Copyright © 2007 by the Genetics Society of America
DOI: 10.1534/genetics.107.078006

# Interspecific Recombinant Congenic Strains Between C57BL/6 and Mice of the *Mus spretus* Species: A Powerful Tool to Dissect Genetic Control of Complex Traits

**Gaëtan Burgio,*,† Marek Szatanik,‡ Jean-Louis Guénet,* Maria-Rosa Arnau,§ Jean-Jacques Panthier* and Xavier Montagutelli*,[1]**

**Unité Postulante de Génétique Fonctionnelle de la Souris, CNRS URA 2578 and ‡Unité des Neisseria, Institut Pasteur, 75015 Paris, France, †Unité d'Éco-Anthropologie, Équipe "Génétique des Populations Humaines," CNRS MNHN, P7, UMR 5145, Musée de l'Homme, 75016 Paris, France and §Universidad de la Laguna, 38207 Teneriffe, Spain*

Manuscript received June 21, 2007
Accepted for publication September 21, 2007

ABSTRACT

Complex traits are under the genetic control of multiple genes, often with weak effects and strong epistatic interactions. We developed two new collections of mouse strains to improve genetic dissection of complex traits. They are derived from several backcrosses of the *Mus spretus* SEG/Pas or STF/Pas strains on the C57BL/6J background. Each of the 55 interspecific recombinant congenic strains (IRCSs) carries up to eight SEG/Pas chromosomal segments with an average size of 11.7 Mb, totalizing 1.37% of the genome. The complete series covers 39.7% of the SEG/Pas genome. As a complementary resource, six partial or complete interspecific consomic strains were developed and increased genome coverage to 45.6%. To evaluate the usefulness of these strains for QTL mapping, 16 IRCSs were compared with C57BL/6J for seven hematological parameters. Strain 66H, which carries three SEG/Pas chromosomal segments, had lower red blood cell volume and higher platelet count than C57BL/6J. Each chromosomal segment was isolated in a congenic strain to evaluate individual effects. Congenic strains were combined to assess epistasis. Our data show that both traits were controlled by several genes with complex epistatic interactions. IRCSs are therefore useful to unravel QTL with small effects and gene-by-gene interactions.

INDIVIDUALS of every species exhibit more or less profound differences that affect all aspects of their anatomy, physiology, development, behavior, and susceptibility to various diseases. Most of these phenotypic differences are under the control of multiple genes and environmental factors and are therefore referred to as complex traits, by contrast with monogenic traits, where a single gene modification controls phenotypic variation. Complex traits are assessed in individuals as one or more quantitative measurements, so that genes controlling these variations are quantitative trait loci (QTL). Several notable features of QTL which make their identification difficult are the lack of strict relationship between phenotypes and genotypes, low penetrance, epistasis, and pleiotropy.

Because of their genetic architecture, complex traits require that sophisticated mapping populations be developed and studied (Abiola *et al.* 2003; Flint *et al.* 2005). Model organisms such as the mouse offer the advantage of providing conditions resembling human pathology (*e.g.*, diabetes mellitus, asthma, hypertension, and atherosclerosis) together with the ability to produce segregating populations optimized for the dissection and functional analysis of the underlying genes.

Complex traits are often analyzed using two-generation crosses, such as backcrosses and $F_2$'s. While such crosses can be produced between almost any pair of inbred strains, their analysis results in low power to detect weak QTL, produces large confidence intervals (Darvasi *et al.* 1993), and requires that presumptive QTL be confirmed in congenic strains (Abiola *et al.* 2003). Heterogeneous stocks (HSs) represent an outbred population produced from an eight-way intercross followed by several generations of random breeding to increase the density of recombination events (Mott and Flint 2002). This strategy has yielded the identification of an incredibly large number of QTL with very high LOD scores and extremely small confidence intervals (Valdar *et al.* 2006). However, neither two-generation crosses nor HSs can be perpetuated, hence preventing further studies on the same population.

Recombinant inbred strains (RISs) are permanent resources derived from an $F_2$ cross (Bailey 1971, 1981; Taylor 1978). Their inbred status allows replication of phenotype assessment in genetically identical individuals and for correlations between traits that cannot be measured in the same individuals. However, most existing sets, which were developed between a restricted

[1]*Corresponding author:* Unité Postulante de Génétique Fonctionnelle de la Souris, 25 Rue du Docteur Roux, 75724 Paris Cedex 15, France. E-mail: xmonta@pasteur.fr

number of laboratory inbred strains, are too small to provide the required power for QTL detection and fine mapping, especially when small-effect QTL are involved, with some degree of epistasis (Flint *et al.* 2005). To overcome these limitations, a much larger reference population is currently under development. The collaborative cross aims to produce a series of >1000 inbred strains with a balanced contribution from eight parental inbred strains, two of which are derived from a different subspecies (Churchill *et al.* 2004). Consomic strains are an ordered series of inbred strains that share the same inbred genetic background and differ by only one entire chromosome, which has been introgressed from another inbred strain by repeated backcrosses (Singer *et al.* 2004). While any phenotypic variation between a consomic strain and the background strain can readily be attributed to one or more genes on the differential chromosome, epistatic interaction between nonsyntenic genes may be missed. Other genetic tools developed for genetic analysis of complex traits include sets of overlapping congenic strains covering the entire genome (Iakoubova *et al.* 2001; Davis *et al.* 2005).

Recombinant congenic strains (RCSs) were introduced by Demant and Hart (1986) as inbred populations derived after two or three backcrosses, resulting in unequal genomic contribution from the two parental inbred strains. With three backcrosses, each RCS is expected to carry 12.5% (one-eighth) of its genome from the "donor" strain. Theoretical considerations predict that ~95% of the genome of the donor strain can be transferred and distributed within a series of 20 RCSs (Demant and Hart 1986; Moen *et al.* 1997). Five sets of such RCSs have been produced between pairs of laboratory inbred strains (Groot *et al.* 1992; Martin *et al.* 1992b; Fortin *et al.* 2001b) and the results have fulfilled the theoretical expectations (Stassen *et al.* 1996). Since each RCS carries only a small fraction of the donor genome, multiple genes controlling a trait are likely to be isolated in different strains (Demant and Hart 1986; Moen *et al.* 1991). QTL mapping is achieved in an $F_2$ population produced between the "background" strain and an RCS with a contrasting phenotype (Van Wezel *et al.* 1996). Because only one-eighth of the genome is segregating in such a cross, the total genetic variance is reduced and the power to identify QTL is higher, as is the ability to detect epistatic interactions since a lower value of the statistics is required to declare significant an interaction between two loci. Recombinant congenic strains have been successfully used to study a variety of complex traits (Martin *et al.* 1992a; Moen *et al.* 1992, 1996; Serreze *et al.* 1994; Fortin *et al.* 2001a; Banus *et al.* 2005; Lemay and Haston 2005; Lee *et al.* 2006; Roy *et al.* 2006; Gouya *et al.* 2007).

Besides genetic populations of appropriate structure, genetic analysis of complex traits also requires phenotypic polymorphism. While most genetic tools have been developed using laboratory mice, wild-derived inbred strains have proven highly valuable in providing original genotypic and phenotypic variations (Guénet and Montagutelli 1994; Guénet and Bonhomme 2003). They offer a much higher level of diversity (Ideraabdullah *et al.* 2004), and polymorphism rate at the sequence level is estimated to be 1–1.5% between C57BL/6J and *Mus spretus*-derived strains, such as SEG/Pas (Guénet and Bonhomme 2003), the same order of magnitude as that between human and chimpanzee (Newman *et al.* 2005). Interspecific and intersubspecific crosses have been at the origin of high-density, high-resolution genetic maps (Avner *et al.* 1988; Rhodes *et al.* 1998). Wild-derived inbred strains also differ from laboratory mice in their susceptibility to various traits (Melanitou *et al.* 1998; Staelens *et al.* 2002; Yi *et al.* 2004), including viral (Mashimo *et al.* 2002), bacterial (Sebastiani *et al.* 2002; Turcotte *et al.* 2006), and parasitic (Bagot *et al.* 2002) infections. It has been suggested, however, that, because of this high genetic polymorphism, phenotypic differences might be under the control of many QTL, which could hamper the identification of the genes.

