GETTING CLOSER TO THE ENCYLOPEDIA OF MAMMALIAN GENE FUNCTION

Ismael Hinojo Muñoz

EHEA Degree in Biotechnology, Universitat Autònoma de Barcelona

INTRODUCTION

Human Genome Project took a giant step towards the understanding of human genetics. However, much is yet to be discovered regarding genetic information. Issues such as gene functions and epigenetic modifications are great unknowns to the scientific community. For this reason, two big consortia have arisen to take the lead on the discovery of gene function on the mammalian genome. The International Knockout Mouse Consortium (IKMC) and the International Mouse Phenotyping Consortium (IMPC) aim, through mice gene knockout and phenotyping, to create an encyclopedia of mammalian gene function.

THE IKMC – A KNOCKOUT CONSORTIUM

The IKMC is a consortium of different members of different countries whose aim is to mutate all protein-encoding genes in the mouse using a combination of gene trapping and gene targeting in C57BL/6 mouse embryonic stem (ES) cells. Generated data is available at www.knockoutmouse.org.

MEMBERS

KOMP — The Knockout Mouse Project (KOMP) is a trans-NIH initiative. Using its two production centers, KOMP aims to generate up to 8,500 knockout ES cells and, to a limited extent, knockout mice.

EUCOMM — The European Conditional Mouse Mutagenesis Program (EUCOMM) is founder member of the IKMC and it contributes with up to 12.000 conditional mutations across the mouse genome.

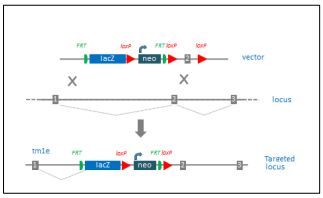
EUCOMM TOOLS — EUCOMM: Tools for Functional Annotation of the Mouse Genome (EUCOMMTOOLS) is the successor project of EUCOMM.

NorCOMM — The North American Conditional Mouse Mutagenesis project (NorCOMM) is a large-scale research initiative from Canada that and aims to generate 2,000 targeted ESC clones.

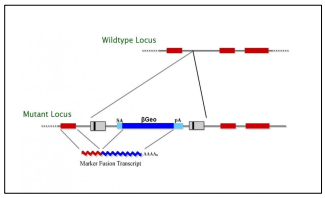
TIGM — The Texas A&M Institute for Genomic Medicine (TIGM) is a research institute whose resources include the world's largest gene trap library that contains mutated ES cell clones representing more than 10,000 genes.

STRATEGIES

Gene targeting

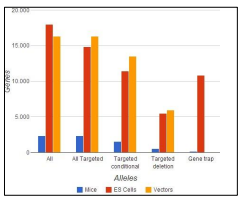


Gene trapping



The current strategies used for knockout are gene targeting and gene trapping, being the first one the most used. Each member of the IKMC has its own constructs with variations between them.

CURRENT PROGRESS



Targeted alleles

Total Genes	KOMP CSD	KOMP Regeneron	EUCOMM/ EUCOMMTools	NorCOMM	mirKO
Vectors available	6596	4731	9909	839	
ES cells available	5518	4086	7862	569	221
Mutant mice available	890	465	904	4	

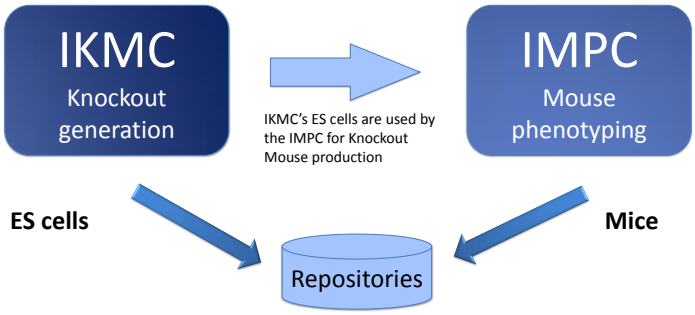
Gene trap alleles

Total Genes	TIGM	EUCOMM	NorCOMM
ES cells available	9346	4352	4587
Mutant mice available	159	19	

The main focus has been producing ES cells, although some mice and vectors have been produced to a lesser extent. Gene targeting has been preferred over gene trap due to its precision.

REPOSITORIES

Repositories were created to host all the production of ES cells, mice and the rest of derivative products from the IKMC. All the materials are publically available and can be purchased by individual researchers or organisations. KOMP mice, vectors and ES cells are stored at the KOMP Repository. EUCOMM and EUCOMMTOOLS mice are stored at the European Mouse Mutant Archive (EMMA) while ES cells and vectors are stored at the European Mouse Mutant Cell Repository (EuMMCR). The Canadian Mouse Mutant Repository (CMMR) is the main repository of NorCOMM and the TIGM stores its own products.



THE IMPC – A PHENOTYPING CONSORTIUM

The IMPC builds on the efforts of IKMC to produce knockout mice and carry out high-throughput phenotyping of each line in order to determine the function of every gene in the mouse genome. Data is available at www.mousephenotype.org.

GOALS

- Establish a world-wide consortium of centres with capacity for large-scale phenotyping
- Test each mutant mouse line through a broad based primary phenotyping pipeline aiming to ascribe biological function to each gene
- Provide a centralized data centre and a portal for free, unrestricted access to data
- Generate an encyclopedia of mammalian gene function

The project started in 2011 with the expected duration of 10 years and divided in two phases.

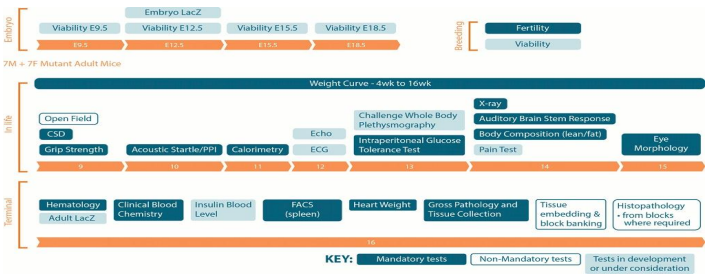
Phase 1 2011-2016

The main deliverables are selection of IMPC core phenotypic tests and platforms, release of the database, population of the database with information on at least 4,000 KO mouse lines.

Phase 2 2016-2021

The objectives are to assess performance of Phase I, adjust pipelines and operations as necessary, and scale operations to complete the genome i.e. standardized production and phenotyping of 16,000 targeted genes.

THE PIPELINE



The IMPC core pipeline is the one agreed upon by the different research institutions that form part of the Consortium. The pipeline involves the analysis of cohorts of mice with a broad spectrum of tests. The pipeline is still in development and probably will be until 2016 and its protocols are still being developed. The data can be accessed through the IMPReSS webpage (http://www.mousephenotype.org/impress).

SUMMARY

The IKMC is well on its way to accomplish the objective of knocking out every protein-encoding gene on the mouse genome. The focus is now on the IMPC, who has the task of phenotyping all the knockouts produced by the IKMC. Although still in development, it is starting to achieve a well-defined pipeline and when that's finished, it will accelerate the analysis rate greatly towards the end goal, the generation of an encyclopedia of mammalian gene function. Furthermore, the public availability of all the materials generated by the IKMC and IMPC with no added costs opens a new way for SCIENCE.