Title

A resource of targeted mutant mouse lines for 5,061 genes

Journal

Nature Genetics, 53(4)

ISSN

1061-4036

Authors

Birling, Marie-Christine
Yoshiki, Atsushi
Adams, David J
et al.

Publication Date

2021-04-01

DOI

10.1038/s41588-021-00825-y

Peer reviewed

HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2021 October 01.

Published in final edited form as:

Nat Genet. 2021 April ; 53(4): 416–419. doi:10.1038/s41588-021-00825-y.

A resource of targeted mutant mouse lines for 5,061 genes

A full list of authors and affiliations appears at the end of the article.

Summary

The International Mouse Phenotyping Consortium reports the generation of new mouse mutant strains for over 5,000 genes, including 2,850 novel null, 2,987 novel conditional- ready, and 4,433 novel reporter alleles.

Despite thirty years of mouse targeted mutagenesis, *in vivo* function of the majority of genes in the mouse genome are still unknown. This reflects the observation that a small number of genes have been the object of intensive study including the development of multiple mouse models, while a significant proportion of the coding genome remains

[40] Correspondence: l.teboul@har.mrc.ac.uk. [41] Correspondence: steve.murrav@jax.org.
[*]A list of authors and their affiliations appears at the end of the paper.

Complete list of additional contributors available in the Supplemental Material

Author Contributions

M.-C. B., A.Y., S.A., J.B., A.Bu., W.B., F.C., S.C., G.F.C., F.J.D., B.D., M.D.F., A.G., M.G., A.G-S., L O G., P.K. R.E.K., R.K., H.L., Y.J.L., I.L., A-M.M., T.F.M., V.M.F., S.N., L.M.J.N., G.T.O., G.P., R.R-S., B.R., E.J.R., L.A.S., J.S., J.R.S., C.S., M.T., H.W-J., M.W-D., M.V.W., B.J.W., J.A.W., L.T., and S.A.M generated data, developed data tools and databases, and/or carried out data and statistical analyses; A.Y., D.J.A., A.L.B, A.B., S.D.M.B., H-J.G.C., MED., L.R.D., X.G., J.D.H., Y.H., M.H.A., S-T.J., M.J.J., Z.L., K.C.K.L., C M., A-M.M., T.F.M., L.M.J.N., R.S., J.K.S., W.C.S., T.S., K.P.S., G.P.T-V., C- K.L.W., S.W., W.W., Y.X., L.T., and S.A.M. directed research at their respective institutions; M.-C. B., L.T., and S.A.M wrote the paper.

Competing Interests: The authors declare no competing interests.