In an attempt to combine the power of interspecific crosses and that of recombinant congenic strains, we developed a set of interspecific recombinant congenic strains (IRCSs), designated BcG, with the original aim of introgressing 12.5% of the *M. spretus* genome (strain SEG/Pas) in a C57BL/6J inbred background. A total of 55 strains were derived from three to five backcrosses, bred to homozygosity, and genotyped for 673 genetic markers spread over the genome. Surprisingly, the final proportion of the *M. spretus* genome retained in each strain varied from 0 to 3.8%, with an average of 1.37%. This sixfold reduction compared to expectations based on the actual crosses was due to strong selection against *M. spretus* alleles in the course of inbreeding generations. In parallel, we attempted to develop a set of interspecific consomic strains (ICSs) from the same strain combination. Only fractions of six chromosomes could be retained. Altogether, IRCSs and ICSs cover 45.6% of the genome. We have analyzed a few IRCSs for several quantitative traits and we demonstrate that they offer a useful tool to explore the complexity of the genetic control of multigenic traits.

## MATERIALS AND METHODS

**Mice and crosses:** IRCSs were developed at the Institut Pasteur in Paris. C57BL/6JIco (B6) mice were purchased from IFFA-CREDO (now Charles River Laboratories, L'Arbresle, France). The SEG/Pas colony was established from breeders provided by François Bonhomme (Montpellier, France) and maintained by relaxed inbreeding (closed colony).

SEG/Pas males were mated to B6 females to produce $F_1$ females. $F_1$ females were mated to B6 males to produce an $N_2$ generation that became part of the EUCIB mapping panel (Breen *et al.* 1994; Rhodes *et al.* 1998). $N_2$ males ($N = 157$)

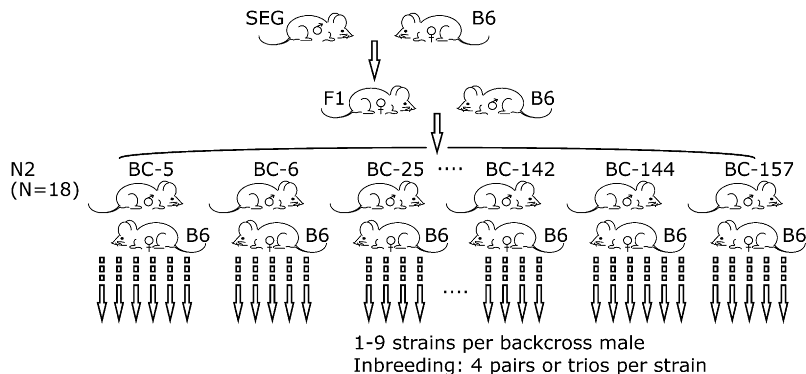


FIGURE 1.—IRCS breeding scheme. SEG/Pas males were bred to B6 to produce $F_1$ females that were crossed to B6 males. Resulting backcross males were selected for fertility and mated with B6 females. From their progeny, breeding pairs were established as a starting point for the development of inbred strains. As soon as possible, each strain was maintained as four pairs or trios, with relaxed inbreeding. One to nine breeding pairs could be established from each backcross male. Each strain was named after the number of the backcross male, followed by a letter.

were mated to two $OF_1$ outbred females IFFA-CREDO to assess their fertility. Eighteen fertile males were then mated to B6 females to yield N3 progeny. Between four and eight N3 breeding pairs were established from the progeny of each $N_2$ male, as the starting point for recombinant congenic strains (Figure 1). As many as four breeding pairs or trios were set with the progeny of each N3 intercross. In the following generations, each strain was maintained as a set of four breeding pairs or trios. Whenever possible, mating involved full brothers and sisters.

Approximately half of the strains derived from N3 mice stopped breeding after one to four generations of inbreeding. About half of them were definitely lost. For the other half, one additional backcross to B6 was performed by mating one or more males of the strain to B6 females. Inbreeding was restarted from their progeny.

We initially estimated the number of IRCSs needed to cover the genome as follows. Let $K$ be the number of $N_2$ males used as progenitors, $N$ be the total number of strains, $N_i$ be the number of strains derived from $N_2$ male $i$ ($N = \sum N_i$), and $P$ be the probability that all $N$ IRCSs are homozygous for the B6 allele at a given locus. The probability that a given strain will be B/B at a locus is the sum of two probabilities: the probability that the $N_2$ male was B/B and the probability that it was B/S but the *M. spretus* allele was not fixed after inbreeding. Because $N_2$ males were crossed again to B6 females before inbreeding starts, $P = 0.5 + 0.5 \times (0.5 + 0.5 \times 0.5) = 0.875$. If all strains were derived independently ($N_i = 1$ and $K = N$), then $P = (0.5 + 0.5 \times 0.75)^K = 0.875^K$. If all strains are not independent, the probability that all the $N_i$ strains derived from male $i$ are B/B is $P = 0.5 + 0.5 \times 0.75^{N_i}$. If inbreeding was started from N4 individuals (instead of N3), the value of $P$ for the strain would be $P = 0.5 + 0.5 \times 0.875 = 0.9375$. For a strain derived from N5, $P = 0.5 + 0.5 \times 0.9375 = 0.96875$. If $NA$, $NB$, and $NC$ strains were derived from N3, N4, and N5 progeny of the same backcross male, respectively, the probability that all $NA + NB + NC$ strains are homozygous for the B6 allele at a given locus would then be $P = 0.5 + 0.5 \times (0.75)^{NA} \times (0.875)^{NB} \times (0.9375)^{NC}$. The expected fraction of the genome not covered by the set of IRCSs is the product of the $P$'s calculated for each contributing backcross male. Theoretically, deriving a single N3 strain from each of 18 $N_2$ males would provide 91% genome coverage. Ninety-six percent genome coverage can be achieved with either three N4 or six N5 strains from each of 18 $N_2$ males.

Consomic strains were developed both at the Institut Pasteur and at the Universidad de la Laguna, Teneriffe (Spain) by backcrossing five or six times SEG/Pas to B6 mice. Chromosome 19 consomic was produced using the STF/Pas strain, another unrelated *M. spretus* moderately inbred strain. At every generation, mice carrying the chromosome of interest were selected by genotyping four to six microsatellites evenly spread along the chromosome.

**Strain nomenclature:** The set of IRCSs was named BcG. Each strain was named after the number of the fertile $N_2$ male it was derived from, followed by a letter indicating the strain order. For example BcG-122C was the third strain derived from the 122 $N_2$ male. However, for simplicity, it is often referred to as 122C. Consomic strains were also named BcG followed by the number of the chromosome (*e.g.*, BcG-14 for chromosome 14), with the exception of strain BcF-19 in which chromosome 19 was introgressed from STF/Pas.

**Genotyping:** IRCSs were genotyped three times at different stages of inbreeding. Forty-six strains were genotyped for 95 microsatellites at an average $F = 8.7$. Fifty strains were later genotyped for 183 microsatellites at an average $F = 24.6$. All 55 IRCSs were finally genotyped for 490 SNPs and 183 microsatellites at an average $F = 45.1$. The list of markers is available from the website mentioned below. To identify all alleles present in every strain, genotyping was performed on DNA pools prepared from tail clips collected from all breeding pairs or trios (8–12 mice per strain). The presence of the two parental alleles in a given sample (scored as "heterozygous") reflected the fact that they were still segregating in the strain.

