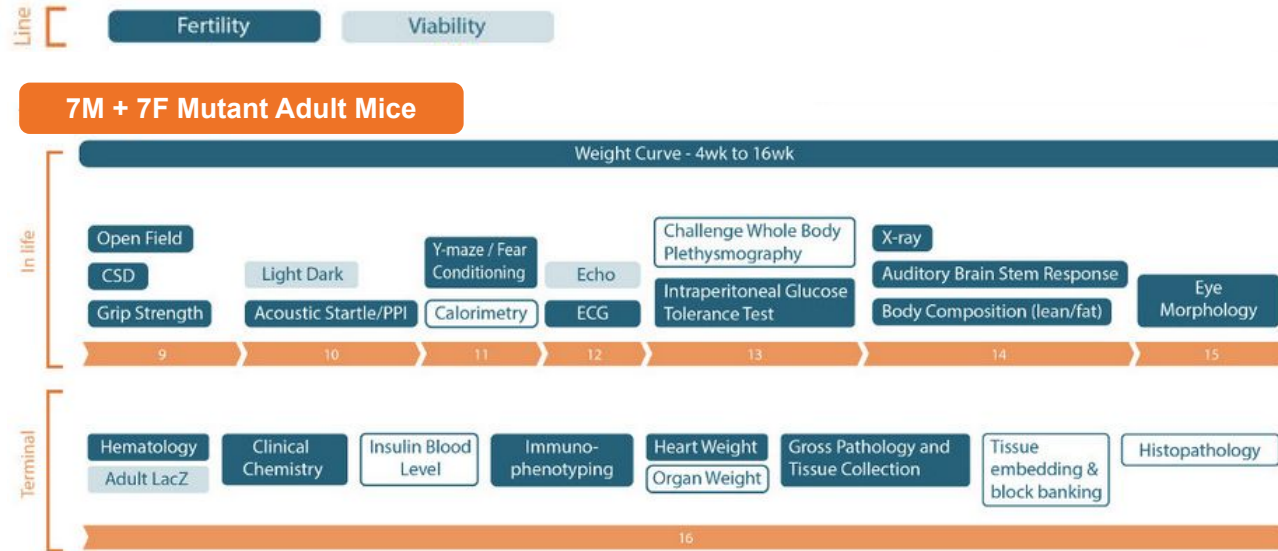
Phenotyping adult and embryo mice



- **Standardized phenotyping pipelines** across IMPC partner centers
- **Early adult** pipeline: applied to all; depending on outcome: selected lines are phenotyped in the embryonic or late adult pipelines



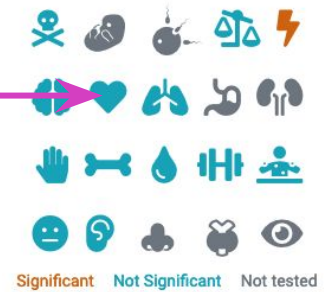
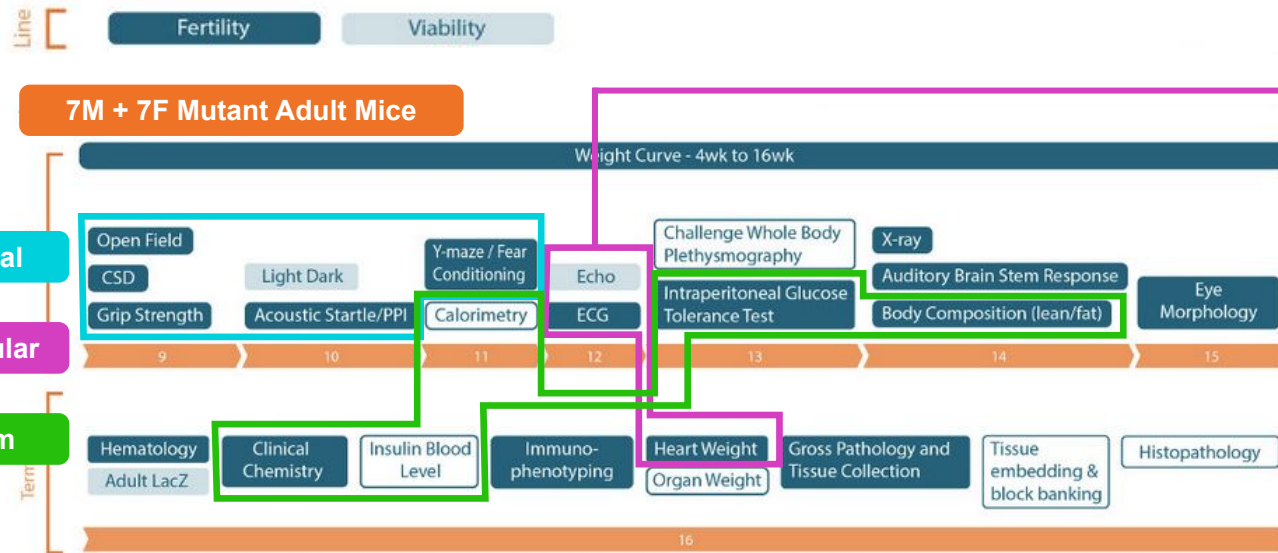
Early adult phenotyping pipeline



- Comprehensive, cover **multiple physiological systems**
- Applied to **mutants and wildtypes** (controls)
- Deviant parameters inform about the **physiological system** that is disrupted when a gene is disabled



Early adult phenotyping pipeline



- Procedures grouped together to inform about **a physiological system**
- Represented by **icons** in the gene pages of the website

Gene pages


- **Cib2^{tm1b}(EUCOMM)Wtsi**
- Phenotype summary with icons
- Significant phenotypes at a quick glance
- Based on the Mammalian Phenotype Ontology
- Register interest in a gene


<https://www.mousephenotype.org/data/genes/MGI:1929293>





Gene: Cib2 ? [Log in to follow](#)


Name calcium and integrin binding family member 2
MGI ID MGI:1929293
Synonyms calcium binding protein Kip2 2810434I23Rik
Viability Homozygous - Viable
Embryo viewer N/A
Other links [MGI](#) [Ensembl](#)


 Significant phenotypes (9)


 Measurements chart (245)

 All data table (472)

 Expression & images (67)




 Disease models (643)

 Order (4)








Significant
 Not Significant
 Not tested

View body weight measurements

 Significant phenotypes (9/9)
 Measurements chart (245/245)
 All data table (472/472)

✕

Phenotype	System	Allele	Zyg	Sex	Life Stage	P Value
polycystic kidney		Cib2 ^{tm1b} (EUCOMM)Wtsi	HET	♂	Late adult	0.00
increased circulating HDL cholesterol level		Cib2 ^{tm1b} (EUCOMM)Wtsi	HOM	♀♂	Early adult	1.04x10 ⁻⁰⁷
increased circulating cholesterol level		Cib2 ^{tm1b} (EUCOMM)Wtsi	HOM	♀♂	Early adult	1.67x10 ⁻¹⁰
abnormal ear morphology		Cib2 ^{tm1b} (EUCOMM)Wtsi	HOM	♀♂	Early adult	4.18x10 ⁻¹⁰
decreased startle reflex		Cib2 ^{tm1b} (EUCOMM)Wtsi	HOM	♀♂	Early adult	5.47x10 ⁻¹⁴

Gene pages

- In-depth view of the data

<https://www.mousephenotype.org/data/genes/MGI:1929293>



Cib2 data chart

? Log in to follow

Home > Genes > Cib2 > Clinical Chemistry / Total cholesterol ?

Description of the experiments performed

A **Clinical Chemistry** phenotypic assay was performed on **4249 mice**. The charts show the results of measuring **Total cholesterol** in **9 female, 13 male mutants** compared to **2106 female, 2121 male controls**. The mutants are for the **Cib2^{tm1b(EUCOMM)Wtl}** allele.

* The high throughput nature of the IMPC means that large control sample sizes may accumulate over a long period of time. See the [animal welfare guidelines](#) for more information.

Life stage **Early adult** ?

Associated Phenotype **Increased circulating cholesterol level**

Testing protocol **Clinical Chemistry**

Measured value **Total cholesterol**

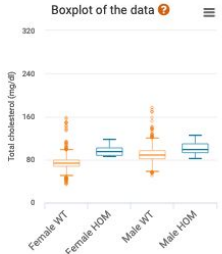
Testing environment **Lab conditions and equipment**

Background Strain **involves: C57BL/6NTac**

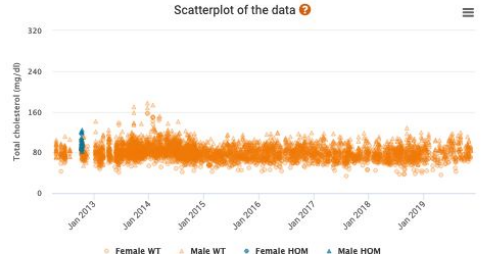
Phenotyping center **MRC Harwell**

Mouseover the charts for more information. Click and drag to zoom the chart. Click on the legends to disable/enable data.

Boxplot of the data ? ≡



Scatterplot of the data ? ≡



Results of statistical analysis

P value
1.67×10⁻¹⁰

Classification
Both genders equally

Genotype effect	P Value	Effect size	Standard Error
Female WT			
Female HOM			
Male WT			
Male HOM			

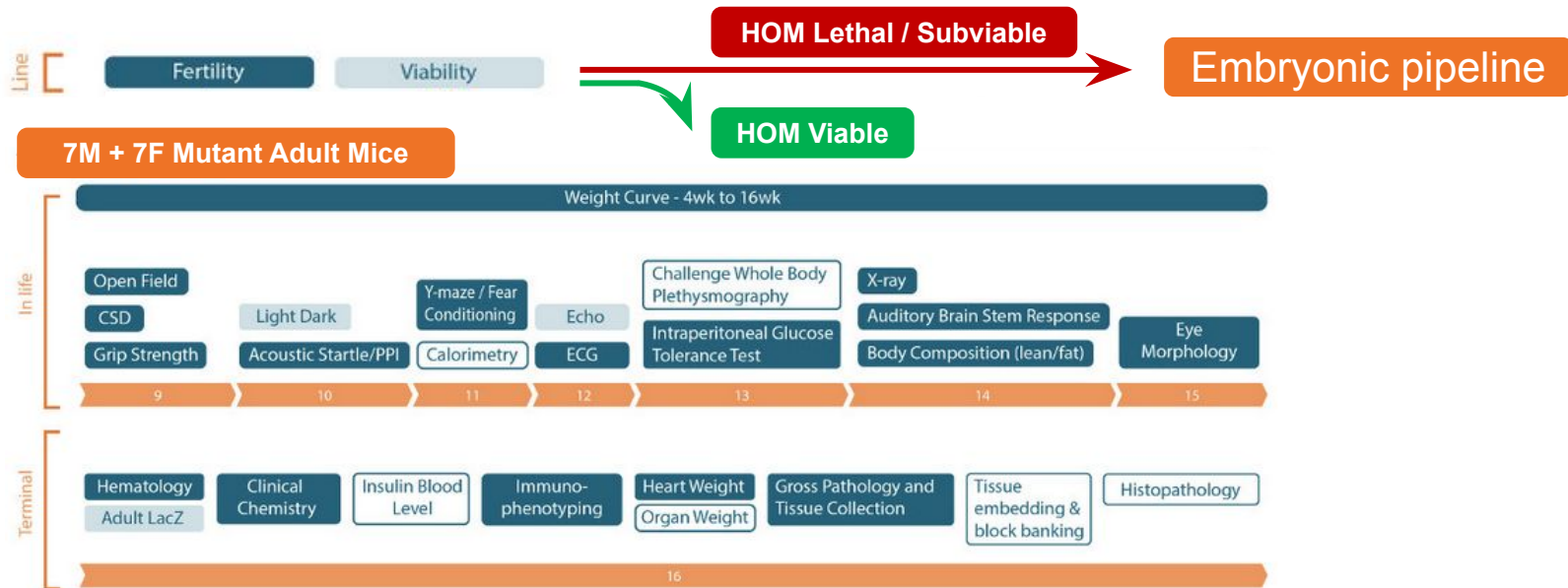
Summary statistics of the dataset

	Mean	Stdev	# Samples
Female Control	73.3795	11.4818	2106
Female homozygote	95.8385	10.4063	9
Male Control	89.223	13.1834	2121
Male homozygote	101.1581	13.6379	13

Genotype	P Value
Female Control	0.00
Female homozygote	1.04×10 ⁻⁰⁷
Male Control	1.67×10 ⁻¹⁰
Male homozygote	4.18×10 ⁻¹⁰
Female homozygote vs Male homozygote	5.47×10 ⁻¹⁴