* IMPC Consortium Authors

Juan J Gallegos[4], Jennie R Green[4], Ritu Bohat[4], Katie Zimmel[4], Monica Pereira[12], Suzanne MacMaster[12], Sandra Tondat[12], Linda Wei[12], Tracy Carroll[12], Jorge Cabezas[12], Qing Fan-Lan[12], Elsa Jacob[12], Amie Creighton[12], Patricia Castellonos-Penton[12], Ozge Danisment[12], Shannon Clarke[12], Joanna Joeng[12], Deborah Kelly[12], Christine To[12], Rebekah van Bruggen[12], Valerie Gailus-Durner[15], Helmut Fuchs[15], Susan Marschall[15], Stefanie Dunst[15], Markus Romberger[15], Bernhard Rey[15], Sabine Fessele[42], Philipp Gormanns[42], Roland Friedel[6], Cornelia Kaloff[6], Andreas Hörlein[6], Sandy Teichmann[6], Adriane Tasdemir[6], Heidi Krause[6], Dorota German[6], Anne Könitzer[6], Sarah Weber[6], Joachim Beig[6], Matthew McKay[10], Richard Bedigian[10], Stephanie Dion[10], Peter Kutny[10], Jennifer Kelmenson[10], Emily Perry[10], Dong Nguyen-Bresinsky[10], Audrie Seluke[10], Timothy Leach[10], Sara Perkins[10], Amanda Slater[10], Michaela Petit[10], Rachel Urban[10], Susan Kales[10], Michael DaCosta[10], Michael McFarland[10], Rick Palazola[10], Kevin A. Peterson[10], Karen Svenson[10], Robert E. Braun[10], Robert Taft[10], Mark Rhue[22], Jose Garay[22], Dave Clary[22], Renee Araiza[22], Kristin Grimsrud[22], Lynette Bower[22], Nicole L Anchell[22], Kayla M Jager[22], Diana L Young[22], Phuong T Dao[22], Wendy Gardiner[9], Toni Bell[9], Janet Kenyon[9], Michelle E Stewart[9], Denise Lynch[9], Jorik Loeffler[9], Adam Caulder[9], Rosie Hillier[9], Mohamed M Quwailid[5], Rumana Zaman[5], Luis Santos[5], Yuichi Obata[2], Mizuho Iwama[2], Hatsumi Nakata[2], Tomomi Hashimoto[2], Masayo Kadota[2], Hiroshi Masuya[2], Nobuhiko Tanaka[2], Ikuo Miura[2], Ikuko Yamada[2], Tamio Furuse[2], Mohammed Selloum[1], Sylvie Jacquot[1], Abdel Ayadi[1], Dalila Ali-Hadji[1], Philippe Charles[1], Elise Le Marchand[1], Amal El Amri[1], Christelle Kujath[1], Jean-Victor Fougerolle[1], Peggy Mellul[1], Sandrine Legeay[1], Laurent Vasseur[1], Anne-Isabelle Moro[1], Romain Lorentz[1], Laurence Schaeffer[1], Dominique Dreyer[1], Valérie Erbs[1], Benjamin Eisenmann[1], Giovanni Rossi[1], Laurence Luppi[1], Annelyse Mertz[1], Amélie Jeanblanc[1], Evelyn Grau[3], Caroline Sinclair[3], Ellen Brown[3], Helen Kundi[3], Alla Madich[3], Mike Woods[3], Laila Pearson[3], Danielle Mayhew[3], Nicola Griggs[3], Richard Houghton[3], James Bussell[3], Catherine Ingle[3], Sara Valentini[3], Diane Gleeson[3], Debarati Sethi[3], Tanya Bayzetinova[3], Jonathan Burvill[3], Bishoy Habib[3], Lauren Weavers[3], Ryea Maswood[3], Evelina Miklejewska[3], Ross Cook[3], Radka Platte[3], Stacey Price[3], Sapna Vyas[3], Adam Collinson[3], Matt Hardy[3], Priya Dalvi[3], Vivek Iyer[3], Tony West[3], Mark Thomas[3], Alejandro Mujica[3], Elodie Sins[3], Daniel Barrett[3], Michael Dobbie[43], Anne Grobler[44], Glaudina Loots[44], Rose Hayeshi[44], Liezl-Marie Scholtz[44], Cor Bester[44], Wihan Pheiffer[44], Kobus Venter[44], Fatima Bosch[45]

42 INFRAFRONTIER GmbH

43 Phenomics Australia, John Curtin School of Medical Research, The Australian National University, Acton, Canberra, Australia

44 PCDDP, North-West University, Potchefstroom, South Africa

45 Universitat Autònoma de Barcelona

A full list of members and their affiliations appears in the Supplementary Information.

entirely unexplored [1] The completion of the sequencing of the mouse genome, coupled with the use of mouse embryonic stem (ES) cells for gene targeting to create complex mutant alleles, presented an opportunity to functionally analyze all the protein coding genes of a mammalian species [2,3] Taking advantage of comprehensive manual annotation of the genome [4], the International Knockout Mouse Consortium (IKMC) systematically generated single-gene, reporter-tagged null alleles for protein-coding genes by homologous recombination in mouse ES cells [5,6] Subsequently, large-scale mouse production and phenotyping programs deployed these unique resources, establishing the feasibility of genome-scale mouse production and phenotyping [7–9] Building upon these successes, the International Mouse Phenotyping Consortium (IMPC) was established to coordinate a network of programs around the globe, assuring uniformity and reproducibility of these efforts, including standardization of phenotyping protocols and the use of a single inbred mouse strain background, C57BL/6N, with the ultimate goal of generating and phenotyping a single-gene knockout (KO) mouse line for every protein-coding gene in the genome.