Primers for microsatellite amplification were purchased either from Research Genetics (Huntsville, AL) or from Genset (Evry, France). Amplification was performed according to the manufacturer's instructions (35 cycles with 55° or 52° annealing temperature and 1.5–2 mM $MgCl_2$). For each marker, controls included DNAs from B6, SEG/Pas, and 1:1 and 1:20 mixtures of B6 and SEG/Pas, to verify that one B/S heterozygous mouse could be detected among a group of 10 B/B homozygous mice. PCR products were resolved on 4% Nu-Sieve agarose gels stained with ethidium bromide. SNP genotyping was performed at the Centre National de Génotypage (Evry, France) with the Illumina platform.

**Hematology:** All measurements of blood parameters were made on 60 ± 5-day-old males. Blood was collected by puncture of the orbital sinus and EDTA was immediately added to prevent clotting. Complete blood counts were determined using a Vet'ABC counter (SCIL, Viernheim, Germany).

**Analysis:** Statistical analysis was performed using StatView F-5.0 (SAS Institute, Cary, NC). Groups were compared by one-way ANOVA after checking distributions of values for normality by comparison with a normal distribution with the same mean and standard deviation using a Kolmogorov–Smirnov test. QTL analysis in the $F_2$ cross was performed using R-QTL (BROMAN *et al.* 2003).

A computer program was developed in Turbo-Pascal (Borland, version 7.0) to analyze pedigree data and genotypes, to generate the maps and diagrams shown in this article, and to create html files. All maps, genotypes, and statistics are

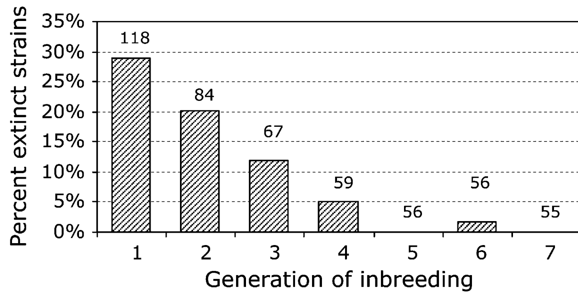


FIGURE 2.—Percentage of strain extinction during the first generations of inbreeding. A large number of strains did not survive the first four generations of inbreeding. Of 118 breeding pairs, 34 did not yield any progeny. The rate of extinction decreased steadily and the number of strains (given at the top of the bars) stabilized after $F_4$, as the result of both the elimination of very deleterious allelic combinations and the implementation of additional backcrosses to B6 that were able to rescue most infertile strains.

publicly available at http://www.pasteur.fr/recherche/unites/Gfons/ircs/ircshome.htm.

## RESULTS

**Establishment of IRCSs:** All $F_1$ and most backcross males derived from a cross between SEG/Pas and B6 mice are sterile due to hybrid-sterility genes (HALDANE 1922). To avoid the deleterious effect of these genes in further generations, the production of IRCSs was started from fertile backcross males. Of 157 $N_2$ males, 18 (11%, approximately one-eighth) proved to be fertile, consistent with the previous report of three genetically independent hybrid sterility genes ( GUÉNET *et al.* 1990; PILDER 1997; ELLIOTT *et al.* 2001).

To achieve the highest possible genome coverage, several strains were initiated from the progeny of each fertile $N_2$ male (Figure 1). We estimated that, with an average of three strains for each of the 18 males, <1% of the genome would be missed (except for regions containing hybrid sterility genes). A total of 118 strains were started, with 1 (male 49) to 9 (male 122) strains per $N_2$ male (average 6.6).

Breeding performances were very poor during the first generations of inbreeding (Figure 2). To avoid even more severe losses, one (and sometimes two) additional backcrosses to B6 were performed. Forty-eight strains underwent one additional backcross. Among those, 14 were unsuccessful due to male sterility. For 6 other strains, a second backcross was necessary and allowed to rescue the strain. Interestingly, 10 strains required an additional backcross at a late stage of inbreeding ($F>10$) due to poor breeding, 4 of which were between $F_{31}$ and $F_{53}$. In total, 53% of the strains were lost during inbreeding, with large variations across males. The collection of IRCSs currently consists of 55 strains (Figure 3).

Because of poor breeding performance, strict brother–sister mating was sometimes impossible and, in such cases, crosses involved males and females originating from different cages of the strain. This relaxed inbreeding resulted in delayed progression to homozygosity, so that we considered that inbreeding would be achieved after ~40 generations. The number of inbreeding generations varies between 22 and 83 (average 52). Forty-six IRCSs have passed $F_{40}$.

**Establishment of ICSs:** The breeding of ICSs was also very difficult and most strains were lost sooner or later during backcrosses. For many chromosomes, mice could not be bred past the second or third generation. Even for the remaining strains, it was often necessary to

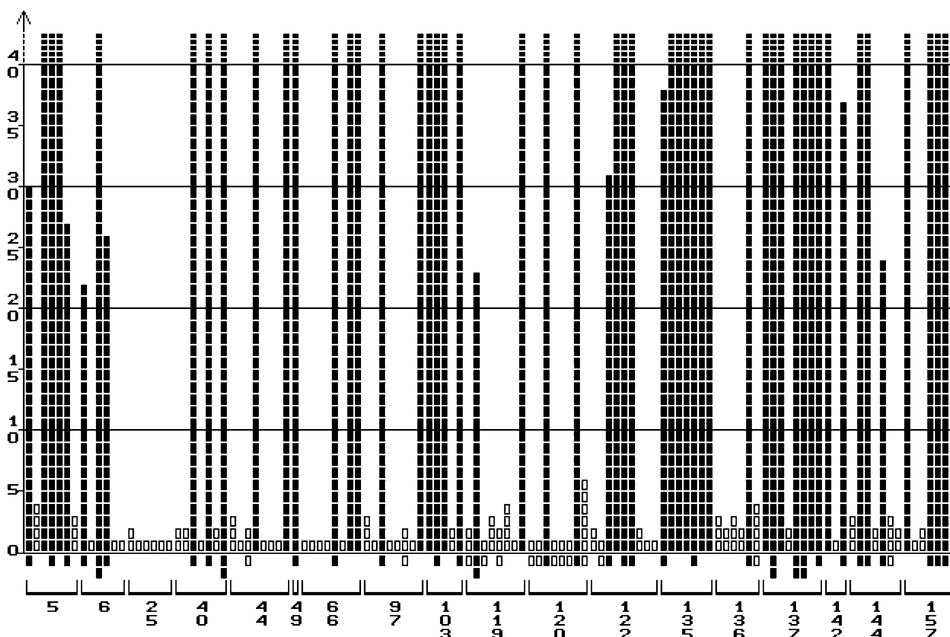


FIGURE 3.—Schematic of the current status of IRCSs. Each vertical bar with solid boxes represents one alive IRCS, while open boxes represent extinct strains. Strains derived from the same backcross male are grouped within a horizontal square bracket with the number of the backcross male at the bottom. Each box of a vertical bar represents one generation of inbreeding, up to 40. For extinct strains, the number of boxes indicates the number of generations after which the strain stopped breeding. Boxes below the baseline show the additional backcrosses (one or two) that were performed to avoid strain extinction.