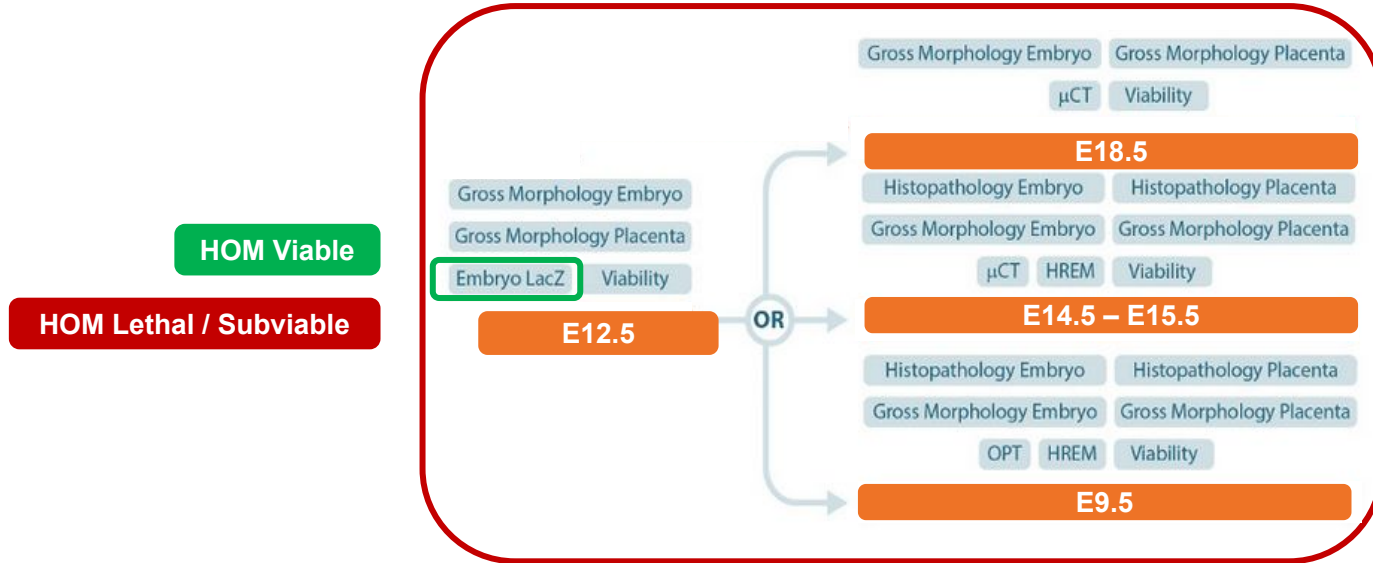
Early adult phenotyping pipeline



- **Embryonic lethals** assessed in specialized pipeline
- **Heterozygotes** of embryonic lethals phenotyped in adult pipeline



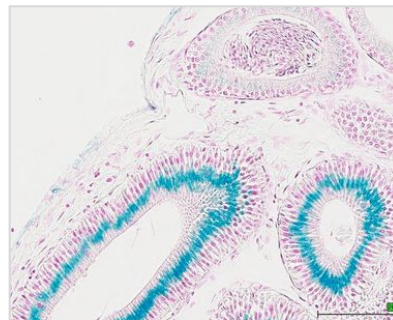
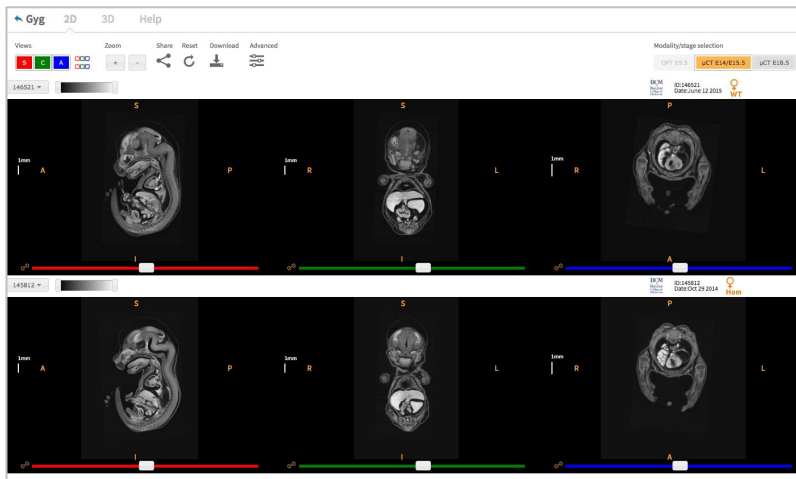
Embryonic pipeline



- Approx. **third of lines** are embryonic lethal
- Specialized pipeline created
- Structural changes during **embryonic development**

Embryo development and viability

- Determine defects that occur during development
 - Embryonic data focus pages
 - 2D and 3D images
- LacZ expression

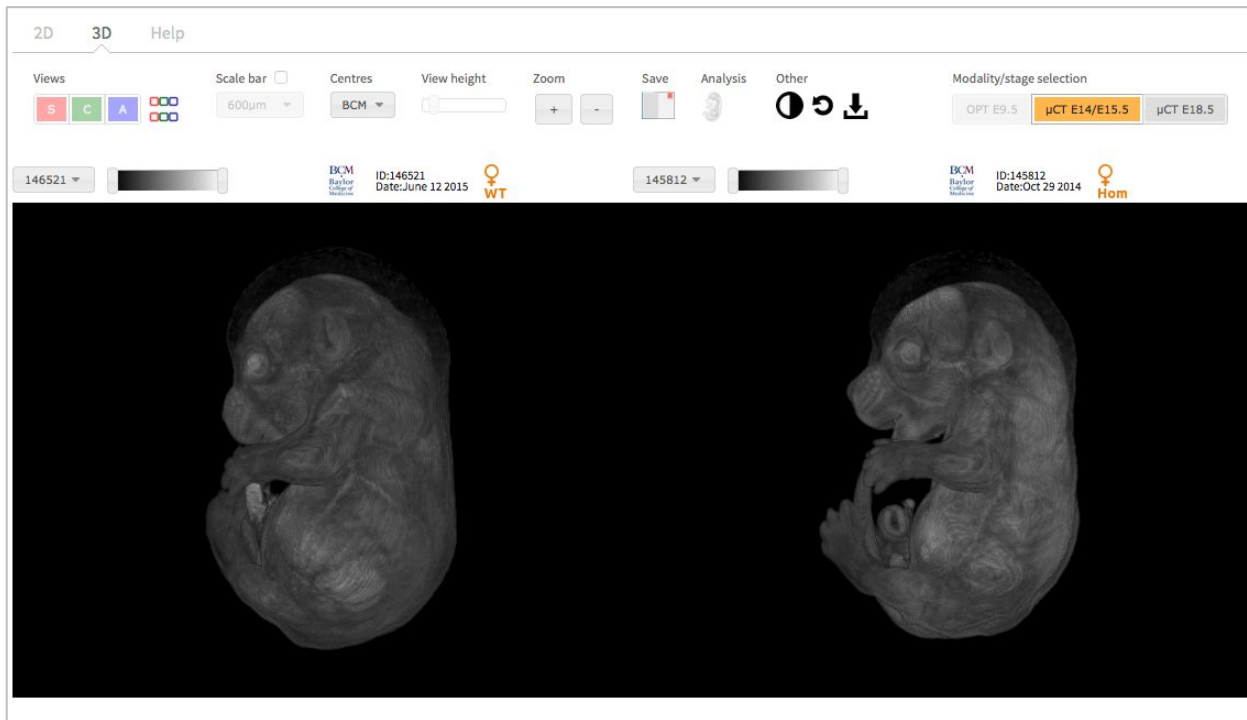


LacZ expression in the epididymis of *Rspo1^{tm1.1(KOMP)VICg}* heterozygotes. Gyg KO embryo viewed using the Interactive Embryo Viewer (IEV) viewer.



Single images from E18.5 microCT volumes showing spinal cord abnormalities (arrow), enlarged thymus (asterisk) and thickened myocardium (arrowhead) in homozygous null embryos compared to wild-type littermates.

3D image view



The screenshot displays the IMPC 3D image view interface. At the top, there are tabs for "2D", "3D", and "Help". Below the tabs is a control panel with several sections:

- Views:** Includes buttons for "S" (Screenshot), "C" (Center), "A" (Analyze), and a grid of view icons.
- Scale bar:** A checkbox and a dropdown menu showing "600µm".
- Centres:** A dropdown menu currently set to "BCM".
- View height:** A slider control.
- Zoom:** "+" and "-" buttons.
- Save:** A button with a document icon.
- Analysis:** A button with a brain icon.
- Other:** Buttons for "Refresh", "Reset", and "Download".
- Modality/stage selection:** A row of buttons for "OPT E9.5", "µCT E14/E15.5" (highlighted in orange), and "µCT E18.5".

Below the control panel, there are two specimen information blocks:

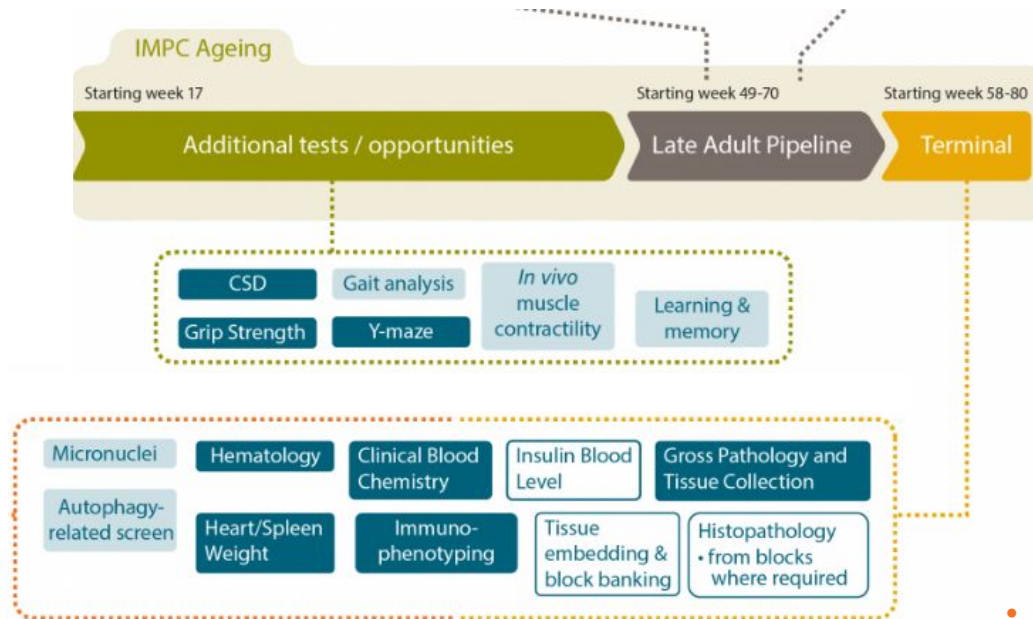
- Specimen 1:** ID: 146521, Date: June 12 2015, Genotype: WT (indicated by a female symbol and "WT").
- Specimen 2:** ID: 145812, Date: Oct 29 2014, Genotype: Hom (indicated by a female symbol and "Hom").

The main area of the interface shows two 3D rendered mouse embryos in a dark background, one on the left and one on the right, both shown in a lateral view.





Late adult phenotyping pipeline



- Effects of **ageing**
- **Adult pipeline** repeated
- Specialized analysis implemented

Late Adult phenotypes

- Gene page
- Specialized heat map

Gene: Cib2 ? Log in to follow

Name: calcium and integrin binding family member 2
 MGI ID: MGI:1929293
 Synonyms: calcium binding protein Kip2 2810434I23RIK
 Viability: Homozygous - Viable
 Embryo viewer: N/A
 Other links: MGI, Ensembl


Significant phenotypes (9) | Measurements chart (245) | All data table (472)

In the All data table, "Late Adult" can be selected

Viewing: all phenotypes

All Embryo Early adult Mid adult **Late adult**

Allele	Center	Procedure / Parameter	Life stage	Zyosity	Significant	P Value	F
Cib2 ^{tm1MEUCOMM/wvl}	MRC Harwell	Gross Pathology and Tissue Collection / Kidney	Late adult	HET	Significant	-	po kid
Cib2 ^{tm1MEUCOMM/wvl}	MRC	Gross Pathology and Tissue	Late adult	HOM	Not significant	-	-


Help IMPC Cloud Contact us Search

[MY GENES](#)
[ABOUT THE IMPC](#)
[DATA](#)
[HUMAN DISEASES](#)
[PUBLICATIONS](#)
[NEWS](#)
[BLOG](#)

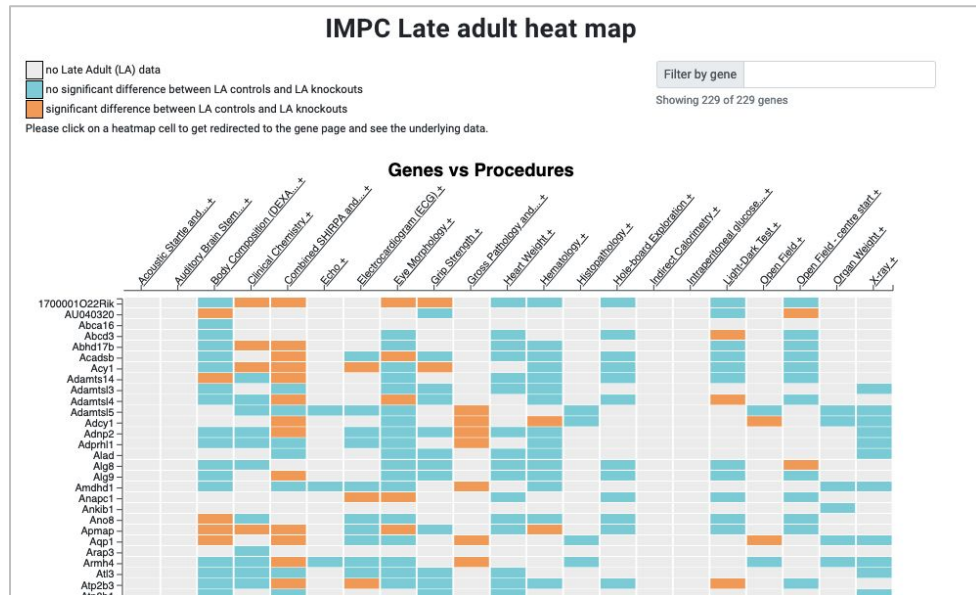
COVID-19

- Understanding Our Data
- Allele design
- Citing IMPC Data
- How We Generate Our Data
- Phenotyping Process