Production of KO mice began in concert with the expansion of the ES cell library, but rapidly accelerated after 2011 with the funding of multiple IMPC programs. To date, more than 17,500 individual production attempts (microinjection or aggregation) have resulted in the germline transmission of KO alleles for 5,061 unique genes (Figure 1a; Supplementary Table 1). These lines have been expanded for phenotyping, providing key insights into mammalian biology and disease [10–15]; www.mousephenotype.org). To date, phenotype data for these lines shows that overall 72% of lines display at least one phenotype, revealing extensive pleiotropy (Supplementary Figure 1). This includes the 35.8% of lines that show partial or complete lethality, consistent with our earlier finding [11]. The IMPC contribution extends the total number of genes with targeted KO alleles produced by the scientific community from the 8,391 reported and curated by Mouse Genome Informatics (MGI; [16]) (Figure 1b; Supplementary Table 1), to 11,241, or more than half of the genome. Much of the overlap (2,211 genes) reflects specific community requests for the production of novel complex alleles (see below), targeting on an inbred C57BL/6N background, or for mutant mouse lines unavailable through public repositories. The growing use of CRISPR/Cas9 editing to produce null alleles for the IMPC led to the decrease in ES cell-based production beginning in 2015.

While the primary goal of the IKMC and IMPC was to generate and phenotype a null allele for every protein-coding gene, the mutant alleles included additional functional features. All alleles included a *lacZ* reporter cassette to facilitate analysis of gene transcription *in situ* (Supplementary Figure 2; [5,6]). A large proportion of the alleles have conditional potential, providing future users with a useful tool for detailed, mechanistic analyses (Supplementary Figure 2a). The multifunctional utility of the alleles produced by the IMPC has greatly expanded the repertoire of genetic resources available to the scientific community. Of the 3,674 unique gene, conditional-ready mouse models generated and validated, 2,987 were novel alleles for genes without an existing conditional allele (81.3%). These nearly double the total number of genes with conditional KO alleles produced by the scientific community as a whole (2,987 IMPC conditional alleles added to the 3,295 conditional alleles reported in MGI as mouse lines; Figure 1c). The impact is even more significant for reporter alleles. The IMPC has produced reporter alleles for 5,059 unique genes, of which 4,433 are novel

(87.6%), complementing the 2,733 produced by the scientific community (Figure 1d). This has nearly tripled the total number of genes with reporter alleles available to the community as mouse lines.

The generation of mouse lines was underpinned by comprehensive quality control strategies for both ES cell karyotype and targeted allele, which ensured efficient production and integrity of the targeting event (Supplementary Table 1; Supplementary Figures 3, 4). Further quality control (QC) analysis also showed that part of the ES cell collection contained an additional insertion of a wild-type nonagouti (*A*) gene on chromosome 8, likely introduced with the targeted reversion event in these cell lines (subclone JM8A3[17]). However, as the insertion of the wild-type nonagouti gene results in an agouti coat color, this allele can be easily segregated from the mutant allele in most cases (Supplementary Figure 5). High- throughput allele validation of ES cells was performed using either a suite of quantitative and endpoint PCR-based tests or a combination of Southern blot[18] and PCR-based analysis[19], depending on production center (Supplementary Figures 3, 4 and 6; Supplementary Tables 2, 3; Supplementary Note). Despite these efforts, we found that additional quality control (QC) on the mouse lines themselves was required to ensure all IMPC lines resulted from the transmission of the correctly targeted allele (Supplementary Figure 4; [19]). This additional QC at the mouse level identified a small but significant proportion of incorrect alleles that transmitted through the germline of chimera mice derived from clones that had passed initial and secondary validation QC testing in the ES cell. Our experience highlights the importance of careful allele validation before and after mouse production.