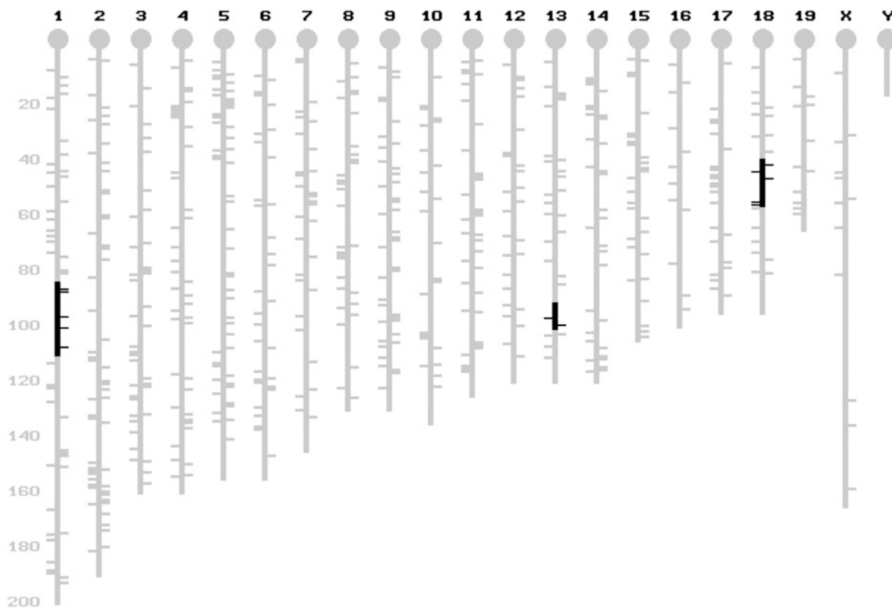


FIGURE 4.—Genetic map of IRCS 66H. IRCSs were genotyped for a total of 673 microsatellite and SNP markers represented as tick marks along chromosomes, with position in megabases. Regions of B6 origin are shaded, while SEG segments are solid. Segment boundaries are set halfway between adjacent markers with different genotypes. Strain 66H carries SEG alleles on chromosomes 1, 13, and 18. An interactive version of the maps is available at http://www.pasteur.fr/recherche/unites/Gfons/ircs/ircshome.htm.

split the chromosome to get homozygous, viable, and fertile progeny. For chromosomes 6, 13, 14, 16, and 18, only part of the chromosome of SEG/Pas origin could be bred to homozygosity. However, the entire chromosome 19 from STF/Pas could be retained.

**Genotyping of IRCSs:** Genotyping at high density revealed that each strain carries a small number of chromosomal segments of *M. spretus* origin. Figure 4 shows the example of IRCS 66H with three segments on chromosomes 1, 13, and 18. The boundaries of each segment were set halfway between neighbor markers with B/B and S/S genotypes. The fraction of the genome covered by the entire set was estimated by adding up the chromosomal segments of *M. spretus* origin carried by each strain. For example, eight strains were found to have at least one *M. spretus* allele—either heterozygous or homozygous—on chromosome 10 (Figure 5). The resulting segments were calculated and superimposed to estimate that ~57.9% of this chromosome is covered with IRCSs. There was large heterogeneity in genome coverage among chromosomes. No SEG allele was found on chromosome X, probably as a consequence of the selection of fertile $N_2$ males that were hence B6 homozygous. The percentage of the SEG genome was <12% on chromosomes 5, 8, and 9. By contrast, it was ~90% for chromosomes 18, and 19, the two smallest chromosomes of the genome. It ranged from 28 to 73% for all other chromosomes.

Maps of the entire genome were established similarly for IRCSs and ICSs and are presented in Figure 6. An interactive version of these maps can be found on a dedicated website (http://www.pasteur.fr/recherche/unites/Gfons/ircs/ircshome.htm), which provides an easy access to genotyping data and selection of strains of interest for a particular genomic region. Overall, IRCSs and ICSs taken separately provide genome coverage of 39.7 and 18.4%, respectively, and 45.6% when combined, of which 42.9% was fixed as homozygous and 2.7% was segregating at the time of genotyping. On the basis of the actual genealogy of the strains, it was

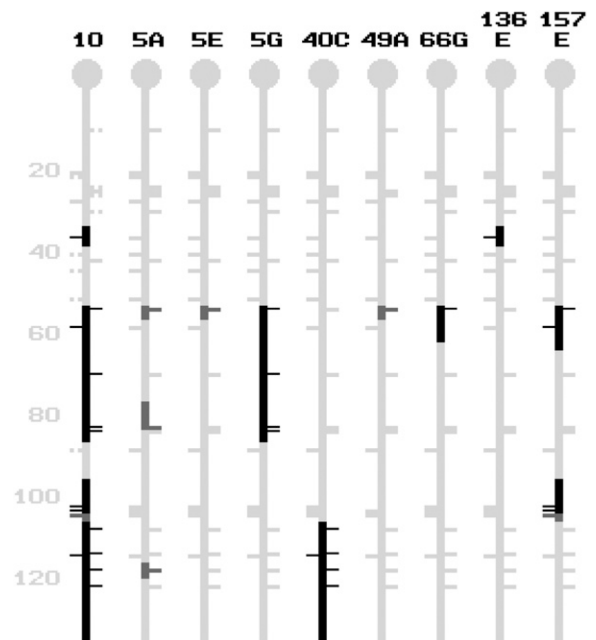


FIGURE 5.—Regions of chromosome 10 covered by IRCSs. All strains carrying a fragment of SEG origin on chromosome 10 are represented on the right, as in Figure 4. Regions segregating for B6 and SEG alleles are shown in dark shading. The leftmost chromosome combines the contribution of all IRCSs. Dashed lightly shaded tick marks represent markers with missing genotypes in one or more strains, hence for which there might be undetected SEG alleles. Segment boundaries are set halfway between adjacent markers with different genotypes. It is estimated that 57.9% of this chromosome is covered in the set of IRCSs.