In the Data menu, "Late Adult" can be selected

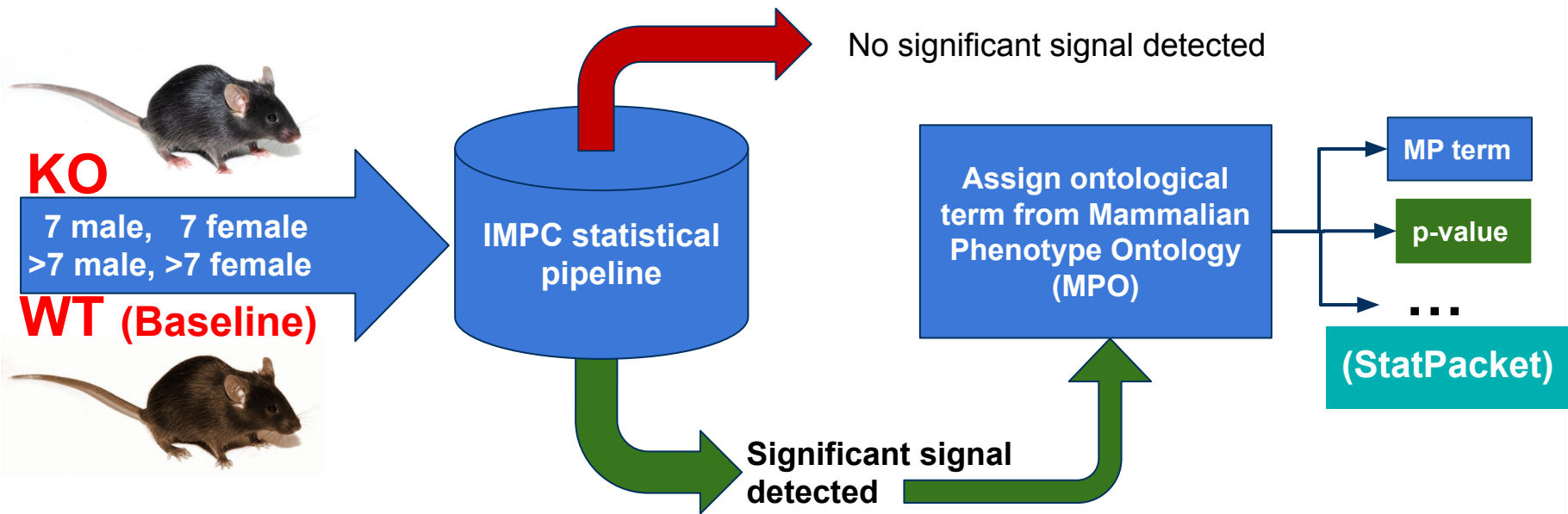
Data Focus

- Cardiovascular
- Embryo Development
- Embryo Imaging
- Embryo Vignettes
- Essential genes
- Histopathology
- Late Adult Data
- Pain



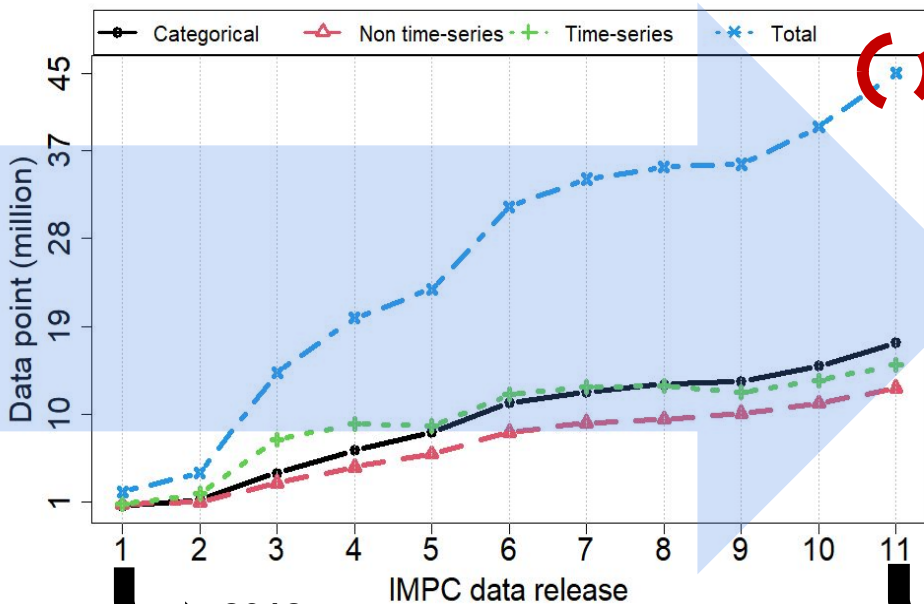
IMPC Statistical pipeline

IMPC statistical pipeline is meant to detect phenotype effect



IMPC Statistical pipeline

Fast growing data in the IMPC



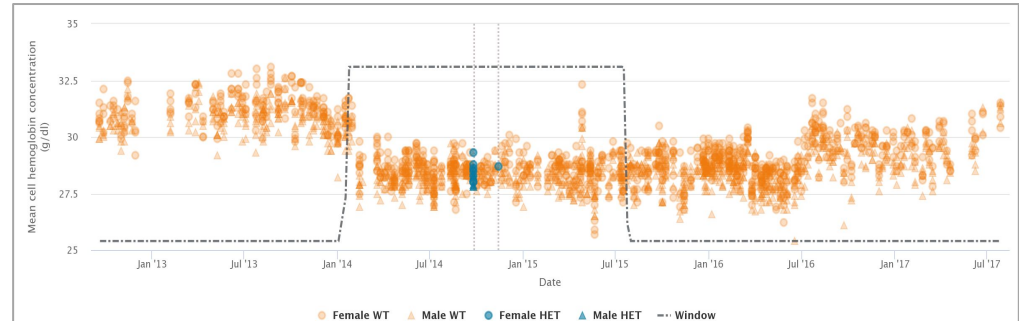
80M+ data points processed by the IMPC statistical pipeline in data release (DR) 16.0 April 2022

Statistical pipeline

- Scalable
- Reproducible
- Accurate
- Reliable

Soft windowing

- **Soft windowing application to improve analysis of high-throughput phenotyping data**
- **New methodology:**
 - Select control data collected proximally in time to that of mutants
 - Number of false positives generally reduced
 - Permits fixing the phenotype calls so that they are stable over time



Annotating phenotypes: ontologies

Equipment

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

Procedure

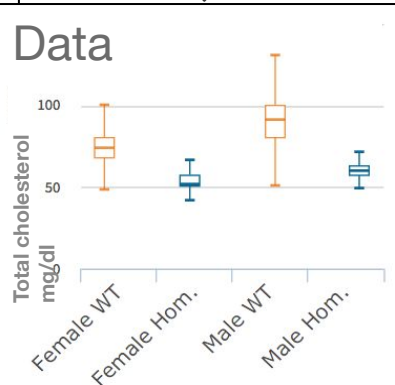
1. Fast mice overnight for approximately 16 hours by transferring mice to clean cages with no food or feces in hopper or bottom of cage. Ensure that they have access to drinking water at all times.
2. Prepare an equipment record sheet, syringe and sticks for glucose measurement and glucose solution.
3. Weigh the mouse.
4. Calculate and record the volume of 20% glucose solution required (2g of glucose/kg body mass) for IP injection as follows: volume of IP glucose injection (µl) = 10 x body weight (g).
5. Optional application of topical anesthetic cream:

 - a. Apply a small amount of topical anesthetic cream to the tail of the mouse, spreading over the tail evenly
 - b. Gently massage it in for ~15 seconds to enhance the effect of the anesthetic cream ensuring that the proposed incision site is fully covered
 - c. Allow an appropriate length of time for the local anesthetic to take effect.

6. Optional: Restrain the mouse in the restraining device with the tail exposed.
7. Score the tip of the tail using a hot or sterilized scalpel blade.
8. Discard first small drop of blood. A small drop of blood (~5µl) is placed on the test strip of the blood glucose meter. This is the baseline glucose level (t = 0) and is recorded on the equipment record sheet.
9. Remove the mouse from the restraining device.
10. Inject the mouse intraperitoneally with the appropriate amount of glucose solution, as previously determined (panel 8) and note the time-point of injection.
11. The blood glucose levels are measured at 15, 30, 60 and 120 minutes (t = 15, t = 30, t = 60 and t = 120) after glucose injection, by placing a small drop of blood on a new test strip and recording the measurements. Start the timing again by removing the test from the first incision, massage the tail (3 blood flows to indicate). Results are recorded on the record sheet.
12. Ensure that further blood from the incision is minimal by briefly applying pressure to the incision after each measurement. At the end of the experiment add food to the cage and make sure that a plentiful supply of water is available to the animals.
13. At the end of the experimental session, place the mouse in a clean cage with water and food available ad libitum.
14. Monitor the animals carefully to observe any abnormal behaviors.




Parameters

	Version	Type	Req. Upload	Req. Analysis	Annotated
Sodium [IMPC_CBC_001_001]	1.3	simpleParameter			✓
Potassium [IMPC_CBC_002_001]	1.3	simpleParameter			✓
Chloride [IMPC_CBC_003_001]	1.4	simpleParameter			✓
Urea (Blood Urea Nitrogen - BUN) [IMPC_CBC_004_001]	1.5	simpleParameter	✓		✓
Creatinine (enzymatic method preferred) [IMPC_CBC_005_001]	1.4	simpleParameter	✓		✓



Statistical analysis

Ontologies

	Option	Increment	Ontology Term	Ontology ID
INCREASED			increased circulating cholesterol level	MP:0005178 
DECREASED			decreased circulating cholesterol level	MP:0005179 
ABNORMAL			abnormal circulating cholesterol level	MP:0000180 

IMPC statistical pipeline: useful resources



OpenStats: Publicly available from Bioconductor:

<http://www.bioconductor.org/packages/release/bioc/html/OpenStats.html>

SoftWindowing: Publicly available from CRAN

<https://cran.r-project.org/web/packages/SmoothWin/index.html>



SoftWindowing: Publication

<https://pubmed.ncbi.nlm.nih.gov/31591642/>

OpenStats: Publication (under the review)

<https://www.biorxiv.org/content/10.1101/2020.05.13.091157v1>



IMPC Statistical pipeline: Repository

<https://git.io/JUBXF>

IMPC Annotation pipeline: Repository

<https://git.io/JUBXN>

IMPC wiki page:

https://github.com/mpj2/impc_stats_pipeline/wiki



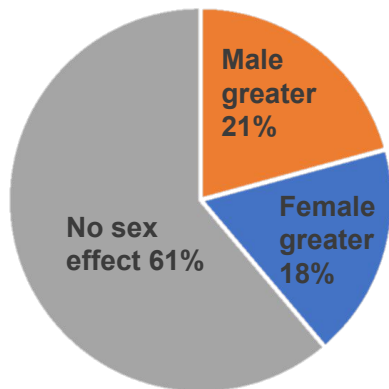
IMPC Statistical pipeline: Docker container

https://hub.docker.com/r/hamedhm/impc_stats_pipeline

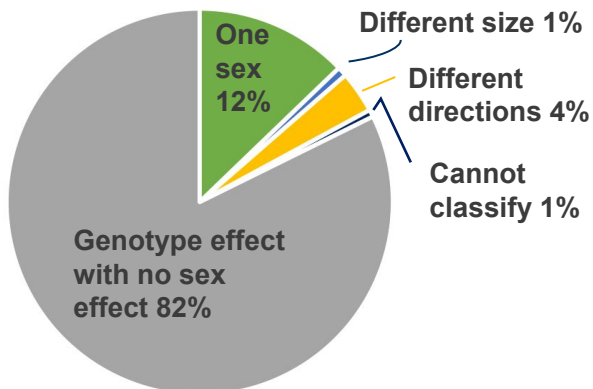
Key findings and achievements

Sexual dimorphism

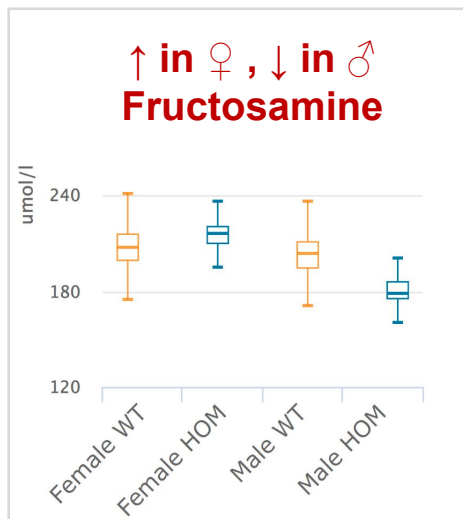
Wildtype mice



KO lines



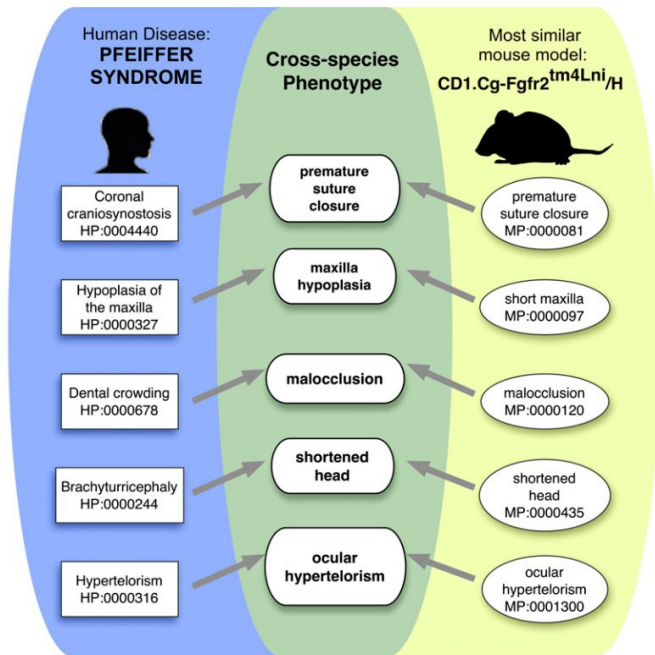
Galnt18^{tm1b(KOMP)Wtsi}



- Screening of 14,250 wildtype mice and 2,186 knockout strains
- IMPC demonstrates the **effect of sex** on many phenotypes
- Support for including **both females and males in biomedical research**