As a result of this effort, mouse lines with targeted alleles for more than 5,000 genes on a C57BL/6N genetic background with extensive and documented genetic validation of the targeted locus are now available to the biomedical research community, supporting high standards of reproducibility for future investigations. This resource nearly triples the number of genes with reporter alleles and almost doubles the number of conditional alleles available to the scientific community. When combined with more than 30 years of community effort, the total mutant allele mouse resource covers more than half of the genome. The IMPC resource has shown its usefulness through the continued and robust uptake of mutant mouse lines by investigators around the world. This includes both KO and conditional alleles with mouse lines distributed as live mice and cryopreserved stocks. To date, over 5,000 orders for mutant mice for 3,301 unique genes have been processed and shipped to more than 4,000 investigators around the world (Figure 1e). To date, more than 1,900 publications acknowledge the use of EUCOMM/KOMP alleles (for example [20–30]). This demonstrates the utility of these resources, the cumulative use of which continues to grow over time, and complements the systematic phenotyping efforts of IMPC centers. In the new era of genome editing, this ES cell-derived collection remains of unique value as it offers particularly sophisticated and quality-controlled alleles representing a cornerstone of the collective development of a null allele resource for the complete mammalian genome [2].

All data are freely available from the IMPC database hosted at EMBL-EBI via a web portal (mousephenotype.org), ftp (ftp://ftp.ebi.ac.uk/pub/databases/impc) and automatic programmatic interfaces. An archived version of the database will be maintained after

cessation of funding (exp. 2021) for an additional 5 years. Information on alleles, together with phenotype summaries, are additionally archived with Mouse Genome Informatics at the Jackson Laboratory via direct data submissions (J: 136110, J:148605, J:157064, J:157065, J:188991, J:211773).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Authors

Marie-Christine Birling[1], Atsushi Yoshiki[2], David J Adams[3], Shinya Ayabe[2], Arthur L Beaudet[4,27], Joanna Bottomley[3], Allan Bradley[3,28], Steve DM Brown[5], Antje Bürger[6], Wendy Bushell[3,29], Francesco Chiani[7], Hsian-Jean Genie Chin[8], Skevoulla Christou[9], Gemma F Codner[9], Francesco J DeMayo[4,30], Mary E Dickinson[4], Brendan Doe[3], Leah Rae Donahue[10], Martin D Fray[9], Alessia Gambadoro[7], Xiang Gao[11], Marina Gertsenstein[12], Alba Gomez-Segura[14], Leslie O Goodwin[10], Jason D Heaney[4], Yann Hérault[1], Martin Hrabe de Angelis[15,16,17], Si-Tse Jiang[8], Monica J Justice[4,12,13], Petr Kasparek[18], Ruairidh E King[5], Ralf Kühn[6,31,32], Ho Lee[19], Young Jae Lee[20], Zhiwei Liu[21], K C Kent Lloyd[22], Isabel Lorenzo[4], Ann-Marie Mallon[5], Colin McKerlie[12,13], Terrence F Meehan[14,33], Violeta Munoz Fuentes[14], Stuart Newman[3,34], Lauryl MJ Nutter[12,13], Goo Taeg Oh[23], Guillaume Pavlovic[1], Ramiro Ramirez-Solis[3], Barry Rosen[3,35], Edward J Ryder[3,36], Luis A Santos[5], Joel Schick[6,37], John R Seavitt[4], Radislav Sedlacek[18], Claudia Seisenberger[6], Je Kyung Seong[24], William C Skarnes[3,38], Tania Sorg[1], Karen P Steel[3,39], Masaru Tamura[2], Glauco P Tocchini-Valentini[7], Chi-Kuang Leo Wang[8], Hannah Wardle-Jones[3], Marie Wattenhofer-Donzé[1], Sara Wells[9], Michael V. Wiles[10], Brandon J Willis[22], Joshua A Wood[22], Wolfgang Wurst[6,25,26], Ying Xu[21], IMPC Consortium*, Lydia Teboul[9,40], Stephen A Murray[10,41]

Affiliations

[1] Université de Strasbourg, CNRS, INSERM, PHENOMIN-ICS, IGBMC, Illkirch, France

[2] RIKEN BioResource Research Center, Tsukuba, Ibaraki, Japan

[3] Wellcome Sanger Institute, Hinxton, UK

[4] Baylor College of Medicine, Houston, Texas, USA

[5] MRC Harwell Institute, Mammalian Genetics Unit, Didcot, UK

[6] Institute of Developmental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, Neuherberg, Germany

[7] Monterotondo Mouse Clinic, Italian National Research Council (CNR), Institute of Cell Biology and Neurobiology, Monterotondo Scalo, Italy

[8] National Laboratory Animal Center, National Applied Research Laboratories (NARLabs), Taipei, Taiwan

[9] MRC Harwell Institute, The Mary Lyon Centre, Didcot, UK

[10] The Jackson Laboratory, Bar Harbor, Maine, USA

[11] SKL of Pharmaceutical Biotechnology and Model Animal Research Center, Collaborative Innovation Center for Genetics and Development, Nanjing Biomedical Research Institute, Nanjing University, Nanjing, China.

