**National Advisory Council for Human Genome Research (NACHGR)**

**May 17, 2021**

**Concept Clearance for FOAs**

**Knockout Mouse Phenotyping Program (KOMP2) renewal**

## Purpose

This concept proposes a Limited Competition Request for Applications (RFA) in order to renew the Knockout Mouse Phenotyping Program (KOMP2) for a final phase. The proposed KOMP2 renewal will have trans-NIH funding support from 18 NIH Institutes, Centers and Offices (ICOs), with joint administration by the National Human Genome Research Institute (NHGRI) and the Office of Research Infrastructure Programs (ORIP). KOMP2 is part of the global International Mouse Phenotyping Consortium (IMPC) effort, which aims to knock out and characterize all protein-coding genes in the mouse genome. The purpose of this limited competition final phase will be to support production and phenotyping of 1,000-1,200 mouse knockouts (KOs), representing a substantive contribution to completion of all high priority mouse:human orthologs.

## Background

Knocking out the activity of a gene provides valuable clues about what that gene normally does; however creating mice with gene KOs is a time-consuming and difficult process, and often is done by individual research labs with varied approaches. The original knockout mouse program (KOMP) was launched in 2006 as a trans-NIH effort stemming from discussions at a 2003 CSHL Banbury Center meeting titled "Mouse Genome-wide Targeted Mutagenesis". The KOMP project created 8,500 embryonic stem cell lines in C57BL/6N. The program was expanded as the Knockout Mouse Phenotyping Program (KOMP2) in 2011 as a joint trans-NIH and Common Fund program with the goal of generating and phenotyping mice lines. Incorporation of CRISPR technology greatly enhanced this phase of the project, and over 5,500 mouse strains have been produced by the NIH-funded groups. Since the founding of the program, KOMP2 scientists established and adhered to uniform characterization – or phenotyping – protocols, data collection, and reporting standards, and made all data available via a data portal. KOMP2 awardees, in conjunction with the IMPC, published over 120 peer-reviewed publications, including papers systematically describing genes with previously unknown function in hearing, embryonic development, and metabolism. Overall, this effort helps scientists explain the genetic basis of many different types of diseases in mice that also occur in humans, including rare diseases that have been under-studied as well as some common chronic diseases.

Last spring, participating ICOs approved support for a one-year funded extension of the KOMP2 program at 40% of the FY20 levels, allowing the program to recover from COVID-related impacts. The Common Fund will no longer be able to contribute to the extension year or renewal, since it has already provided 10 years of support for this project, the maximum allowed.

KOMP2 serves as the flagship project of the IMPC. The IMPC has created a distributed infrastructure for the analysis of mouse gene function, allying all the major mouse genetics centers worldwide to leverage an extraordinary breadth of expertise in mouse genetics from genome editing and phenotyping to data analysis. KOMP2 contributes over half of all mice and data in IMPC. The IMPC consortium began work in 2011 with its first and most immediate goal to generate a null mutant and undertake broad-based phenotyping for every gene in the mouse genome. IMPC partners have implemented a standardized phenotyping pipeline that interrogates multiple biological domains. Data are uploaded into a common archive for analysis, annotation, and dissemination.

**Current Status for KOMP2/IMPC:**

1. 9,719 genotype confirmed knockouts
2. 7,455 lines phenotyped – 93% of lines have 1 or more phenotype
3. 98 million data points and 572k images in Data Release 13
4. 3,740 publications using KOMP/IMPC resources

**Noteworthy Achievements and Discoveries:**

1. Production
    a. Surpassing annual production targets
    b. KOMP2/Mutant Mouse Resource & Research Centers (MMRRC) have fulfilled more than 9,100 orders received from the research community
2. Phenotyping
    a. Ageing screen near completion – 600 mutant lines breeding, aging, or phenotyped
    b. Identified 1,718 embryo lethal and 620 subviable lines - ~35% lethal or sub-viable.
    c. 93% of KOs have phenotypes, even for poorly annotated/understudied genes, many are sexually dimorphic
    d. 74% of KOs have 2 or more phenotypes, pleiotropy is common
3. Data & Informatics
    a. Freely available; Delivering integration and access for data clinical informatics resources

**Proposed Scope and Objectives**

Although the KOMP/KOMP2 project has delivered on its goals and milestones during each of its project periods, it is not quite at the finish line. There are 16,847 genes that have clear 1-to-1 homology between human and mouse and are thus considered directly relevant to human health. 13,069 of these have mouse knockouts that can be considered "readily available" either from KOMP2/IMPC (9,719) or high-quality knockouts in other repositories (3,350). Approximately 3,000 of these remaining genes are considered high priority targets. The proposal is to complete the high priority group, with KOMP2 making ~1,200 KOs and IMPC partners making the additional strains.

Two different efforts will be supported through separate, limited competition RFAs:

1) **KOMP2 production and phenotyping centers** will be funded to generate mutant mouse lines using CRISPR/Cas9 technology, perform phenotyping assays, conduct quality control assessments, cryopreserve germplasm, and make mice and data readily available to the research community (via the MMRRC and the IMPC web portal, respectively).

2) The **KOMP2 data coordinating center** will provide informatics support to NIH-funded projects that are performing high-throughput, broad-based phenotyping and to coordinate with international efforts so as to integrate all data into a common database. The Data Coordination Center will perform the curation, analysis, visualization, and dissemination of the phenotype data from the knockout lines. Curation will require integration with other data sources. Analysis will require further development and validation of statistical methods.

The ultimate goals of these efforts are to enhance the ability of the biomedical research community to identify new disease models, to better understand phenotypic patterns, and to gain a more comprehensive understanding of the underlying function of each gene.

**Relationship to Ongoing Activities**
As noted above, KOMP2/IMPC provides KO strains and data that are broadly used by the biomedical research community. Recent publications show preferential use of IMPC alleles in publications reporting on new knockouts. The proposed renewal of the KOMP2 project complements NHGRI projects that focus on cellular or molecular phenotypes (including the MorPhiC project that received Concept Approval at February 2021 Council). Over the years, KOMP2 has formed collaborations with several NIH partners, including the Gabriella Miller Kids First Pediatric Research Program, the Illuminating the Druggable Genome program, the Centers for Mendelian Genomics, the Undiagnosed Disease Network, and NICHD's KOMP-focused R01 projects.

**Mechanism of Support**
We plan to issue two RFAs, each using the UM1 cooperative agreement mechanism. We anticipate supporting 2-3 production and phenotyping centers and one data coordinating center.

**Limited Competition**
Only grantees funded under the existing KOMP2 Knockout Mouse Production and Phenotyping Project and Knockout Mouse Phenotyping Project Database will be eligible to apply. These groups have been involved with this project for 15 years and have developed the infrastructure needed to efficiently and effectively carry out this work. Given the reduced level of funding available for this final stage of this project, it is not feasible to onboard and ramp-up new groups to reach the production levels needed.

**Funds Anticipated**
NHGRI will commit $1.5M total costs/year over 5 years beginning in FY2022 for a total of $7.5M total costs. This project will be organized as a trans-NIH program with co-funding from 17 other ICOs (NCCIH, NCI, NEI, NHLBI, NIA, NIAAA, NIAID, NIAMS, NICHD, NIDA, NIDCD, NIDCR, NIDDK, NIEHS, NINDS, ORIP, and ORWH). The total expected NIH investment will b $8.5M total costs/year for 5 years. Approximately $1M total costs/year will support the data coordinating center and the remainder will be allocated for the Production and Phenotyping Centers. This represents a ~60% reduction from FY20 funding levels (which included Common Fund support).