Translating to other species: ontologies

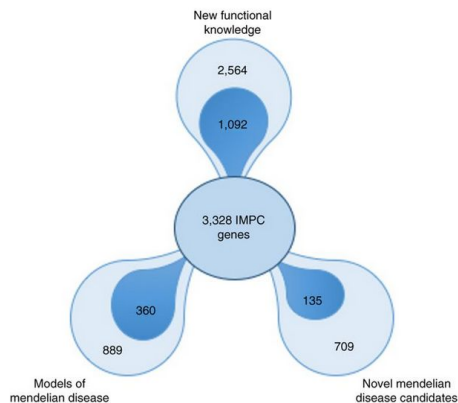
- Ontologies
- Cross-species phenotype comparison by semantic similarity



Disease models by semantic phenotype overlap

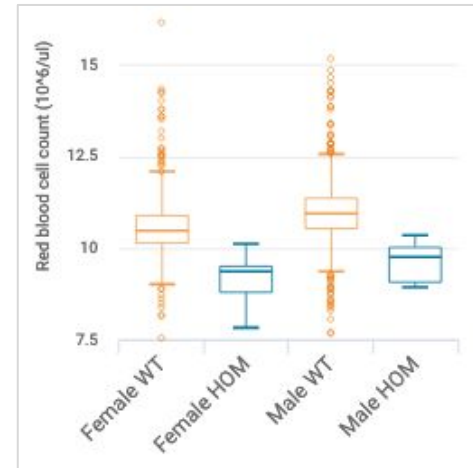
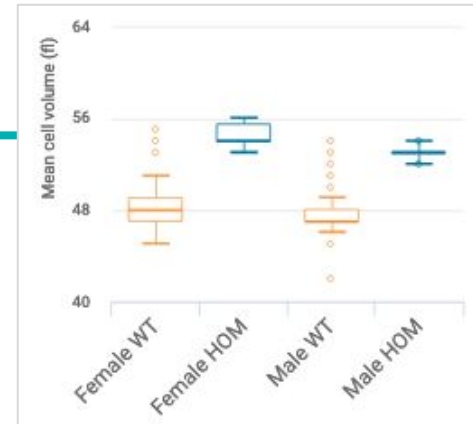
- 3,328 genes surveyed
- Demonstrated semantic phenotype overlap of mouse strains and human diseases >> propose 360 disease models

Source	Disease term	Mouse symbol	Common phenotype
OMIM	Gordon Holmes Syndrome	Rnf216	infertility
OMIM	Bernard-Soulier Syndrome	Gp9	increased mean platelet volume decreased platelet cell number
OMIM	Bone Mineral Density Quantitative Trait Locus 18	Pls3	decreased bone mineral density
OMIM	Dyskeratosis Congenita, Autosomal Recessive, 2	Nhp2	abnormal platelet morphology
OMIM	Epilepsy, Progressive Myoclonic 6	Gosr2	abnormal gait
ORPHANET	Bardet-Biedl Syndrome	Bbs5	abnormal retina morphology
ORPHANET	Cerebellar Ataxia - Hypogonadism	Rnf216	Decreased fertility



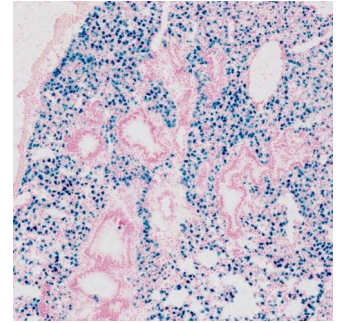
Example: Fam53b^{tm1b(EUCOMM)Hmgu}

- IMPC phenotyping reveals:
 - Increased mean red blood cell volume
 - Decreased red blood cell number
- Candidate gene for macrocytic hyperchromic anemias
- Provide data and models



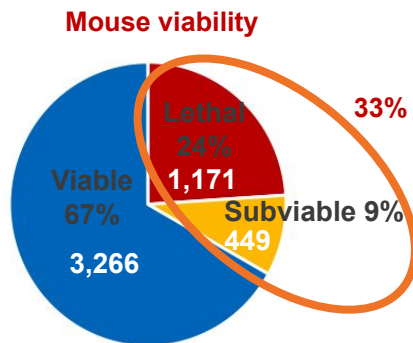
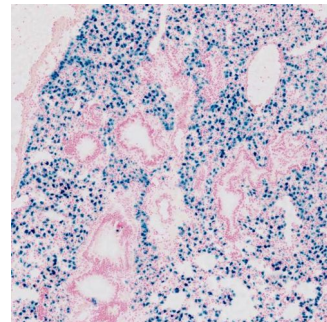
Understanding embryo development

- Embryo and viability data
- Wealth of **image data**: X-ray, LacZ, histopathology, embryo 3D
 - Embryonic landing page
 - Interactive Embryo Viewer
- Genes associated to **developmental phenotypes**
- Human disease genes enriched for **essential genes**



Understanding embryo development

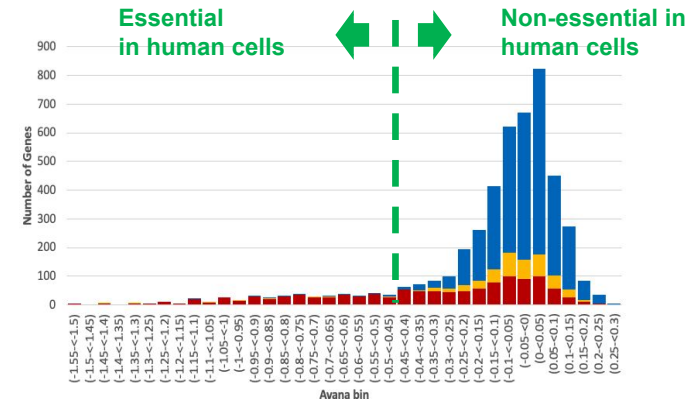
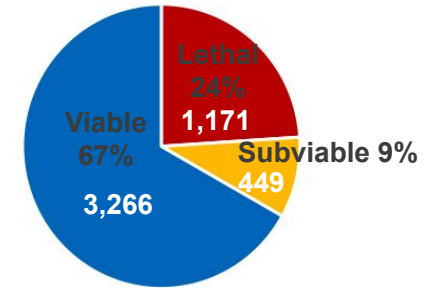
- Embryo and viability data
- Wealth of *image data*: X-ray, LacZ, histopathology, embryo 3D
- Genes associated to *developmental phenotypes*
- Human disease genes enriched for *essential genes*



Human and mouse essentiality screens

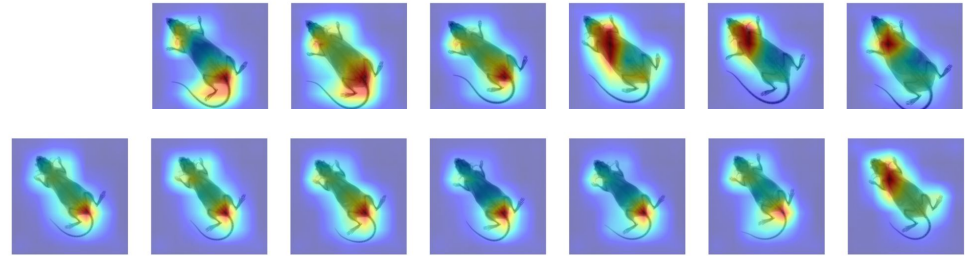
- Combined IMPC *mouse* viability and phenotypic data with *human* cells essentiality screens
- Genes into 5 mutually exclusive categories (from more to less essential): *FUSIL binning*
- Developmental Lethal (essential in mice but not in human cells) are *enriched for genes associated with developmental disease*
- Demonstrate as a tool for disease gene discovery: identified *163 genes as novel candidates* for developmental disease

Mouse viability



Machine learning to automatically annotate images

- Deep learning applied to X-ray images of WT females and males
- Applied to automatically annotate gender in KO strains
- Detect abnormalities in sex-related phenotypes

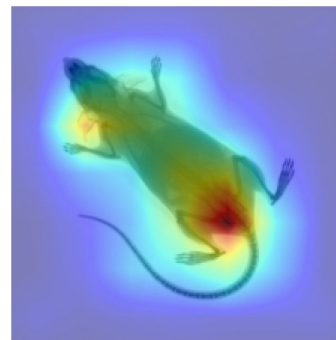
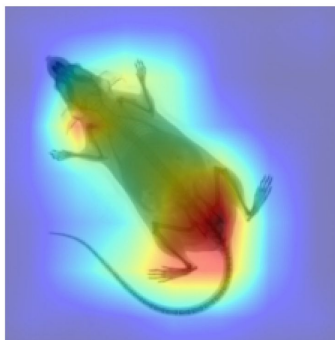
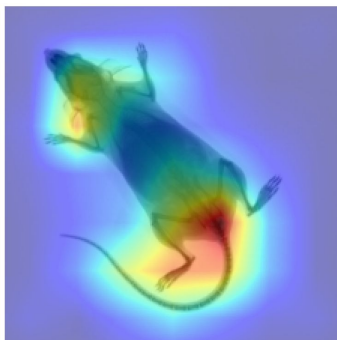


Top: Class activation maps for randomly selected WT females (first three) and males (last three).
Bottom: Class activation maps for null *Duoxa2*^{tm1b(KOMP)Wtsl} individuals, all labelled as males; using machine learning techniques, six were classified as females and one as male.

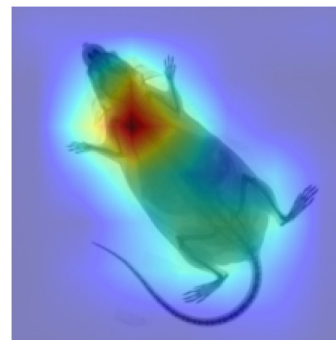
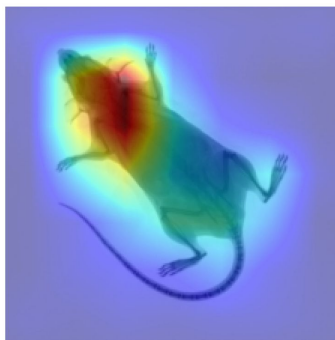
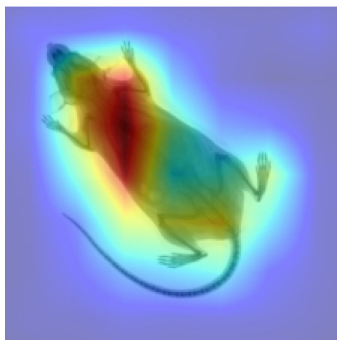
Investigating phenotypes with CNNs

- Built convolutional neural network to classify X-ray images based on sex
- Has 97% accuracy
- Works in an intuitive manner
 - Occlusion sensitivity
 - Class activation maps
- Preliminary results indicate useful to investigate phenotype changes

Class activation maps



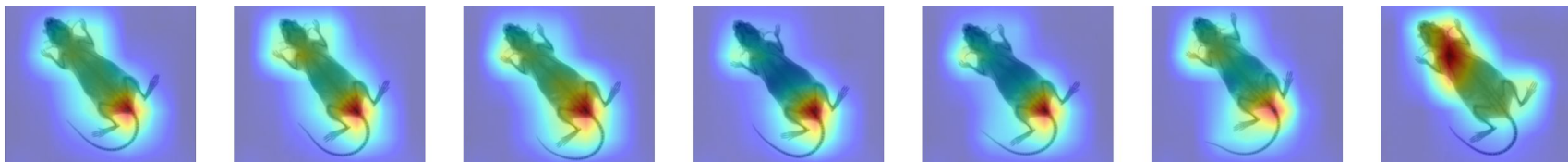
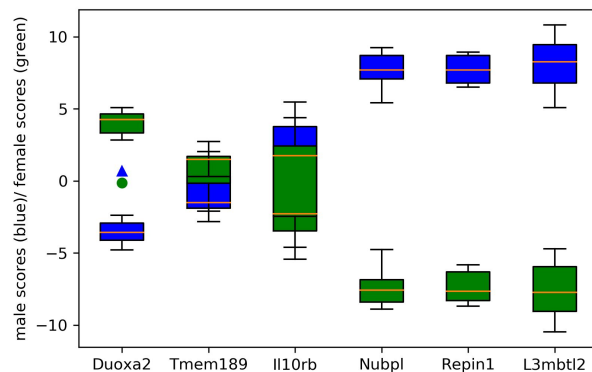
Typical CAMs from
3 female WT



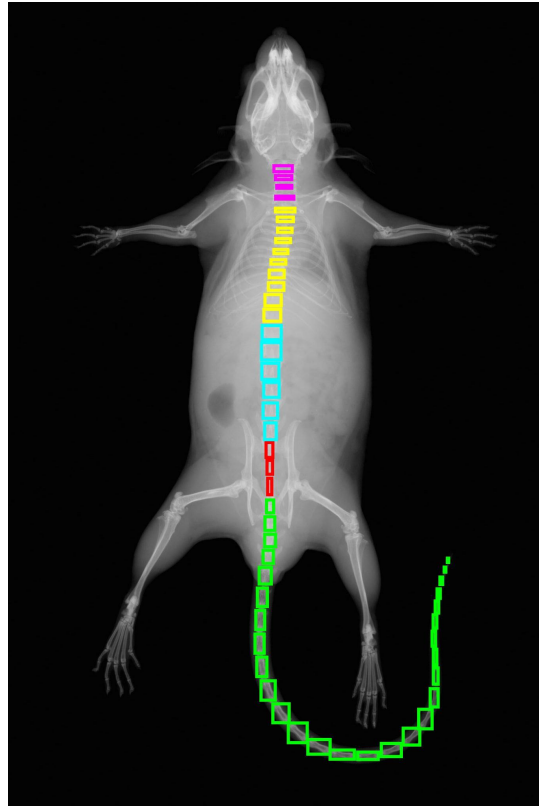
Typical CAMs from
3 male WT

Have we identified phenotypic changes?