[12] The Centre for Phenogenomics, Toronto, Ontario, Canada

[13] The Hospital for Sick Children, Toronto, Ontario, Canada

[14] European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL- EBI), Hinxton, UK

[15] German Mouse Clinic, Institute of Experimental Genetics, Helmholtz Zentrum München, German Research Center for Environmental Health GmbH, Neuherberg, Germany

[16] Experimental Genetics, Center of Life and Food Sciences Weihenstephan, Technische Universität München, Freising-Weihenstephan, Germany

[17] German Center for Diabetes Research, Neuherberg, Germany

[18] Czech Centre for Phenogenomics, Institute of Molecular Genetics of the Czech Academy of Sciences, Vestec, Czech Republic

[19] Korea Mouse Phenotyping Center (KMPC) and Graduate School of Cancer Science and Policy, National Cancer Center, Gyeonggi, Republic of Korea

[20] Korea Mouse Phenotyping Center (KMPC) and Lee Gil Ya Cancer and Diabetes Institute, Gachon University,Incheon, Republic of Korea

[21] CAM-SU Genomic Resource Center, Soochow University, Suzhou, China

[22] Mouse Biology Program, University of California, Davis, California, USA

[23] Immune and Vascular Cell Network Research Center, National Creative Initiatives and Department of Life Sciences, Ewha Womans Univesity, Seoul, Republic of Korea

[24] Korea Mouse Phenotyping Center (KMPC) and BK21 Program for Veterinary Science, Research Institute for Veterinary Science, College of Veterinary Medicine, Seoul National University, Seoul, Republic of Korea

[25] Developmental Genetics, Center of Life and Food Sciences Weihenstephan, Technische Universität München, Freising-Weihenstephan, Germany

[26] German Center for Neurodegenerative Diseases, Munich, Germany

[27] Current affiliation: Luna Genetics, Houston, Texas, USA

[28] Current affiliation: Cambridge Institute of Therapeutic Immunology & Infectious Disease (CITIID), Jeffrey Cheah Biomedical Centre, University of Cambridge, Cambridge, UK

[29] Current affiliation: IONTAS Ltd, London Road, Cambridge UK

[30] Current affiliation: National Institute for Environmental Health Science Research, North Carolina, USA

[31] Current affiliation: Max Delbrueck Center for Molecular Medicine, Berlin, Germany

[32] Current affiliation: Berlin Institute of Health, Berlin, Germany

[33] Current affiliation: Kymab Group Ltd, Cambridge, UK

[34] Current affiliation: PetMedix Ltd, Cambridge, UK

[35] Current affiliation: AstraZeneca, Discovery Sciences, Cambridge, UK

[36] Current affiliation: LGC, Sport and Specialised Analytical Services, Fordham, UK

[37] Current affiliation: Genetics and Cellular Engineering Group, Institute of Molecular Toxicology and Pharmacology, Helmholtz Zentrum Munich, Neuherberg, Germany

[38] Current affiliation: The Jackson Laboratory for Genomic Medicine, Farmington, Connecticut, USA

[39] Current affiliation: Wolfson Centre for Age-Related Diseases, King's College London, London, UK

