Mamm Genome (2008) 19:379–381
DOI 10.1007/s00335-008-9113-1

COMMENTARY

The Collaborative Cross, developing a resource for mammalian systems genetics: A status report of the Wellcome Trust cohort

Fuad A. Iraqi · Gary Churchill · Richard Mott

Received: 10 April 2008 / Accepted: 1 May 2008 / Published online: 3 June 2008

Abstract We report on the progress of a project funded by the Wellcome Trust to produce over 100 recombinant inbred mouse lines as part of the Collaborative Cross (CC) genetic reference panel. These new strains of mice are being derived from a set of eight genetically diverse founders. The genomes of the finished strains will be mosaics of the founder strains' genomes with a high density of independent recombination breakpoints. The CC mice will be available for distribution free of any intellectual property constraints to serve as a community resource for systems genetics studies.

Introduction

In recent years new techniques have emerged for utilizing genetic perturbations as a tool for systems genetics studies (Jansen and Nap 2001). To date such studies have been carried out largely on transient populations of intercross animals, thus limiting our ability to integrate the information obtained. A sufficiently large and genetically diverse population of animals that are replicable can be studied by multiple researchers under a broad array of environmental conditions and phenotyping protocols, providing a common integration point for vast amounts of molecular data that are expected to be generated in the coming decades. Although standard intercrosses between inbred lines of mice are powerful tools for mapping quantitative trait loci (QTL), with some important exceptions the genes underlying the QTLs remain unknown because the intervals are too broad. Recognition of these problems—lack of integration and low mapping resolution—resulted in the realization that a new general-purpose mouse resource was needed to model complex human diseases, with particular emphasis on traits relevant to human health in its broadest aspects. Open discussion among members of the genetics community resulted in the conception and design of the Collaborative Cross (CC) (Churchill et al. 2004; Threadgill et al. 2002).

The CC population is now being created through a community effort by the complex trait consortium (CTC) (www.complextrait.org). The CC is a unique, genetically diverse reference population of recombinant inbred mouse lines descended from eight divergent strains of mice: A/J, C57BL/6J, 129S1/SvImJ, NOD/LtJ, NZO/HiLtJ, CAST/Ei, PWK/PhJ, and WSB/EiJ (Roberts et al. 2007). Once fully inbred, the average distance between recombinants in a CC genome will be approximately 12 Mb and the QTL mapping resolution of the order of 1 Mb (Broman 2005; Valdar et al. 2006).

Initial funding by the Wellcome Trust in 2005 supported the breeding of upward of 100 lines, initially at the International Livestock Research Institute (ILRI), based in Kenya. The project was transferred to its current location at Tel Aviv University in Israel in 2007. The U.S. Department of Energy (DOE), The Ellison Foundation, and the U.S. National Institutes of Health (NIH) supported another

F. A. Iraqi (✉)
Department of Human Microbiology, Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel
e-mail: fuadi@post.tau.ac.il

G. Churchill
The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609, USA
e-mail: gary.churchill@jax.org

R. Mott
Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
e-mail: Richard.Mott@well.ox.ac.uk

colony of upward of 300 lines in Oak Ridge National Laboratory, Tennessee, USA, with further lines planned. The breeding design ensures that the autosomal genomes of each line will have equal input contributions from each of the founder strains and that recombination events that accumulate during the breeding process are independent between lines. Some care is required to ensure approximate balance of the sex chromosomes, the mitochondrial genome, and their pairwise combinations. An eight-way "funnel" breeding design was used to mix the genomes before inbreeding by brother-sister mating. Labeling the eight founders with the letters A to H, the funnel design first chooses four F_1 matings, say A × B, C × D, E × F, G × H. Then two G_2 matings AB × CD and EF × GH are made, and finally one G_3 mating ABCD × EFGH. All of the Wellcome CC lines were descended from a set of F_1 matings made at The Jackson Laboratory, Bar Harbor, Maine, USA. The F_1 hybrids were distributed to multiple sites where they were mated to produce a balanced set of funnels. The F_3 generations were then sib-mated to produce unique incipient CC recombinant inbred lines (RIL). The inbreeding process is still ongoing. Once inbred, these mice will be available for distribution, free of any intellectual property constraint, and will serve as a powerful tool for quantitative trait analysis and systems biology.

Status of the Wellcome Trust CC mouse subpopulation

The Wellcome Trust-funded component of the CC was started in July 2005 with 200 lines, and at the time of writing the 110 remaining lines in this cohort are variously between the 6th and 11th generation of inbreeding, reflecting differences in fertility between the lines. An attrition rate of approximately 50% of the lines is expected by the end of the inbreeding stage.