- **Duoxa2** was particularly interesting as pattern inverse of expectation
- Class activation maps for the 7 males of **Duoxa2** are shown below
 - 6 classified as females
 - Note these have similar CAMs to WT females



Results of counting vertebrae using object detection



- Cervical
- Thoracic
- Lumbar
- Sacral
- Caudal

Other IMPC key achievements

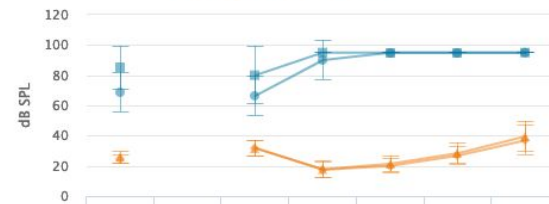
- Extensive *pleiotropy*
- *Systematic analyses* published on
 - Metabolism, hearing, eye... more underway
 - Novel genes associated to hearing loss
- Over 1,500 *CRISPR*-modified lines
- Insight into mammalian genome editing

Novel, severe hearing loss

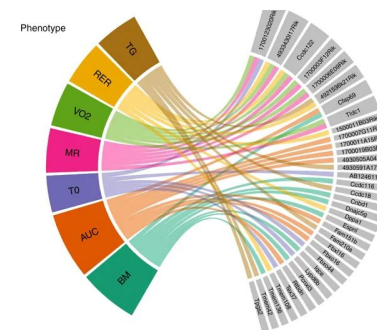
Zfp719^{tm1b(EUCOMM)Wtsi}

Evoked ABR Threshold (6, 12, 18, 24, 30 kHz) ≡

Auditory Brain Stem Response



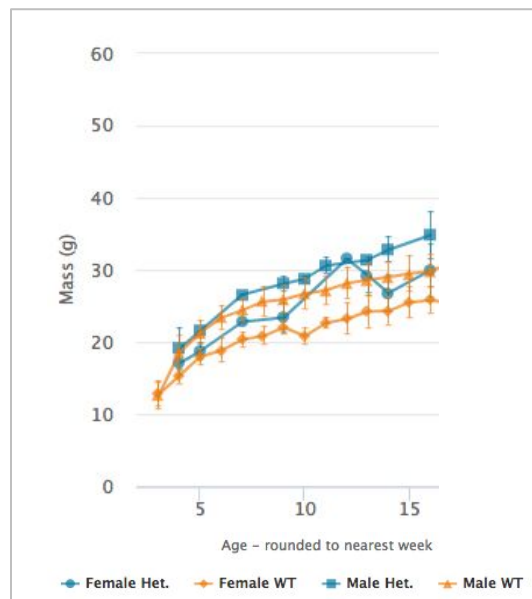
Moore et al. *Communications Biology* 2018; *Scientific Reports* 2019



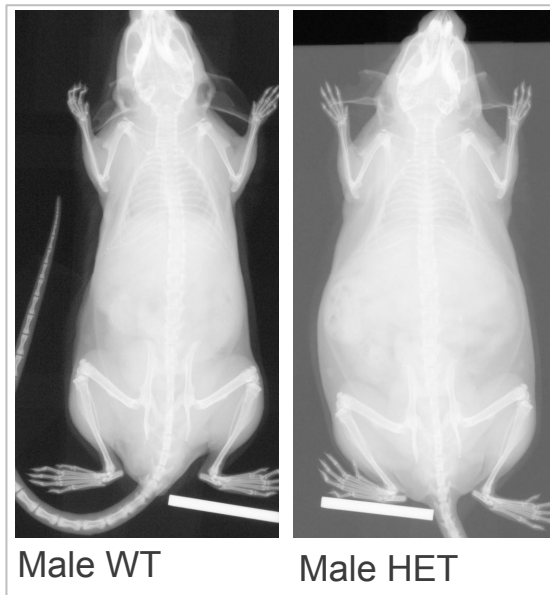
Rozman et al. *Nature Communications* 2018

Rsbn1^{tm1b(EUCOMM)Wtsi}

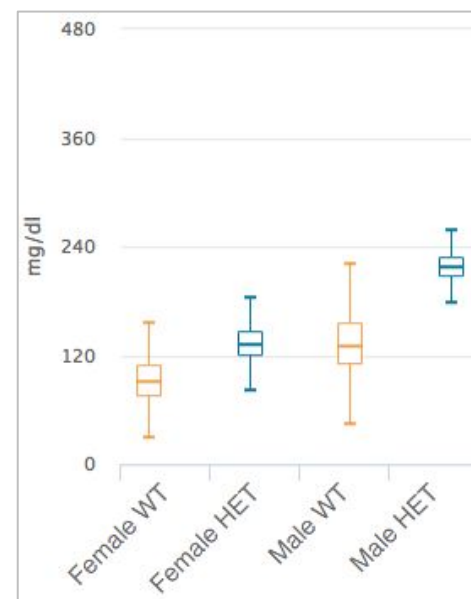
Bodyweight



X-ray



Triglycerides



Impact and integration

IMPC - UK Biobank

- Prospective study of ~500,000 participants recruited in the UK
- Healthy individuals, 40-70 years of age and 88.5% British White
- Genetic, physical and health-related data

STUDY

- At scale, for humans and mice
- Mappings for 20 haematological and 16 blood biochemical measurements
- Whole-exome genetic data for all participants and loss of function associations for mice
- Data continuously growing

Age comparison

Relative comparison between human and mouse age

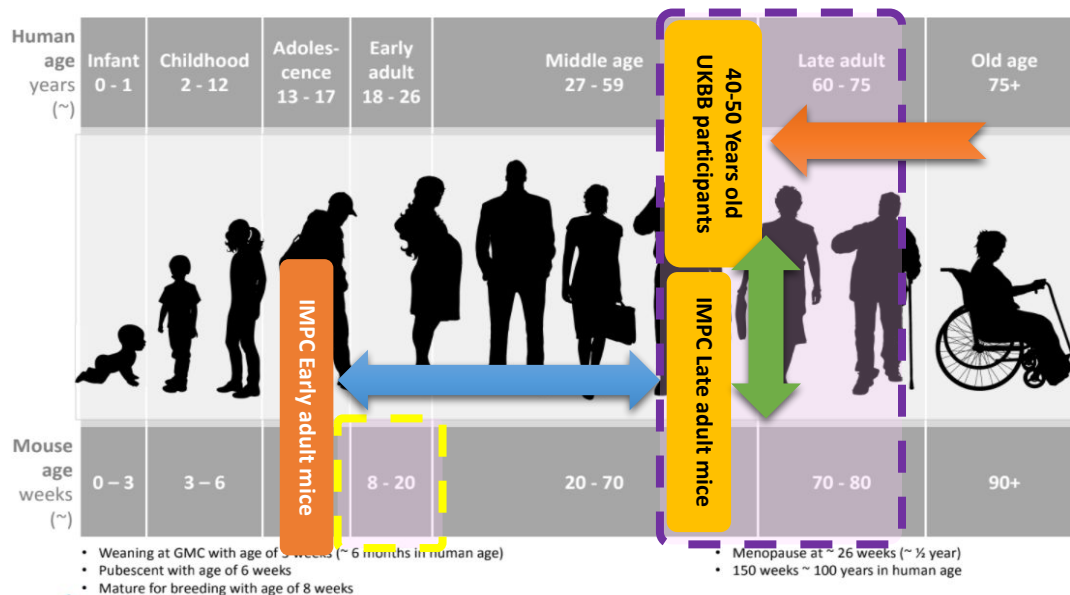


Figure is adapted from: Dutta, S., and Sengupta, P. (2016) Men and mice: Relating their ages. *Life sciences* 152, 244-248
 Mitchell, S. J., Scheibye-Knudsen, M., Longo, D. L., and de Cabo, R. (2015) Animal models of aging research: implications for human aging and age-related diseases. *Annu Rev Anim Biosci* 3, 283-303
 Lister, R., and Mukamel, E. A. (2015) Turning over DNA methylation in the mind. *Front Neurosci* 9, 252
 Flurkey, Currier, and Harrison. (2007) The mouse in biomedical research, American College of Laboratory Animal Medicine series. Elsevier, AP: Amsterdam/Boston.

Modelling sexual dimorphism



- Capturing variation across centres in wildtype mice
- Summarize results from all centres into a single score

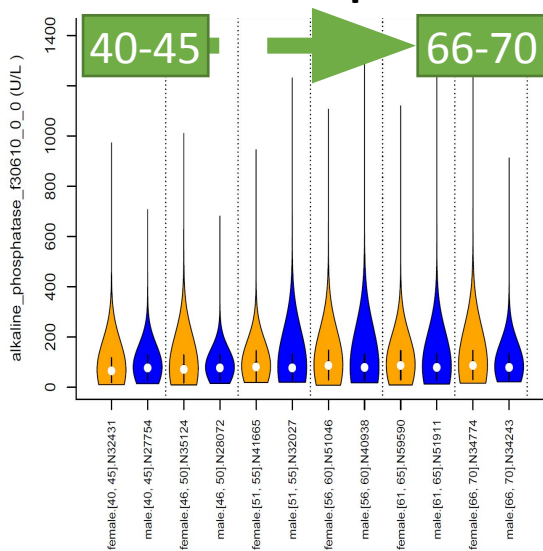


- Binning the human data into 6 age ranges
- Each covers 5 years:
 - [40-45], [46-50], [51-55],
 - [56-60], [61-65], [66-70]



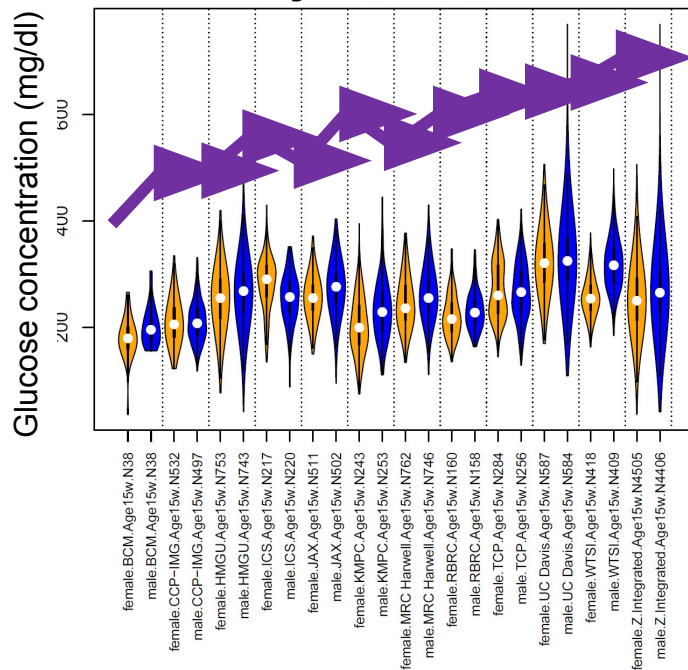
 females
 males

Alkaline Phosphatase

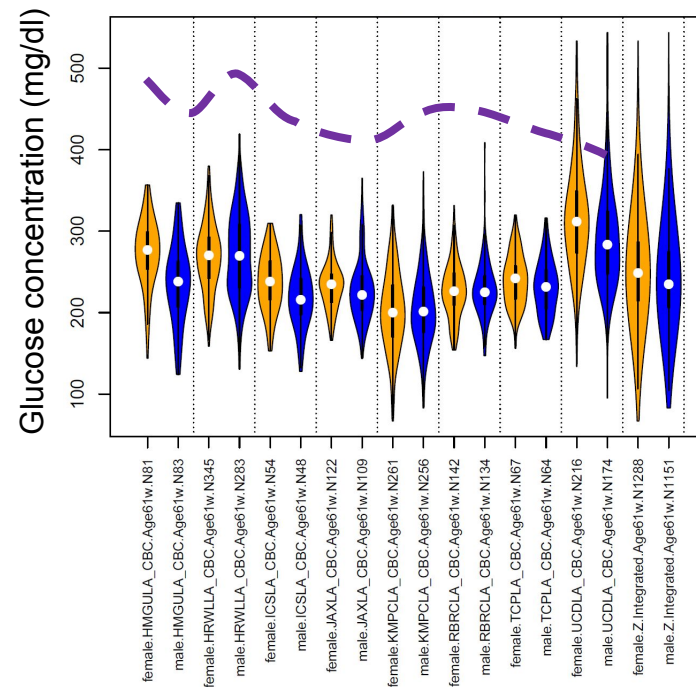


Modelling sexual dimorphism

Early adult mice



Late adult mice



Preliminary results



1. Sexual dimorphism widespread in mice

Sexual dimorphism present in Early Adult and Late Adult

- Males or females higher
- 1 parameter **males** higher, 1 **females** higher

Additionally, sexual dimorphism found in one life stage:

- **Early Adult:**
 - 5 **males** higher, 6 **females** higher



2. Sexual dimorphism present in all parameters

Mostly stable across age groups

- 10 parameters **males** higher, 7 parameter **females** higher

Changes in trend at ~50 years of age

- 6 parameters out of 17

Secondary phenotyping of IMPC mice

- **198 genes** with NeuroAnatomical Phenotypes (NAPs) linked to **brain morphogenesis**: 1,566 IMPC genes were studied for 118 neuroanatomical parameters
- **38 genes** with novel association to **hearing loss**: 1,211 IMPC genes studied for the Auditory Brainstem Response (ABR)
- **140 genes** linked to **immunology phenotypes**: 530 IMPC genes studied in 3i steady state and challenged screens; sexual dimorphism detected
- Review outcomes of **in-depth bone phenotyping** programmes: <http://bonebase.org/>, <http://boneandcartilage.com/>

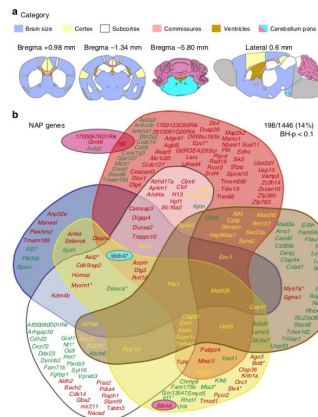
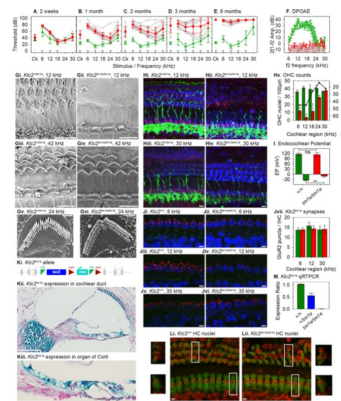
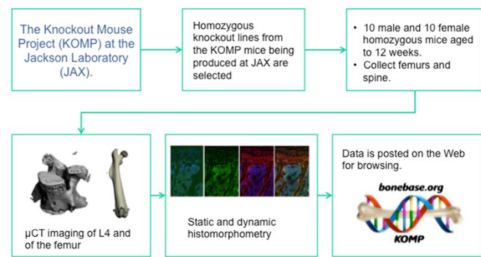

 Collins et al. *Nature Communications* 2019

 Ingham et al. *PLoS Biology* 2019


FIGURE 1 | The data collection pipeline for the BoneBase.org phenotyping project. This is one of two specialized high-throughput phenotyping pipelines that is conducting auxiliary, bone-specific phenotyping of mice generated by the IMPC. The BoneBase.org logo is used with permission from the database owners.