Acknowledgments

We thank all technical personnel at the different centers involved in this project for their contribution. We thank G. Clarke for support with illustrations. M.-C.B., G.P., M. W.-D., Y.H., and T.S. were supported by the Université de Strasbourg, the CNRS, the INSERM and the programmes 'Investissements d'avenir' (ANR-10-IDEX-0002–02, ANR-10-LABX-0030- INRT, ANR-10-INBS-07 PHENOMIN), AY. and M.T. were supported by RIKEN BioResource Research Center (BRC), Management Expenses Grant from the Ministry of Education, Culture, Sports, Science and Technology (MEXT), D.A, J.B., A.B., W.B., B.D., S.N., R.R.S., B.R., E.R., W.S., K.S. and H.W.-J. were supported by the Wellcome Trust, S.B., S.C., G.C., M.F., S.W. and L.T. were supported by the Medical Research Council (A410), M.G., C.McK. and L.N. were supported by Genome Canada and Ontario Genomics (OGI-051),), H.L., Y.J.L., G.T.O., and J.K.S. supported by National Research Foundation (2014M3A9D5A01074636, 2014M3A9D5A01075128, 2013M3A9D5072550), Republic of Korea (KMPC), M B., A.B., S B., A.Bu., W.B., F.C., M.F., A.G., M. H. A., R. K., S.N., G.P., R.R.S., B.R., E.R., J.S., W.S., C.S., T.S., G.T.-V, S.W., W.W. and L.T. were supported by the European Commission (EUCOMM (LSHM-CT-2005–018931), EUCOMMTOOLS (FP7- HEALTH-F4–2010-261492), EUMODIC, Infrafrontier 01KX1012 (M.H.A.); EU Horizon 2020: IPAD-MD funding 653961 (M.H.A.)). R.S. and P.K. were supported by RVO 68378050 by Academy of Sciences of the Czech Republic and by LM2015040 and LM2018126 (Czech Centre for Phenogenomics), CZ.1.05/2.1.00/19.0395, CZ.1.05/1.1.00/02.0109 funded by the Ministry of Education, Youth and Sports and the European Regional Development Fund. Work was supported by the National Key R&D Program of China 2018YFA0801100 to Y.X. Research reported in this publication was supported by the NIH Common Fund, the Office of The Director, and the National Human Genomic Research Institute of the National Institutes of Health under Award Numbers (U42OD011174 supported A.L.B. J.B., A.B., W.B., B.D., M.D., M.F., J. H., M. J., I.L., F.M., S.N., R.R.S., B.R., E.R., W.S., J.S., S.W., and L.T.; U42OD011175 supported M.G., C.McK., L.N., B.W., J.W, K.C.L.; U42OD011185 supported L.R.D., L.G., M.V.W. and S.A.M.; U54HG006370–02 supported A. G-S., R.E.K., A.-M.M., T.M., V.M.F. and L.S.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

1. Stoeger T, Gerlach M, Morimoto RI & Nunes Amaral LA Large-scale investigation of the reasons why potentially important genes are ignored. PLoS Biol 16, e2006643 (2018). [PubMed: 30226837]
2. Austin CPet al.The knockout mouse project. Nat Genet36, 921–4 (2004). [PubMed: 15340423]