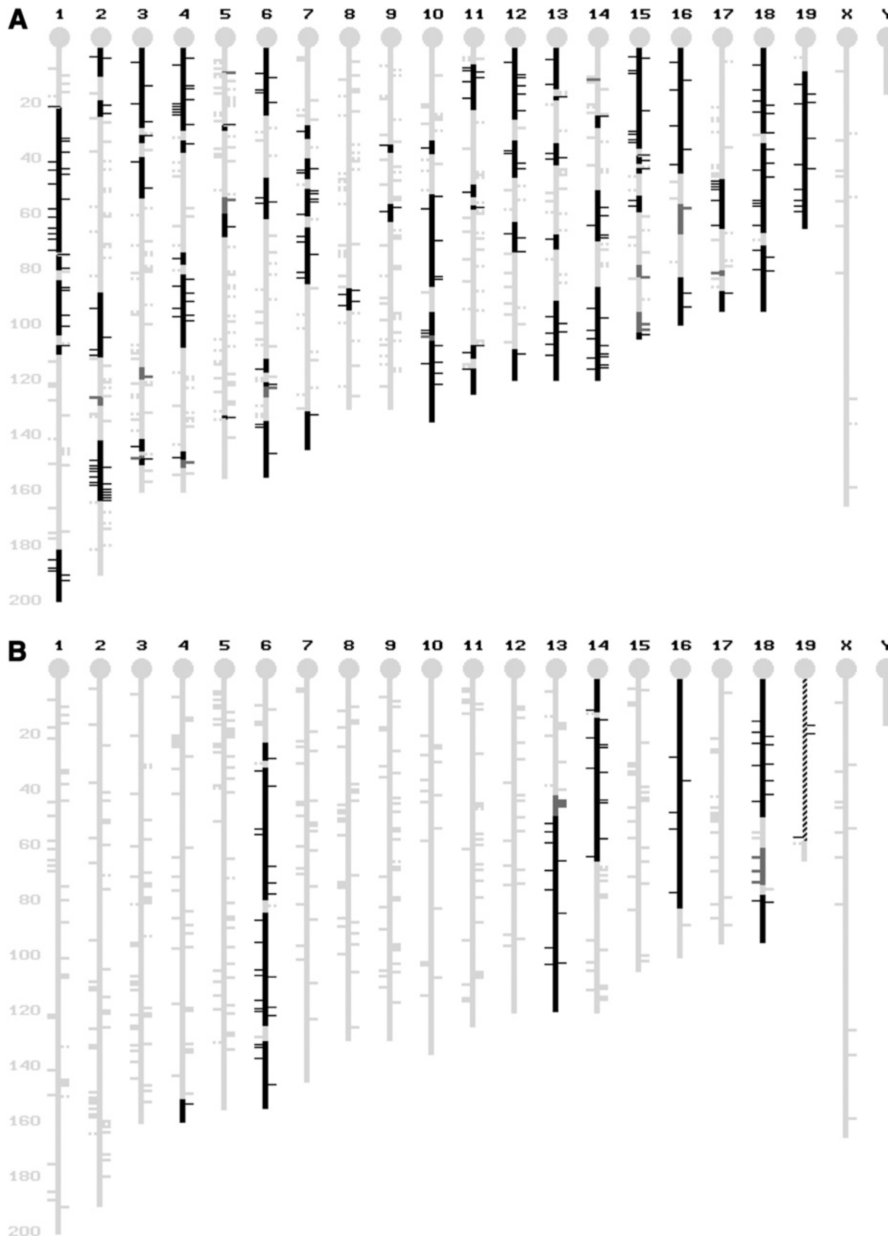


Figure 6.—Composite maps of IRCSs (A) and ICSs (B). Each genetic map summarizes the coverage provided by all IRCSs and ICSs. Solid segments depict regions for which at least one of the strains carries the SEG allele. Segments in dark shading depict regions for which at least one of the strains was segregating for the SEG allele at the time of genotyping. Segments are interrupted by missing B6 homozygous genotypes. Dashed lightly shaded tick marks represent markers with missing genotypes in one or more strains, hence for which there might be undetected SEG alleles. Segment boundaries are set halfway between adjacent markers with different genotypes. The hatched bar (B) illustrates the STF/Pas origin of the segment carried by the chromosome 19 consomic strain. At present, genome coverage is estimated to be 39.7% for IRCSs, 18.4% for ICSs, and 45.6% for both sets combined, of which 42.9% was fixed as homozygous and 2.7% was segregating at the time of genotyping.

expected to be 98% outside of regions containing hybrid sterility genes.

The average percentage of the *M. spretus* genome per strain was 1.37%, spanning from 0 to 3.79% (Figure 7A). Thirty-seven strains carry <1.5% of SEG alleles, and no *spretus* allele was found in 5 IRCSs (103B, 122E, 135D, 137H, and 142C). Because the set of IRCSs is a mixture of strains derived from three, four, or five backcrosses, the average percentage of the *M. spretus* genome was expected to be ~8.3%, six times higher than what was actually observed.

A total of 173 chromosomal segments of SEG origin were identified, resulting in an average of 3.15 and a maximum of 8 segments per strain (Figure 7B). Interestingly, 11 IRCSs carry only 1 segment, which makes them similar to classical congenic strains. Despite the genotyping of >670 markers, there are still several gaps in the map and some small-sized segments may have been missed, resulting in an underestimation of the total number of segments. Seventy-six segments are identified by only one marker. On the basis of the available data, the physical size of each segment (in megabases) was estimated to be 11.7 Mb on average, with a range from 0.5 to 40 Mb (Figure 8).

Since the first two series of genotypes were obtained on strains that were not at the same stage of inbreeding, the relationship between the number of generations of inbreeding and the percentage of SEG genome retained was investigated (Figure 9). Two phases emerged from this analysis. During the first generations of inbreeding, the percentage of the *spretus* genome regularly decreased by ~0.3% per generation and reached

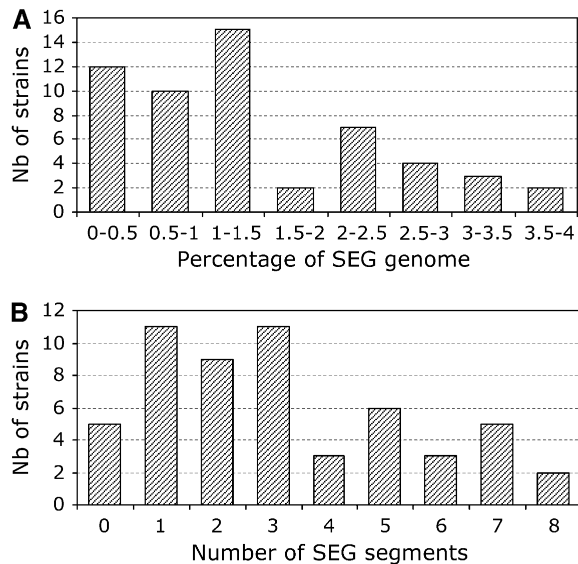


FIGURE 7.—Distribution of SEG contribution across IRCSs. (A) Distribution of the percentage of the SEG genome. For each strain, the percentage of the SEG genome was estimated as the sum of the length of all SEG segments over the total length of the genome (2670 Mb). (B) Distribution of the number of SEG segments in each strain. No SEG segment was identified in 5 strains, while 12 strains carry only one segment and are hence congenic strains (see online maps at http://www.pasteur.fr/recherche/unites/Gfons/ircs/ircshome.htm).

1.5% at $F_{15}$. It remained more or less stable in further generations. This evolution suggests that there was a continuous selection against SEG alleles during the inbreeding process, as long as chromosomal segments were still heterozygous.

**Analysis of quantitative traits using IRCSs:** To evaluate the power of IRCSs for the genetic dissection of quantitative traits, we analyzed a subset of them for seven hematological parameters, measured on 9-week-old males. Fifteen to 31 mice from 16 IRCSs and 14 SEG mice were compared with 28 B6 mice for red blood cell parameters, as well as platelet and white blood cell

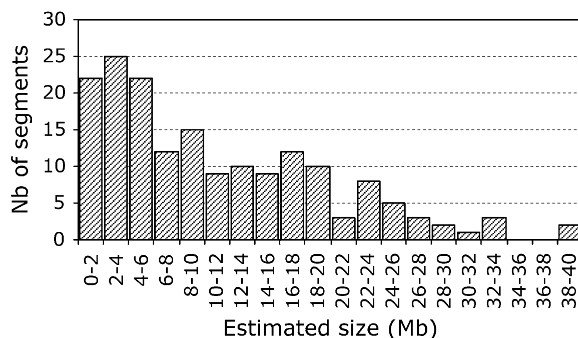


FIGURE 8.—Distribution of the size of SEG segments in IRCSs. Eighty-four percent (146/173) of segments of SEG origin are <20 Mb in size (∼10 cM), and the average size is 11.7 Mb, similar to what can be expected in regular congenic strains after 25 generations of backcrossing.

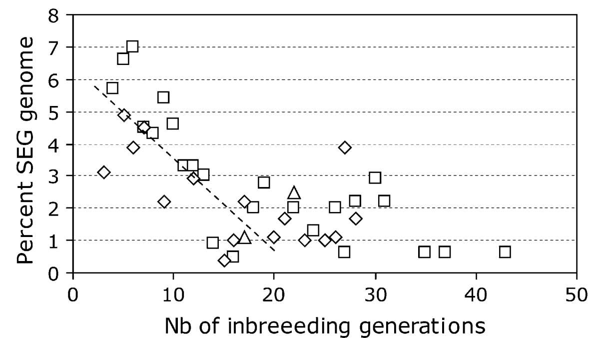


FIGURE 9.—Relationship between the percentage of the SEG genome and the number of generations of inbreeding at which IRCSs were genotyped. Data from the first two genotyping series (see text) are combined. For each strain, the percentage of the SEG genome was estimated as the number of SEG alleles over the total number of genotypes obtained. Symbols reflect the number of backcrosses to B6 performed before inbreeding: squares, N3; diamonds, N4; triangles, N5. The percentage of the SEG genome decreases steadily during the first 20 generations of inbreeding and remains between 1 and 3% in more advanced strains. The dashed line is the regression line calculated from strains between $F_2$ and $F_{20}$ ($y = 6.43 - 0.28x$; $r^2 = 0.6$; $P < 0.0001$).

counts (Table 1 and supplemental Table 1 at http://www.genetics.org/supplemental/). Four strains did not differ from B6 (120C, 135B, 135E, and 157F), while two strains (49A and 137F) showed differences for four of the seven parameters, including hematocrit, which was not statistically different between B6 and SEG. Of 112 comparisons, 21 differences were significant at the 0.05 threshold, 8 of which remained significant after Bonferroni correction for multiple testing (threshold = 0.00045).