 Maynard and Ackert-Bicknell *Frontiers in Endocrinology* 2019

Collaborations and IMPC data integration

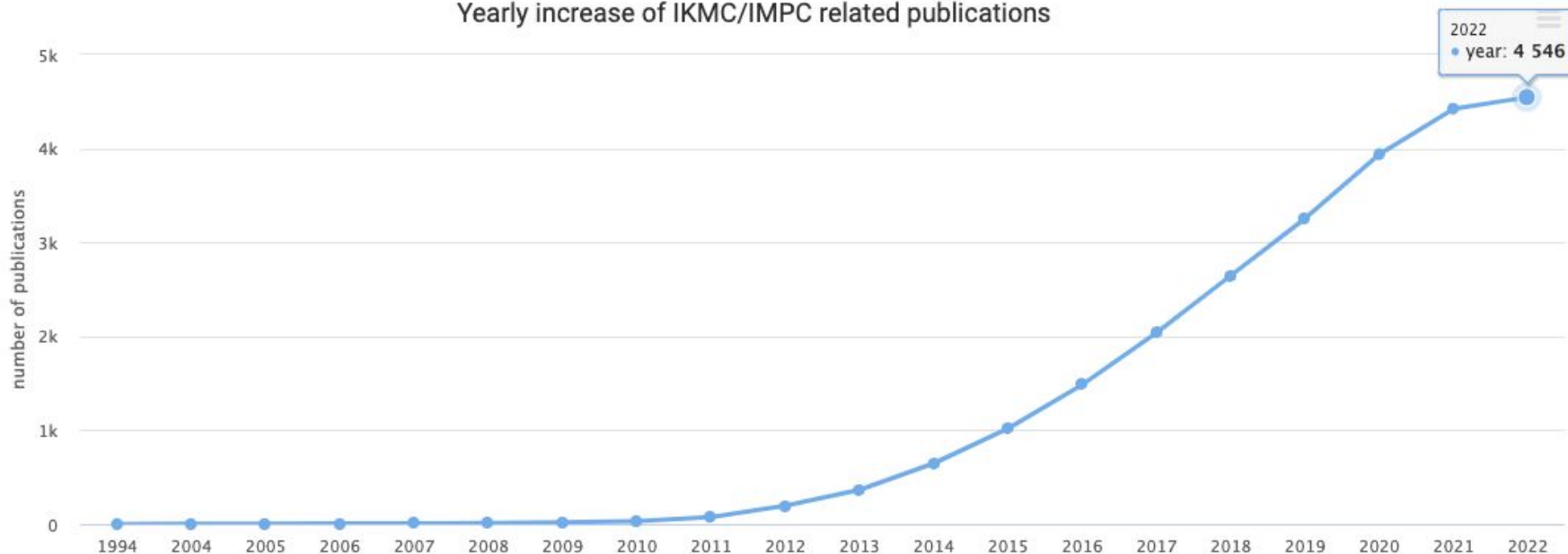
- Support diagnosis of rare disease patients
- Use mouse phenotypes to screen *human variants (Exomiser)*
- Data integrated by other resources

Exomiser (Damian Smedley (QMUL), Peter Robinson (Charite), others)
<http://www.sanger.ac.uk/resources/databases/exomiser>



Publications using IMPC resources

Yearly increase of IKMC/IMPC related publications



Concluding remarks

- High-throughput IMPC phenotyping is revealing ***thousands of new gene-phenotype*** associations
- Connecting clinical phenotype data with mouse phenotypes
- Coordinate mouse production across ***rare disease*** consortia
- New ***tracking system*** to accommodate all types of planned induced mutations (GenTaR)
- Finding best ways to provide ***feedback / communicate*** design, production, phenotyping results
- All ***data freely available*** at mousephenotype.org

Acknowledgments



Robert Wilson



Zina Perova



Hamed Haseli Mashhadi


 Mauricio Martínez
 Jiménez


Violeta Muñoz Fuentes



Tushar Mandloi



Federico López Gómez

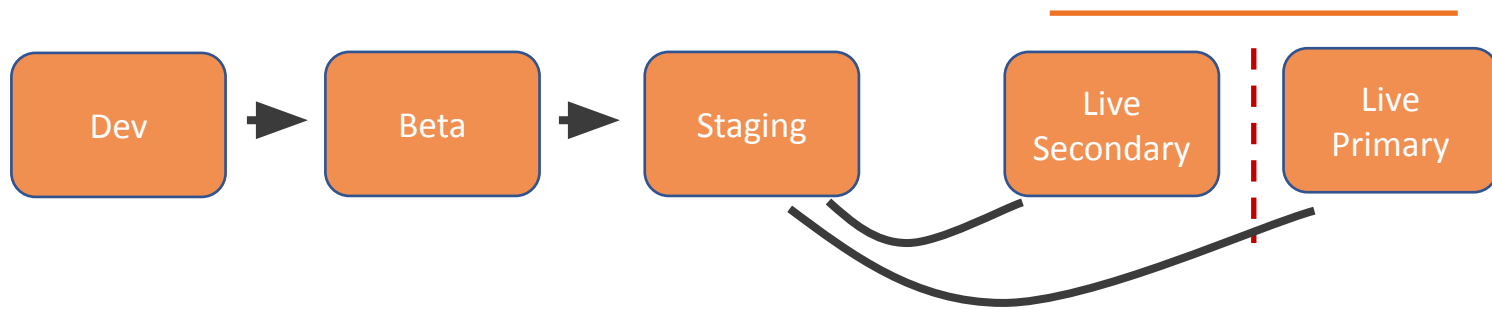


Osman Güneş

National Institutes of Health [NIH U54 - HG006370]
 European Molecular Biology Laboratory Core Funding

BREAK

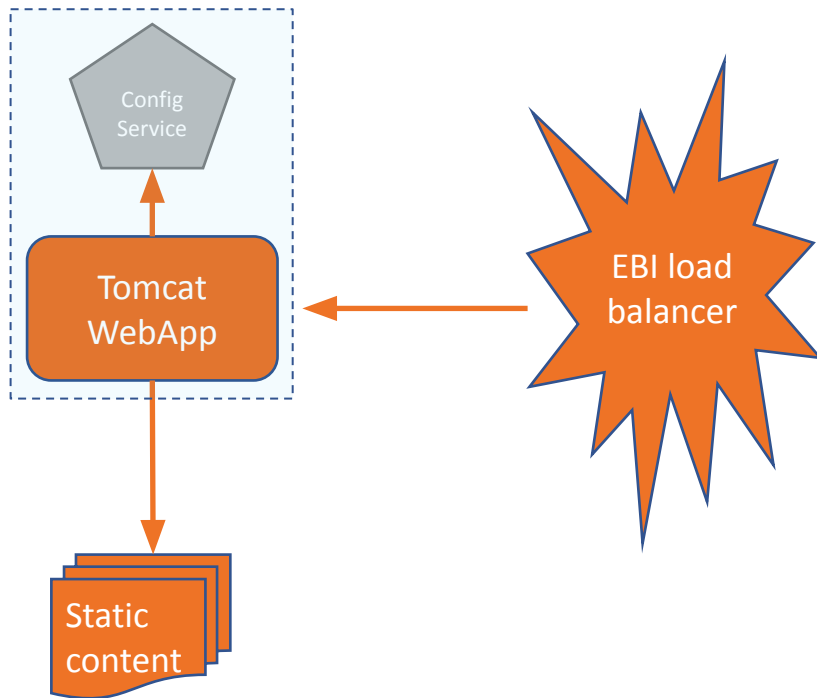
PhenotypeArchive Service



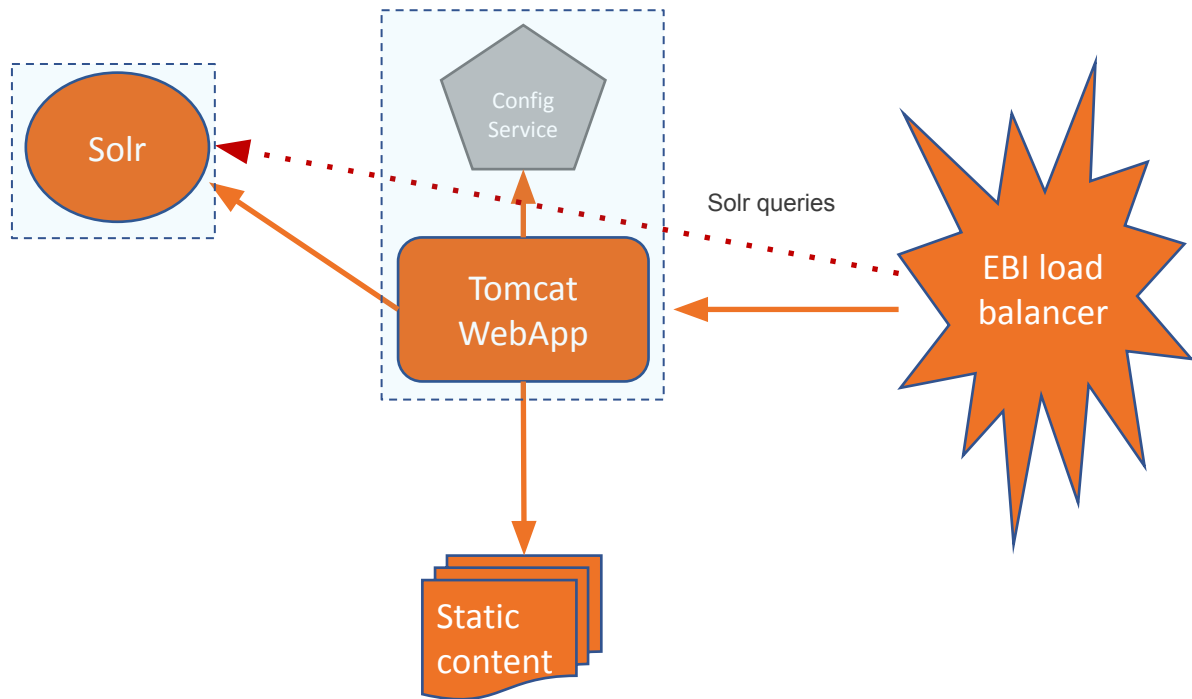
	Dev	Beta	Staging	Live Secondary	Live Primary	Load Balancer
Web	Public/Internal	Public/Internal	Internal	Internal	Internal	Public
Solr API	Public/Internal	Public/Internal	Internal	Internal	Internal	Public