3. Auwerx Jet al.The European dimension for the mouse genome mutagenesis program. Nat Genet36, 925–7 (2004). [PubMed: 15340424]
4. Ashurst JLet al.The Vertebrate Genome Annotation (Vega) database. Nucleic Acids Res33, D459–65 (2005). [PubMed: 15608237]
5. Skarnes WCet al.A conditional knockout resource for the genome-wide study of mouse gene function. Nature474, 337–42 (2011). [PubMed: 21677750]
6. Valenzuela DMet al.High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. Nat Biotechnol21, 652–9 (2003). [PubMed: 12730667]
7. de Angelis MHet al.Analysis of mammalian gene function through broad-based phenotypic screens across a consortium of mouse clinics. Nat Genet47, 969–978 (2015). [PubMed: 26214591]
8. White JKet al.Genome-wide generation and systematic phenotyping of knockout mice reveals new roles for many genes. Cell154, 452–64 (2013). [PubMed: 23870131]
9. West DBet al.A lacZ reporter gene expression atlas for 313 adult KOMP mutant mouse lines. Genome Res25, 598–607 (2015). [PubMed: 25591789]
10. Bowl MRet al.A large scale hearing loss screen reveals an extensive unexplored genetic landscape for auditory dysfunction. Nat Commun8, 886 (2017). [PubMed: 29026089]
11. Dickinson MEet al.High-throughput discovery of novel developmental phenotypes. Nature537, 508–514 (2016). [PubMed: 27626380]
12. Karp NAet al.Prevalence of sexual dimorphism in mammalian phenotypic traits. Nat Commun8, 15475 (2017). [PubMed: 28650954]
13. Meehan TFet al.Disease model discovery from 3,328 gene knockouts by The International Mouse Phenotyping Consortium. Nat Genet49, 1231–1238 (2017). [PubMed: 28650483]
14. Munoz-Fuentes Vet al.The International Mouse Phenotyping Consortium (IMPC): a functional catalogue of the mammalian genome that informs conservation. Conserv Genet19, 995–1005 (2018). [PubMed: 30100824]
15. Rozman Jet al.Identification of genetic elements in metabolism by high-throughput mouse phenotyping. Nat Commun9, 288 (2018). [PubMed: 29348434]
16. Bult CJet al.Mouse Genome Database (MGD) 2019. Nucleic Acids Res47, D801–D806 (2019). [PubMed: 30407599]
17. Pettitt SJet al.Agouti C57BL/6N embryonic stem cells for mouse genetic resources. Nat Methods6, 493–5 (2009). [PubMed: 19525957]
18. Codner GFet al.Universal Southern blot protocol with cold or radioactive probes for the validation of alleles obtained by homologous recombination. Methods (2020).
19. Ryder Eet al.Molecular characterization of mutant mouse strains generated from the EUCOMM/KOMP-CSD ES cell resource. Mamm Genome24, 286–94 (2013). [PubMed: 23912999]
20. Ingham NJet al.Mouse screen reveals multiple new genes underlying mouse and human hearing loss. PLoS Biol17, e3000194 (2019). [PubMed: 30973865]
21. Liu Xet al.The complex genetics of hypoplastic left heart syndrome. Nat Genet49, 1152–1159 (2017). [PubMed: 28530678]
22. Miyata Het al.Genome engineering uncovers 54 evolutionarily conserved and testis-enriched genes that are not required for male fertility in mice. Proc Natl Acad Sci USA113, 7704–10 (2016). [PubMed: 27357688]
23. Pol Aet al.Mutations in SELENBP1, encoding a novel human methanethiol oxidase, cause extraoral halitosis. Nat Genet50, 120–129 (2018). [PubMed: 29255262]
24. Small KSet al.Regulatory variants at KLF14 influence type 2 diabetes risk via a female-specific effect on adipocyte size and body composition. Nat Genet50, 572–580 (2018). [PubMed: 29632379]
25. Akawi Net al.Discovery of four recessive developmental disorders using probabilistic genotype and phenotype matching among 4,125 families. Nat Genet47, 1363–9 (2015). [PubMed: 26437029]
26. Bertero Aet al.Activin/nodal signaling and NANOG orchestrate human embryonic stem cell fate decisions by controlling the H3K4me3 chromatin mark. Genes Dev29, 702–17 (2015). [PubMed: 25805847]

27. de la Rosa Jet al.A single-copy Sleeping Beauty transposon mutagenesis screen identifies new PTEN-cooperating tumor suppressor genes. Nat Genet49, 730–741 (2017). [PubMed: 28319090]

28. Kim JHet al.Regulation of the catabolic cascade in osteoarthritis by the zinc-ZIP8- MTF1 axis. Cell156, 730–43 (2014). [PubMed: 24529376]

29. Koch S, Acebron SP, Herbst J, Hatiboglu G & Niehrs C Post-transcriptional Wnt Signaling Governs Epididymal Sperm Maturation. Cell 163, 1225–1236 (2015). [PubMed: 26590424]

30. Kochubey O, Babai N & Schneggenburger R A Synaptotagmin Isoform Switch during the Development of an Identified CNS Synapse. Neuron 90, 984–99 (2016). [PubMed: 27210552]