Strain 66H showed a mean cell volume (MCV) intermediate between that of B6 and SEG, and platelet counts (Plt) higher than that of both parental strains, with a high level of significance (Table 1). This strain carries three SEG segments on chromosomes (Chr) 1, 13, and 18 (Figure 4). An $F_2$ generation was produced between 66H and B6. Seventy-six males were analyzed under the same conditions as the parental strains and genotyped for microsatellites marking the three genomic regions. However, no significant association could be observed between either traits or the markers.

The three SEG segments were isolated in congenic strains produced in two generations from an intercross between 66H and B6. These congenic strains were then intercrossed to combine two segments in the same strain. However, the Chr 1 + 13 combination was more difficult to breed and could not be included in the analysis for lack of animals. Congenic and bicongenic strains were evaluated for MCV and Plt. For both traits, the phenotypic value of the bicongenic strains could not be anticipated from those of the congenic strains. Indeed, regarding MCV, Chr 18 was similar to B6, while Chr 1 and Chr 13 resembled 66H (Figure 10A and Table 2). However, Chr 1 + 18 was not different from B6, and

**TABLE 1**

**Comparison of SEG/Pas and 16 IRCSs with C57BL/6 ($N = 28$ mice) for seven hematological parameters**

| Strain | No. mice | RBC | Hb | Hct | MCV | MCCH | Plt | WBC |
|---|---|---|---|---|---|---|---|---|
| SEG | 14 | 0.094<br>B6 < SEG | NS | NS | $<10^{-6}$<br>SEG < B6 | 0.0015<br>SEG < B6 | 0.0013<br>SEG < B6 | 0.0019<br>SEG < B6 |
| 5A | 24 | NS | NS | NS | 0.00027<br>SEG < B6 < 5A | NS | NS | NS |
| 6A | 29 | NS | NS | NS | NS | NS | $<10^{-6}$<br>SEG < B6 < 6A | NS |
| 6C | 25 | NS | NS | NS | NS | NS | 7.1 $10^{-6}$<br>SEG < B6 < 6C | NS |
| 49A | 20 | 0.00025<br>B6 < 49A < SEG | NS | 0.039<br>SEG < B6 < 49A | $<10^{-6}$<br>SEG < 49A < B6 | NS | 0.046<br>SEG < B6 < 49A | NS |
| 66H | 20 | NS | NS | NS | 0.015<br>SEG < 66H < B6 | NS | $<10^{-6}$<br>SEG < B6 < 66H | NS |
| 119H | 23 | NS | NS | NS | NS | NS | 0.0016<br>SEG < B6 < 119H | NS |
| 120C | 24 | NS | NS | NS | NS | NS | NS | NS |
| 122C | 22 | NS | NS | NS | NS | NS | 0.0032<br>SEG < B6 < 122C | NS |
| 122D | 17 | NS | NS | NS | $<10^{-6}$<br>SEG < B6 < 122D | NS | 0.0045<br>SEG < B6 < 122D | NS |
| 122F | 15 | NS | NS | NS | NS | NS | 0.0017<br>SEG < B6 < 122F | 0.015<br>SEG < 122F < B6 |
| 135B | 18 | NS | NS | NS | NS | NS | NS | NS |
| 135E | 22 | NS | NS | NS | NS | NS | NS | NS |
| 137E | 16 | NS | NS | NS | 0.014<br>SEG < B6 < 137E | NS | NS | NS |
| 137F | 25 | 0.0005<br>B6 < 137F < SEG | NS | 0.026<br>SEG < B6 < 137F | 1.8 $10^{-6}$<br>SEG < 137F < B6 | NS | 0.0036<br>SEG < B6 < 137F | NS |
| 137G | 23 | NS | NS | NS | NS | NS | 0.013<br>SEG < B6 < 137G | NS |
| 157F | 31 | NS | NS | NS | NS | NS | NS | NS |

*P*-value of the comparison between B6 and the strain (NS, $P > 0.05$), with the relative order of the means.

Chr 13 + 18 had a MCV even higher than that of Chr 18 alone. This suggested that Chr 1 and Chr 13 were mostly responsible for the lower MCV observed in 66H, but their individual effect might be abolished in the presence of Chr 18.

Regarding Plt, Chr 13 alone did not seem to have any effect, while Chr 1 and Chr 18 had intermediate values between those of B6 and 66H (Figure 10B and Table 3). Chr 18 was not affected by the presence of Chr 13 in the Chr 13 + 18 bicongenic, and it seemed that Chr 1 and Chr 18 were able to add up their individual effects in the Chr 1 + 18 bicongenic, although this observation remains preliminary because of small sample size.

These results show MCV and Plt are controlled by shared genetic regions. They also illustrate the importance of epistasis in the genetic control of these traits. The effect of each region, alone or in combination, is different between the two traits, which suggests that MCV and Plt are controlled by linked but distinct genes.

## DISCUSSION

Genetic analysis of complex traits poses special challenges that have hampered so far the characterization of the genes themselves, beyond the identification of QTL-containing regions (Churchill *et al.* 2004; Flint *et al.* 2005). The development of new resources including large reference populations, and the massive accumulation of phenotypic, high-density genotyping, and gene expression data in tens of inbred strains will certainly represent a major instrument toward this goal (Wang *et al.* 2003).

The development of a set of recombinant congenic strains with an interspecific combination of parents was aimed at providing a tool complementary to other, existing resources. Recently, two sets of consomic strains have been successfully developed with the contribution of either *M. musculus musculus* [strain PWD/Ph (Jansa *et al.* 2005)] or *M. musculus molossinus* [strain MSM (Oka *et al.* 2007)] wild-derived inbred strains. Crosses with *M. spretus* have the deserved reputation of being poorly productive, which was the case in the present experiment. Many IRCSs failed to breed very early during the inbreeding process, most likely because mice carried allelic combinations that resulted in reduced fitness, failure to thrive, diseases, or sterility. Since the two species have been separated for >1.5 million years (Guénet and Bonhomme 2003), it is not surprising

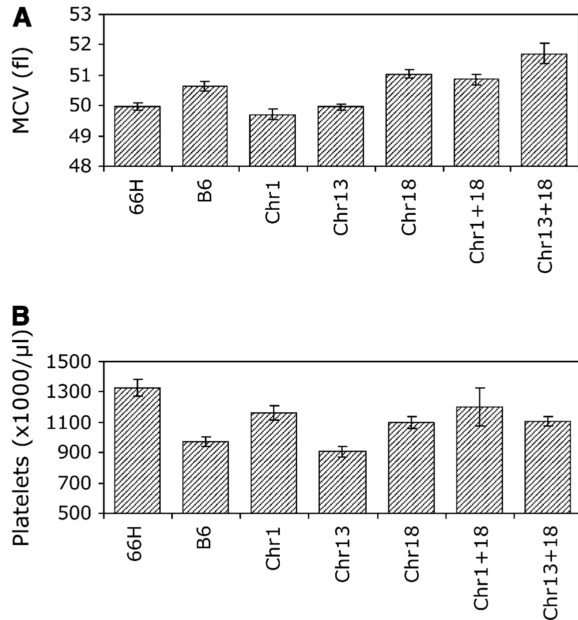


FIGURE 10.—Phenotypes of B6, IRCS 66H, congenic and bicongenic strains for mean cell volume (A) and platelet counts (B). Bars indicate the mean value (± SEM) measured in groups of age-matched males. Group size and statistical significance of comparisons are shown in Tables 2 and 3.