Production

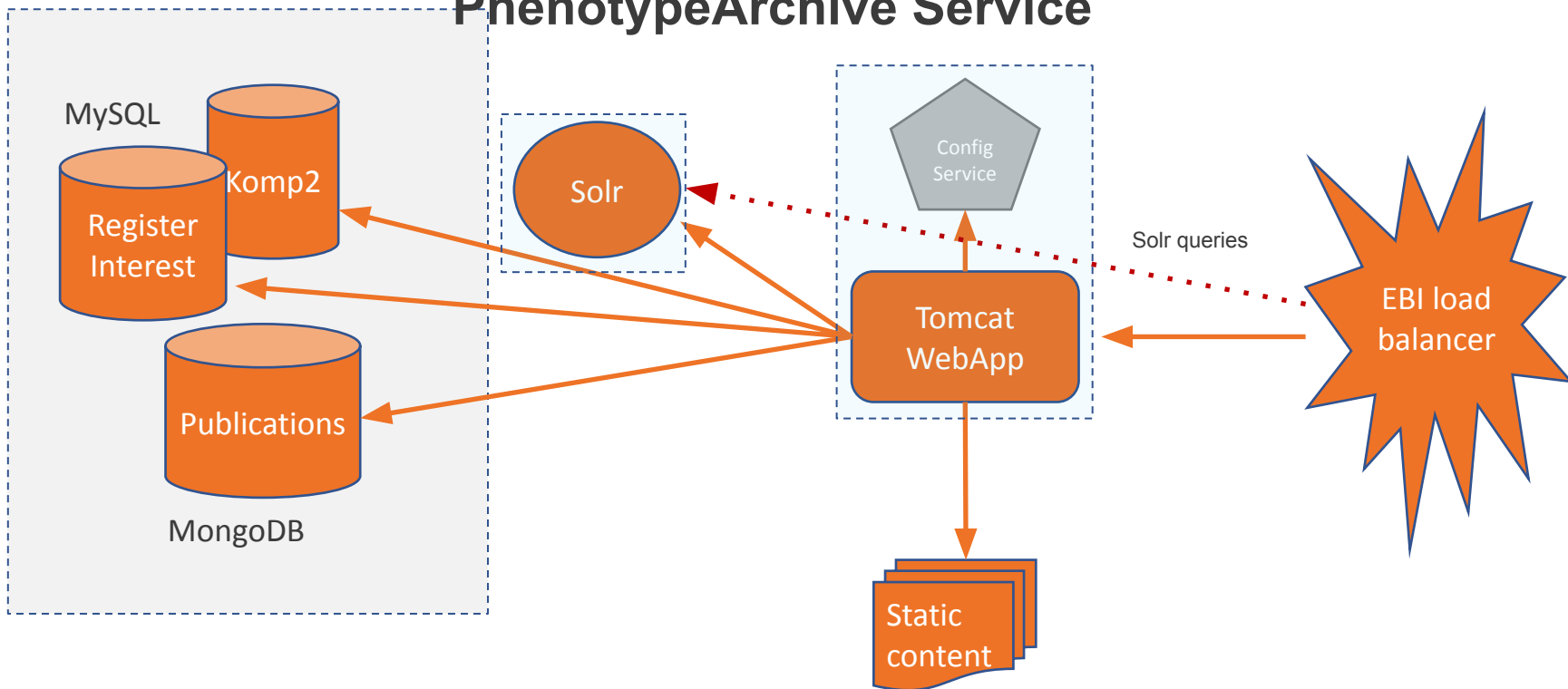
PhenotypeArchive Service



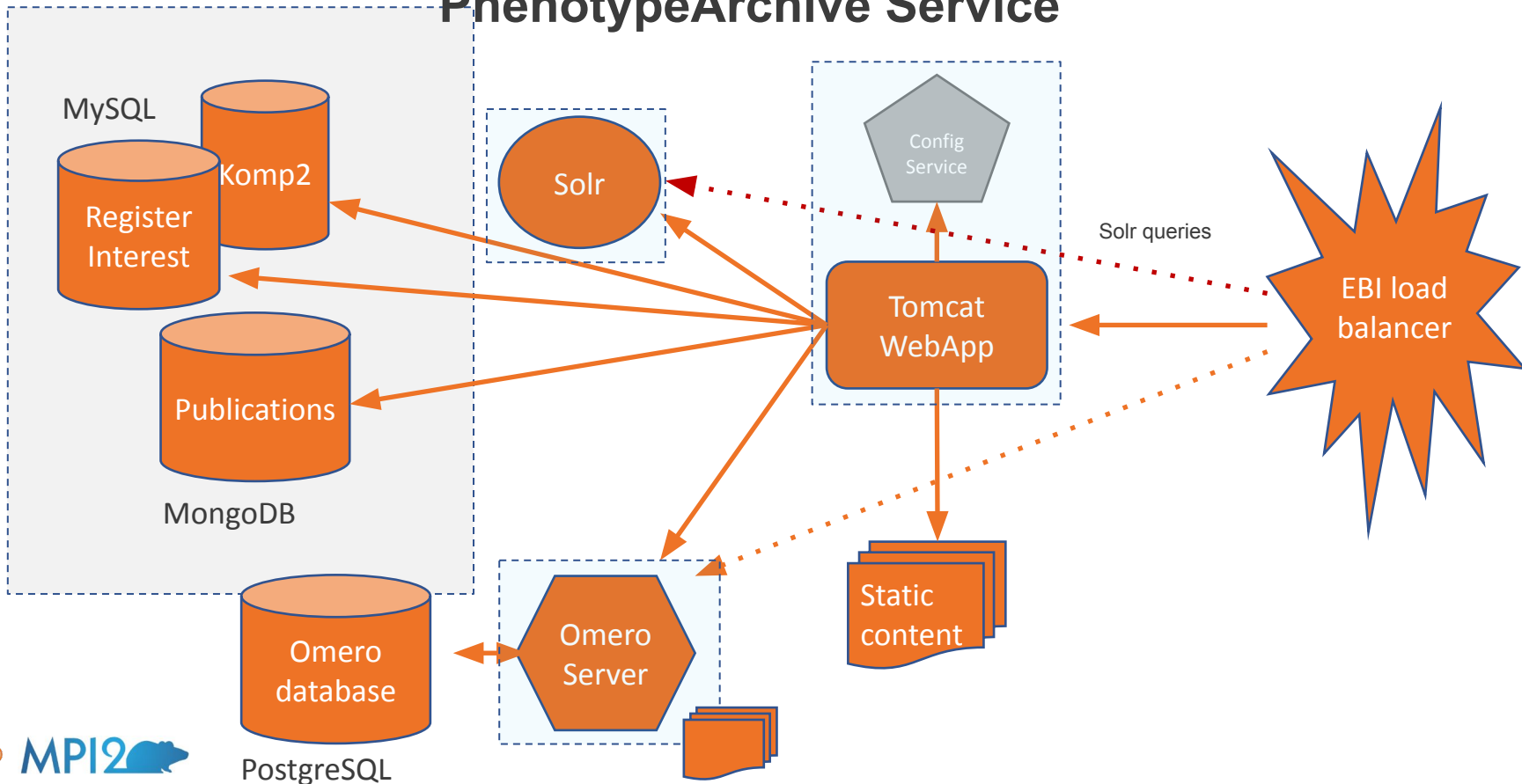
PhenotypeArchive Service



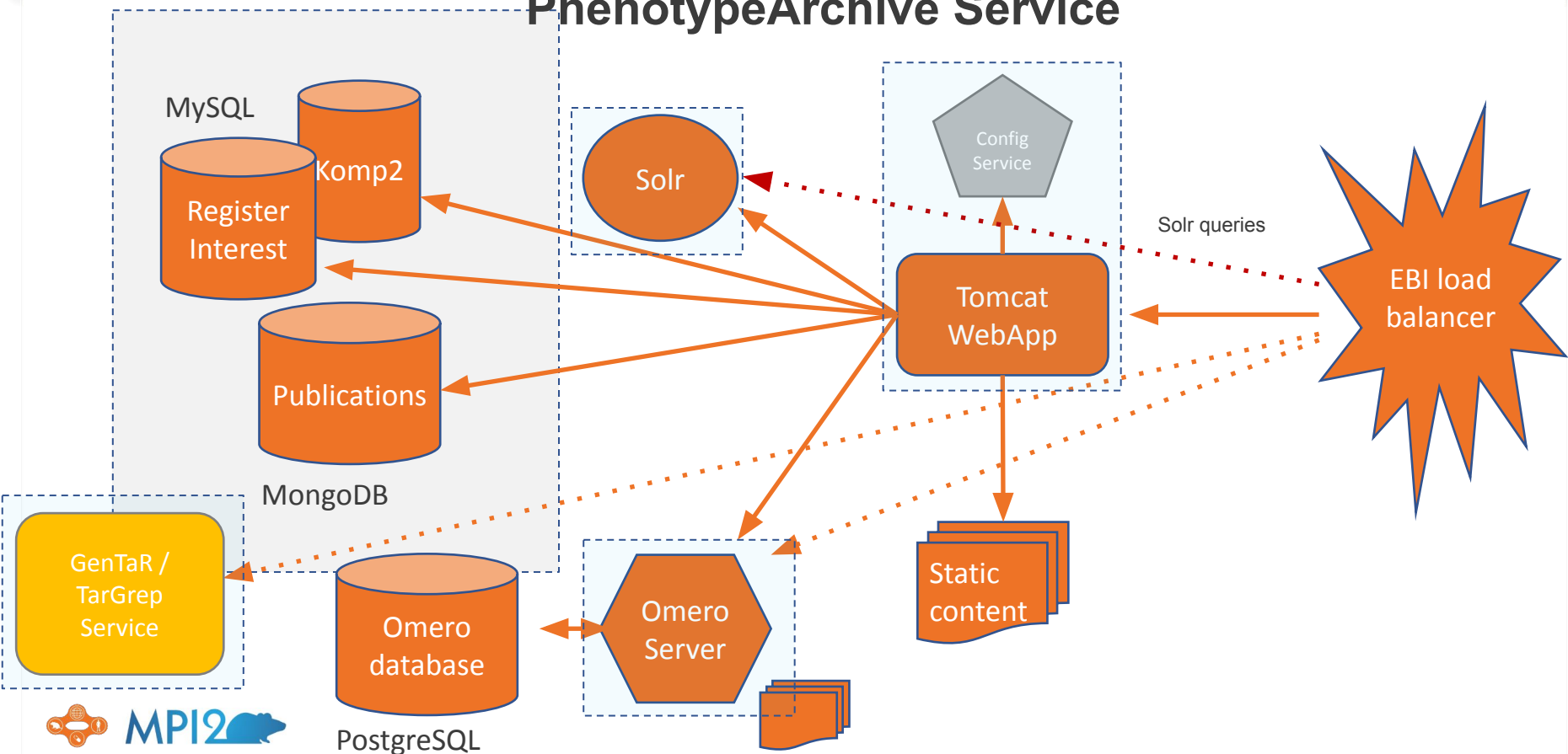
PhenotypeArchive Service

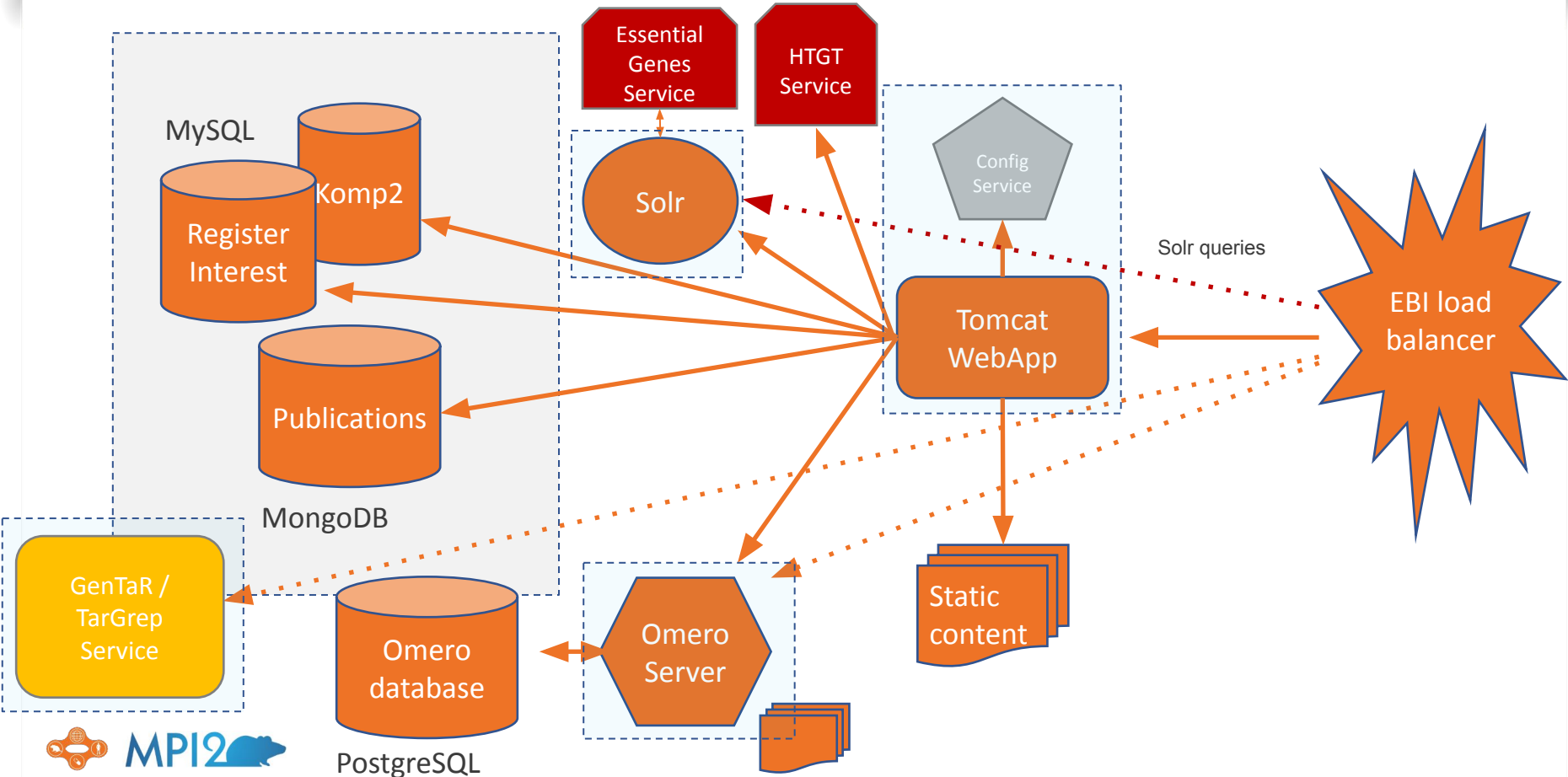


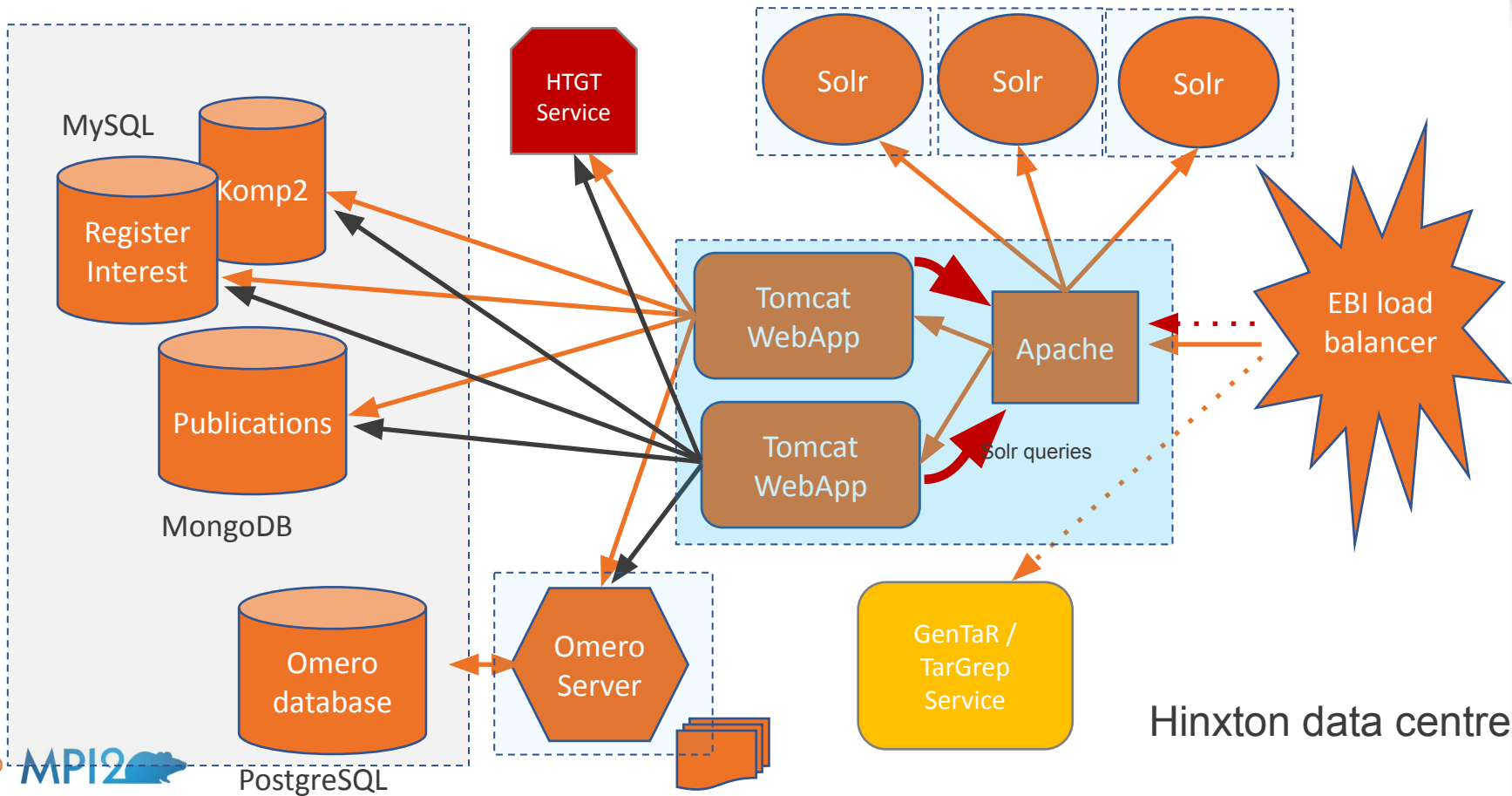
PhenotypeArchive Service



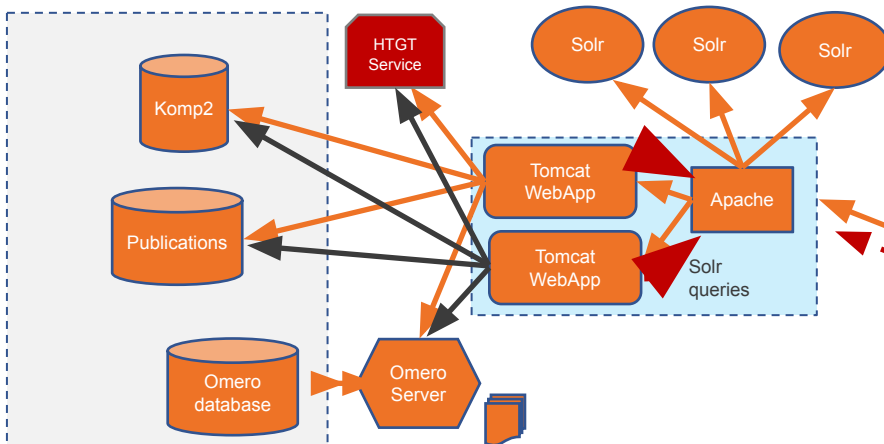
PhenotypeArchive Service



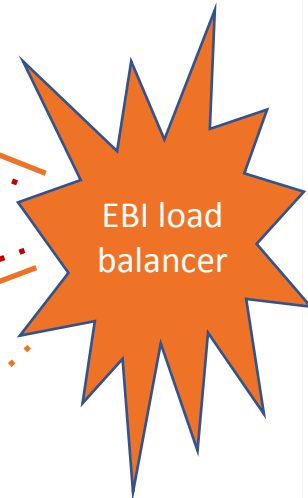
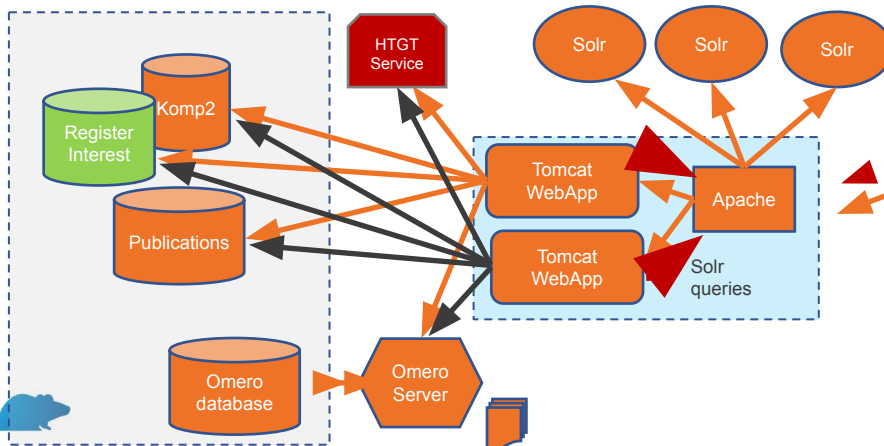




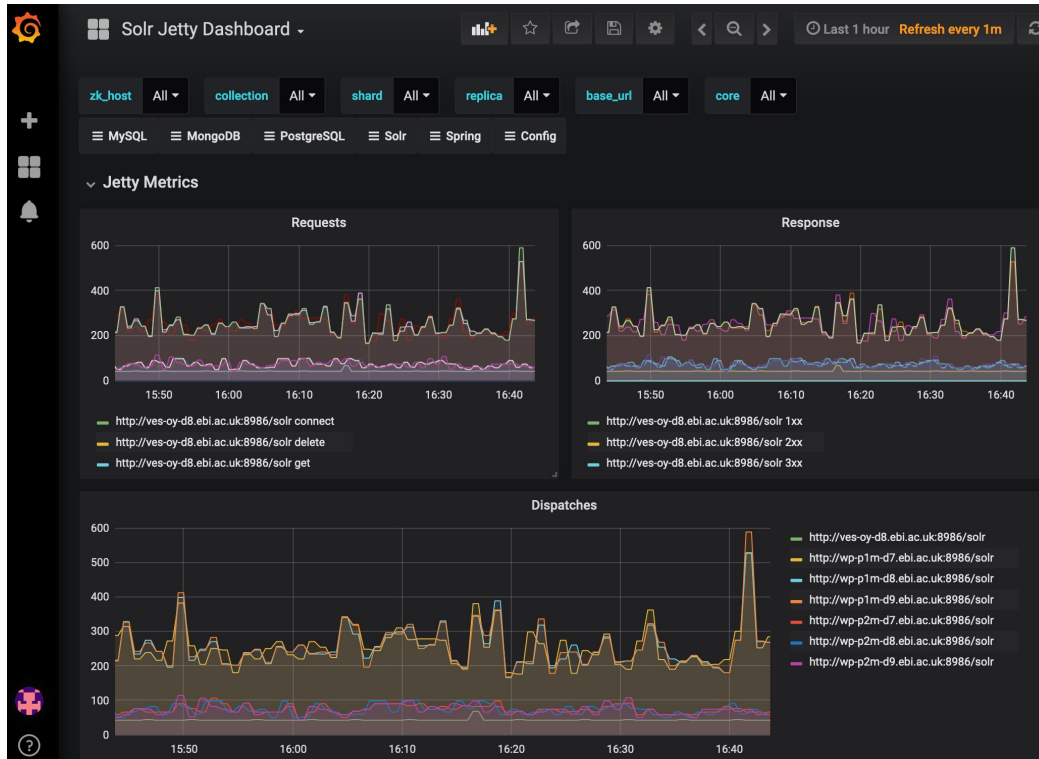
Harlow



Hinxton



Service Monitoring



Docker-compose application
 Self-managed VM to provide docker

- Prometheus
- Grafana
- Spring Exporter
- MySQL Exporter
- PostgreSQL Exporter

Service Management

Jenkins