Phenotypic data

During the course of breeding we have collected DNA samples (tail clips) and phenotypic measurements from every breeding animal. These will serve as a resource for future genetic studies and for mapping QTL. Mothers are mating after weaning at 21 days and we expect first litters to follow in 42–60 days. Some interesting observations on fertility have been made. For example, while some G_2 matings of F_1 pairs were fertile and produced their first litter within the expected age, some lines were fertile at a much later ages, the maximum being 205 days. Nevertheless, more than 70% of the population was fertile between 44 and 60 days, 20% at between 60 and 100 days, and only 10% between 100 and 205 days. Ten F_1 pairs did not litter at all; these mostly carried NZO alleles and consequently were eliminated. The remaining 190 matings were successful. Variation in G_3 fertility age ranged from 46 to 128 days; very late fertility age disappeared. Eleven G_3 lines were infertile and were eliminated so we started the inbreeding process with 179 lines. Inbreeding was performed by selecting brother and sister littermates. The distribution of fertility ages in the G_4 generation (the first generation of inbreeding) was tighter; 64% of lines between 59 and 79 days, 30% between 50 and 58 days, and 6% between 80 and 128 days. From G_5 onward, the fertility age was between 44 and 64 days after mating, which agrees with that of commercially available inbred lines. As noted above, currently 110 lines are at inbreeding generations between 6 and 11 and continue to have a narrow distribution of fertility age. Thus, as the genomes of the animals became more mixed, the phenotypic extremes became rarer.

Litter size and *survival* before and after weaning were recorded and compared between the different generations. There is a decrease in the litter size and survival rate during the progression of inbreeding, and some lines have low fertility, as expected, due to inbreeding depression. Litter sizes among lines ranged between 1 and 15 mice. The average litter size was 8.5, 7.9, 6.4, 6.1, 5.7, 4.9, 5.3, 4.5, and 4 at generations G_2 to G_{10}, respectively. Some lines had a tendency to produce asymmetric sex ratios in the litters, but usually equal numbers of males and females were produced at all generations. The survival rate of mice at different generations was between 68 and 81%, with a significant decrease at G6. The mean survival rate of 8000 recorded births from all generations was 73%.

Body weights of male and female mice varied greatly among lines. On average, males are larger than females and both sexes exhibit a decrease in weight with inbreeding generation number. However, in some lines females have a higher body weight than the males. In 7-week-old G_2 females, body weight ranged from 13 to 32.9 g, while males of the same age were between 9.3 and 41.9 g. Average body weight of males at 8 weeks old on G_3 was 22.6 g (SD = 3.5), while females had a mean weight of 20.0 g (SD = 3.8).

We observed four major *coat colors*, white, black, agouti, and nonagouti brown, in over 7000 mice, and some diluted colors were also observed. The original coat colors of the eight founders are white for A/J and NOD/LtJ, black for C57BL/6J, brown for 129S1/SvImJ, NZO/HiLtJ, CAST/Ei, and PWK/PhJ, and gray for WSB/EiJ. Agouti accounted for 56.1% CC animals, followed by 22.3% white, 12.4% black, 6.1% nonagouti brown, 2.8% brown, and finally 0.3% black diluted (gray) color. The distribution of coat colors was equal in both sexes. The WSB/EiJ strain has a white spot on its head in between the ears (its

name was generated from the definition of *w*hite *s*pot *b*rain) and which some lines have inherited.

Future plans

We expect inbreeding to be at generation 24 in most lines by 2012, at which point we plan to make the lines available through a European distribution center. However, already most lines are more than 80% inbred and therefore usable. Genotypes collected on the current generation of mice will be good predictors of genotypes on subsequent generations. In collaboration with our U.S. partners, we will begin a pilot project in the near future to investigate the use and power of the Collaborative Cross for QTL mapping. In addition, several QTL mapping projects have already been established for understanding host susceptibility to infectious diseases.

Acknowledgments The authors appreciate the generous support of Wellcome Trust for funding this project, and the International Livestock Research Institute (ILRI) and Tel Aviv University for their technical support.

References

Broman K (2005) The genomes of recombinant inbred lines. Genetics 169:1133–1146

Churchill GA, Airey DC, Allayee H, Angel JM, Attie AD et al (2004) The Collaborative Cross, a community resource for the genetic analysis of complex traits. Nat Genet 36(11):1133–1137

Jansen RC, Nap JP (2001) Genetical Genomics: the added value from segregation. Trends Genet 17:388–391

Roberts A, Pardo-Manuel de Villena F, Wang W, McMillan L, Threadgill DW (2007) The polymorphism architecture of mouse genetic resources elucidated using genome-wide resequencing data: implications for QTL discovery and systems genetics. Mamm Genome 18(6–7):473–481

Threadgill DW, Hunter KW, Williams RW (2002) Genetic dissection of complex and quantitative traits: from fantasy to reality via a community effort. Mamm Genome 13:175–178

Valdar W, Flint J, Mott R (2006) Simulating the collaborative cross: power of quantitative trait loci detection and mapping resolution in large sets of recombinant inbred strains of mice. Genetics 172(3):1783–1797