**TABLE 3**

**Comparison of C57BL/6, congenic, and bicongenic strains, for platelets**

| | No. mice | B6 | Chr 1 | Chr 13 | Chr 18 | Chr 1 + 18 |
|---|---|---|---|---|---|---|
| B6 | 28 | | | | | |
| Chr 1 | 20 | 0.002 | | | | |
| Chr 13 | 20 | NS | 0.0001 | | | |
| Chr 18 | 27 | 0.018 | NS | 0.0009 | | |
| Chr 1 + 18 | 6 | 0.02 | NS | 0.0046 | NS | |
| Chr 13 + 18 | 17 | 0.014 | NS | 0.0004 | NS | NS |

*P*-values are calculated from Student's *t*-test.

that genomic incompatibilities have accumulated between unlinked genes. Such epistatic interactions are still compatible with the production of an interspecific backcross, although the small litter size usually observed suggests that there is substantial embryonic lethality.

The strategy used for the development of IRCSs proved to be effective. The deleterious effect of hybrid sterility genes was discarded at the first backcross generation. During later generations, relaxed inbreeding imposed by the small number of progeny available to form subsequent breeding pairs has slowed down the progression to full homozygosity, hence providing more chance for unfavorable allelic combinations to be eliminated. To illustrate this point, let us consider an essential heterodimeric complex. It must be assembled from compatible subunits. The genes encoding the two components must have evolved in a concerted fashion so that these components remain compatible throughout evolution. However, when two divergent genomes are mixed and forced to reassort such as during the development of IRCSs, some zygotes will be B6 homozygous for a subunit and SEG homozygous for another one, which may result in the fatal inability to produce a functional complex. All other allelic combinations are compatible with survival because they allow the formation of a functional complex. When genomes become homozygous, both loci must fix alleles of the same origin. Because the SEG contribution to IRCSs was, from the beginning, seven times less than the B6 counterpart, it was much more likely that both loci become homozygous for the B6 allele. This is a reasonable explanation for the progressive loss of SEG alleles during the first generations of inbreeding.

By comparison, the development of consomic strains forced such essential epistatic interactions between unlinked genes to be disrupted, resulting in the loss of individuals carrying incompatible allelic combinations. As a consequence, only few genomic regions could be retained in consomic strains, especially in the smallest chromosomes (16, 18, and 19).

Deleterious epistatic interactions have been described in very few instances. The best documented example is probably the very strong and consistent transmission ratio distortion observed on chromosome X in [C57BL/6 × *M. spretus*] × *M. spretus* crosses. It was demonstrated that the large excess of *spretus* homozygotes observed for loci located on central chromosome X could be due to the loss of heterozygotes that were not simultaneously heterozygous in a distal region of chromosome 2 (MONTAGUTELLI *et al.* 1996). While such a dramatic situation remains the only observation in interspecific backcrosses and may be rare, interactions with weaker effects can also have a significant impact in combination with others over several generations.

More than half of the strains were lost during the inbreeding phase, mostly, but not only, during the early generations. Notably, four strains stopped breeding

**TABLE 2**

**Comparison of C57BL/6, congenic, and bicongenic strains, for MCV**

| | No. mice | B6 | Chr 1 | Chr 13 | Chr 18 | Chr 1 + 18 |
|---|---|---|---|---|---|---|
| B6 | 28 | | | | | |
| Chr 1 | 20 | 0.0005 | | | | |
| Chr 13 | 20 | 0.0035 | NS | | | |
| Chr 18 | 27 | NS | <0.0001 | <0.0001 | | |
| Chr 1 + 18 | 6 | NS | 0.0018 | 0.0004 | NS | |
| Chr 13 + 18 | 17 | 0.0027 | <0.0001 | <0.0001 | 0.04 | NS |

*P*-values are calculated from Student's *t*-test.

after $F_{30}$ and required an additional backcross. Strain extinction is a common observation during the establishment of recombinant inbred strains, even after full inbreeding (B. A. Taylor, personal communication). Rate of extinction was variable between $N_2$ males, which could reflect differences in the heterozygous regions present in each male. However, no material was collected on these $N_2$ males to perform genotyping.

The average rate of SEG genome in IRCSs is six times less than expected based on the breeding records and results from a progressive loss of *M. spretus* alleles occurring during the first 15 generations of inbreeding, when the genome was still largely heterozygous. We observed that in 87% of cases, strains segregating for the two alleles at a particular locus in the first series of genotyping (average $F = 8.7$) had fixed the B6 allele in the second series (average $F = 24.6$). No such genome-wide selection has been observed in recombinant inbred or recombinant congenic strains established from laboratory strains, which showed the expected proportion from the two parental strains (Groot *et al.* 1992; Stassen *et al.* 1996). The final genetic makeup of IRCSs and ICSs suggests that it is hardly feasible to introduce >4–5% of the *M. spretus* genome in a B6 background, because of the large number of interchromosomal epistatic interactions.

The strong reduction of SEG alleles resulted also in that strains derived from the same backcross male were as genetically different as strains derived from independent pedigrees, unlike intraspecific recombinant congenic strains, where strains derived from the same backcross male share on average half of the contribution from the donor strain.

The proportion of the SEG genome in IRCSs was variable among strains. For five strains, we could not detect any SEG allele, which could be due to either an extreme selection against the *M. spretus* genome, or to the few gaps that still remain between genotyped markers. Since a number of fragments identified are very small in size, it is possible that a few of them were missed. However, it is unlikely that genotyping at higher density will modify the global rate of the SEG genome.

Altogether, this set of IRCSs covers 39.7% of the mouse genome. This rate is highly variable between chromosomes and varies from 0 to >90%. Part of this variation may be explained by the early elimination of regions containing hybrid sterility genes (such as for chromosomes X and 9). The two smallest chromosomes were best covered. The variation of genome coverage between chromosomes cannot be explained by their difference in length, or in gene content. Notably, SEG allelic frequency also varied along chromosomes, with the highest frequency observed almost always for loci located to the proximal of distal quarters of chromosome (data not shown). In particular, the SEG allele was found to be 13.7% at marker *D6Mit201* on distal chromosome 6. To assess whether this higher than expected frequency could be due to transmission ratio distortion favoring the SEG allele, we set up an intercross between IRCS 137C carrying this allele and B6. Distribution of genotypes did not depart from that expected (data not shown), and hence did not provide support for a selective mechanism, although the $F_2$ population ($N = 102$) might have not been large enough to detect weak deviation. Further investigation will be needed to understand the basis for variations in SEG allele frequencies along the genome.

The contribution of *M. spretus* was clustered, in each IRCS, in a small number of small-sized chromosomal segments. These segments were much smaller than those found in classical RCSs (Groot *et al.* 1992; Stassen *et al.* 1996), due to the combination of relaxed inbreeding and counterselection of *M. spretus* alleles. Twelve strains are even congenic strains in that they carry a single SEG fragment. The small number of fragments has two major advantages. First, if a phenotypic difference is observed between B6 and a given strain, its genetic control is readily attributable to a few regions (or even one is some instances). Second, the role of each fragment can be studied in an $F_2$ between this IRCS and B6, where only 1.3% of the genome segregates, on average, which increases the power to detect epistasis. Alternatively, congenic strains carrying each of the segments can be derived directly from $F_2$ progeny, allowing phenotypic studies on a series of genetically identical, sex- and age-matched individuals. Interactions between QTL can be studied by combining the segments in bi- or multiple congenic strains.