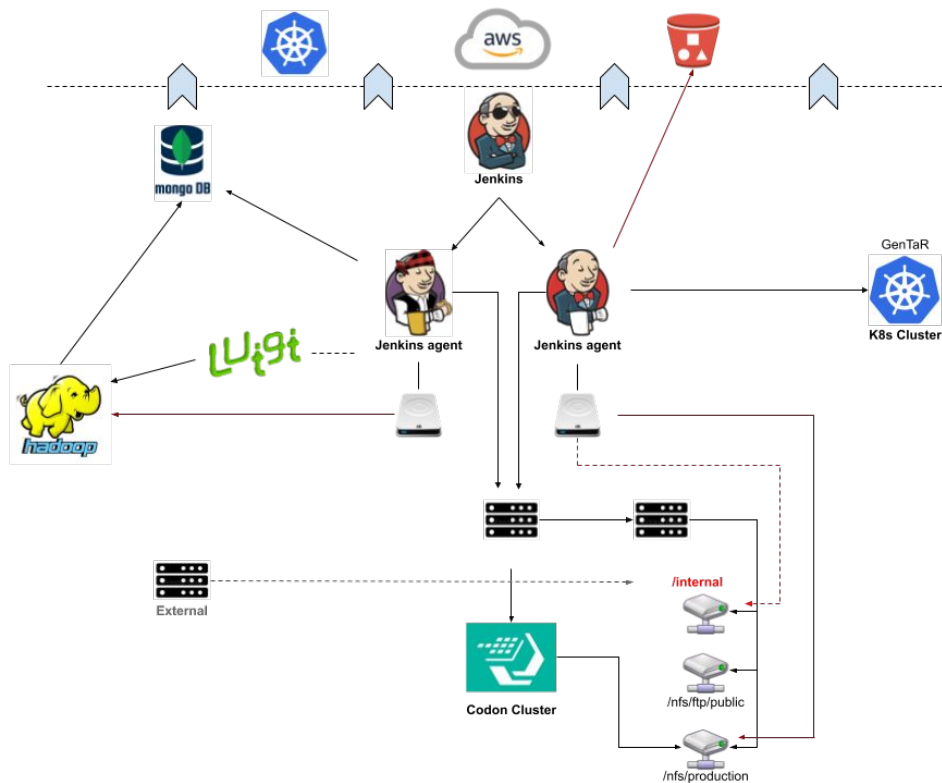
- Jobs to deploy to each environment
 - KOMP2 DB
 - Solr Cores
 - WAR files for Web App
- Imaging - Dev and Production OMERO
- FTP reports and data release archive
- UpTime jobs to check and restart services
- Scheduled Database backups

GitLab

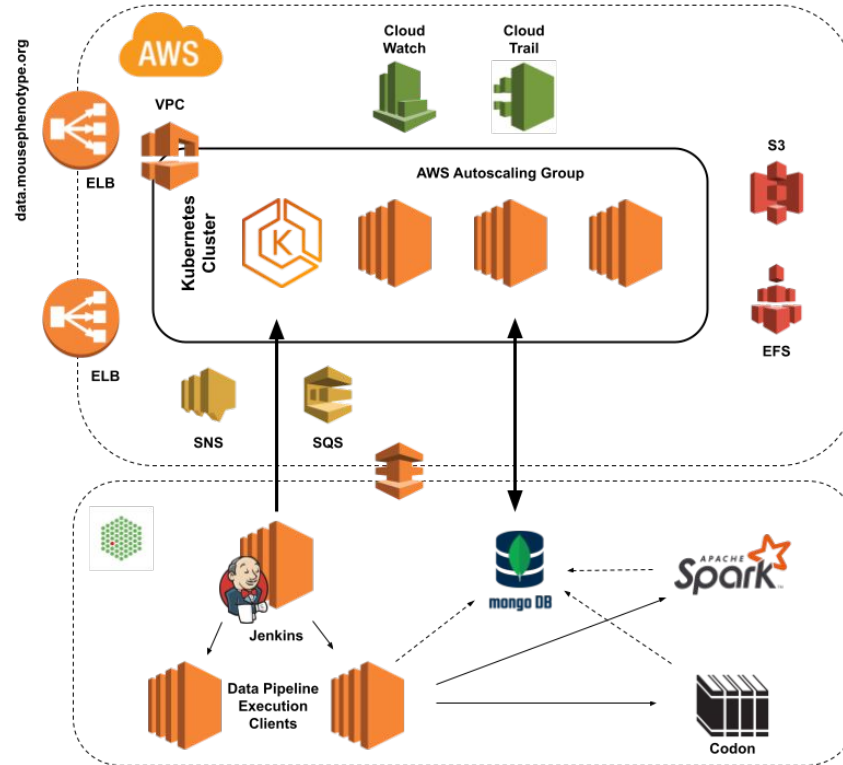


- Used for Kubernetes deployment of GenTaR
- Automated turnover of pods

Future architecture



Public cloud deployment



Acknowledgments



Robert Wilson



Zina Perova



Hamed Haseli Mashhadi


 Mauricio Martínez
 Jiménez


Violeta Muñoz Fuentes



Tushar Mandloi



Federico López Gómez



Osman Güneş

National Institutes of Health [NIH U54 - HG006370]
 European Molecular Biology Laboratory Core Funding