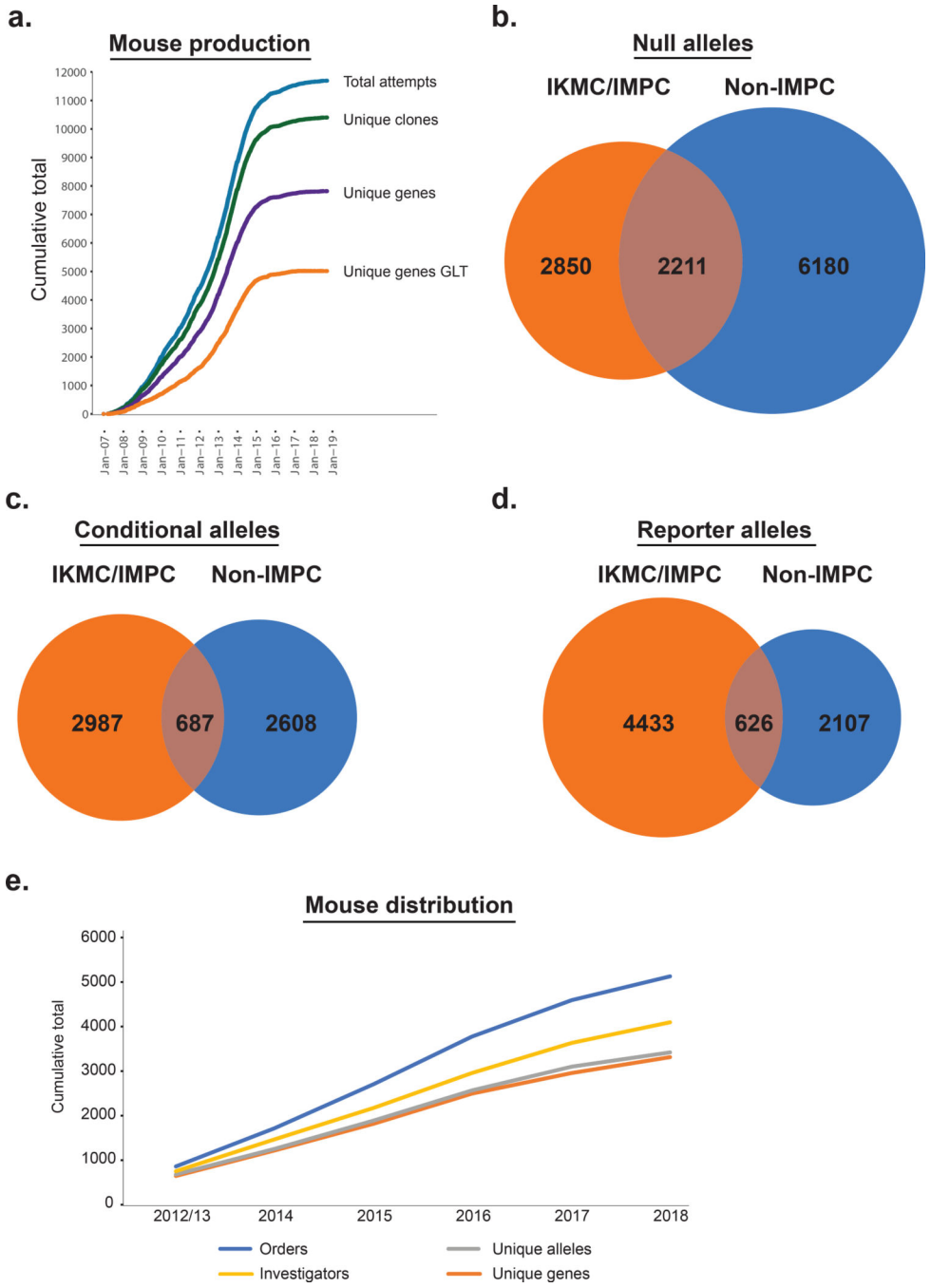

Figure 1:

Generation and impact of targeted alleles for 5,061 unique mouse genes. (a) Cumulative production progress, including all attempts (microinjection or aggregation (black), unique ES cell clones injected (red), unique genes attempted (yellow), and unique genes that achieved germline transmission (GLT; blue). For GLT, the date reflects the date of microinjection, and only reports the first instance of transmission for the small number of duplicate mutations produced. (b) Venn representation of unique gene null alleles produced by the IMPC (orange) and by the rest of the scientific community as reported in MGI ("Non-

IMPC"; blue). (c) Unique gene conditional-ready alleles produced by the IMPC (orange) and by the rest of the scientific community (blue). (d) Unique gene reporter alleles produced by the IMPC (orange) and by the rest of the scientific community (blue). (e) Cumulative mouse orders of IMPC lines processed by production centres and mouse model Repositories from 2012–2018 (blue line). The cumulative number of ordering investigators, unique alleles ordered, and unique genes ordered are shown in yellow, grey, and orange, respectively.