Inbred strains that are genetically highly divergent are more likely to show phenotypic differences for any trait studied than pairs of laboratory strains. However, interspecific crosses have been criticized as possibly involving too many QTL for each of them to be detectable. IRCSs still capture a high rate of polymorphism but it is restricted to small-sized regions. With their reduced proportion of the *M. spretus* genome, it is likely that phenotypic differences are controlled by a small number of QTL that are amenable to genetic dissection. The high density of polymorphic markers in the *M. spretus* chromosomal segments will facilitate fine genetic analysis and positional cloning of QTL. Finally, because the genomes of the two parental strains have diverged for a long time, it is likely that new phenotypes can emerge from the disruption of coadapted allelic combinations.

The SEG strain was only partially inbred at the time of the first cross with B6. One cannot exclude that different IRCSs carrying apparently the same chromosomal region from SEG could actually carry different *M. spretus* segments with some sequence variation, which could confound the interpretation of results. However, high density SNP data have revealed that SEG and STF strains, which originate from Spain and Tunisia, respectively, differ for only 8 of the 660 SNP analyzed (1.2%, data not shown). It is reasonable to assume that

genetic polymorphism present within the original SEG colony was even smaller.

QTL mapping using IRCSs should not be based on the comparison of phenotypes between the parental strains B6 and SEG for at least three reasons. First, less than half of the potential QTL can be identified using IRCSs and ICSs due to partial genome coverage. Second, a difference between B6 and an IRCS may be observed even in the case where B6 and SEG show similar phenotypes. Third, the direction of phenotypic differences between B6 and any IRCS cannot be predicted from the direction observed between B6 and SEG. These last two points have already been reported with RCSs and RISs between laboratory strains (Van Wezel *et al.* 1996; Grisel *et al.* 1997). IRCSs are best used by assessing the phenotypic trait of interest in all strains of the set by comparison with B6. Any phenotypic difference observed must be controlled by one or more of the chromosomal segments carried by the IRCS which can then be submitted to genetic analysis.

To evaluate the potential of IRCSs, we investigated seven hematological traits on 16 IRCSs and found eight significant differences in six strains, after Bonferroni correction for multiple testing. This high rate of phenotypic variation between strains likely reflects both the polygenic control of the traits and the high genetic polymorphism of the B6–SEG cross. Similar observations have been made on other traits, related to male fertility (L'Hôte *et al.* 2007) and skull morphology (G. Burgio and X. Montagutelli, unpublished results). In all cases, it is surprising to note that phenotypes distinct from that of B6 are found despite the low percentage of the SEG genome.

Strain 66H was found to differ from B6 for MCV and Plt. None of the three chromosomal regions known to be of SEG origin in 66H showed a significant association with either trait in the $F_2$ cross between 66H and B6. This could be explained by the small size of the $F_2$ cross ($N = 76$), the dominant or recessive mode of inheritance of the SEG alleles, the weak effect of each QTL, and the complex genetic interactions between the three regions. This was confirmed by the analysis of the three congenic and two of the three bicongenic strains. The effect of each region was weak, but significant. Statistically significant epistatic interactions were found and the phenotype of the bicongenic strains could not be deduced from that of the single congenic strains. For MCV, both Chr 1 and Chr 13 congenic strains appeared to reproduce the phenotype of 66H. Every strain that carried the chromosome 18 segments had a significantly higher MCV than B6, while both chromosome 1 and chromosome 13 congenic strains appeared to reproduce the phenotype of 66H. QTL for MCV have been reported on chromosomes 1 and 13 by Valdar *et al.* (2006) using heterogeneous stocks. However, their confidence intervals do not overlap the SEG regions of strain 66H and it is therefore likely that they are not identical to the QTL identified in the present study. Fine mapping of these QTL is under progress and involves the production of subcongenic strains, taking advantage of the high density of polymorphic markers.

None of the congenic or bicongenic strains had a platelet count as high as that of 66H. Chr 1 and 18, separately or combined in the bicongenic strain, conferred a significant increase compared with B6, and are therefore potential targets for QTL identification. We were not able so far to combine the three regions in a single tricongenic strain to reproduce the phenotype of 66H, as exemplified in the work of Morel *et al.* (2000). We cannot rule out that 66H carries a small SEG fragment that has remained so far undetected and that would explain the difference between the chromosome 1 + 18 bicongenic strain and 66H. QTL for Plt have been described on chromosomes 1 and 18 by Valdar *et al.* (2006) and by Cheung *et al.* (2004) but in regions absent from 66H.

In the present work, we have demonstrated the feasibility and usefulness of interspecific recombinant congenic strains for the genetic evaluation of complex traits, even for QTL with small individual effects and complex epistatic interactions. In addition to being a unique and valuable genetic resource, IRCSs provide an experimental example of genomic incompatibilities that develop when two lineages become separated by more than one million years of evolution.

We are grateful to Isabelle Lanctin for highly dedicated and careful breeding of the IRCSs and to Stéphanie Voegeling and Murielle Rocancourt for expert assistance with microsatellite genotyping. We thank the Centre National de Génotypage for SNP genotyping.

## LITERATURE CITED

Abiola, O., J. M. Angel, P. Avner, A. A. Bachmanov, J. K. Belknap *et al.*, 2003 The nature and identification of quantitative trait loci: a community's view. Nat. Rev. Genet. **4:** 911–916.

Avner, P., L. Amar, L. Dandolo and J.-L. Guénét, 1988 Genetic analysis of the mouse using interspecific crosses. Trends Genet. **4:** 18–23.

Bagot, S., S. Campino, C. Penha-Goncalves, S. Pied, P. A. Cazenave *et al.*, 2002 Identification of two cerebral malaria resistance loci using an inbred wild-derived mouse strain. Proc. Natl. Acad. Sci. USA **99:** 9919–9923.

Bailey, D. W., 1971 Recombinant-inbred strains. An aid to identify linkage and function of histocompatibility and other genes. Transplantation **11:** 325–327.

Bailey, D. W., 1981 Recombinant inbred strains and bilineal congenic strains, pp. 223–239 in *The Mouse in Biomedical Research*, edited by H. L. Foster, J. D. Small and J. G. Fox. Academic Press, New York.

Banus, H. A., H. J. van Kranen, F. R. Mooi, B. Hoebee, N. J. Nagelkerke *et al.*, 2005 Genetic control of Bordetella pertussis infection: identification of susceptibility loci using recombinant congenic strains of mice. Infect. Immun. **73:** 741–747.

Breen, M., L. Deakln, B. Macdonald, S. Miller, R. Sibson *et al.*, 1994 Towards high resolution maps of the mouse and human genome—a facility for ordering markers to 0.1 cM resolution. Hum. Mol. Genet. **3:** 621–627.

Broman, K. W., H. Wu, S. Sen and G. A. Churchill, 2003 R/qtl: QTL mapping in experimental crosses. Bioinformatics **19:** 889–890.

Cheung, C. C., I. C. Martin, K. R. Zenger, J. A. Donald, P. C. Thomson *et al.*, 2004 Quantitative trait loci for steady-state platelet count in mice. Mamm. Genome **15:** 784–797.

Churchill, G. A., D. C. Airey, H. Allayee, J. M. Angel, A. D. Attie *et al.*, 2004 The Collaborative Cross, a community resource for the genetic analysis of complex traits. Nat. Genet. **36:** 1133–1137.

Darvasi, A., A. Weinreb, V. Minke, J. I. Weller and M. Soller, 1993 Detecting marker-QTL linkage and estimating QTL gene effect and map location using a saturated genetic map. Genetics **134:** 943–951.

Davis, R. C., E. E. Schadt, D. J. Smith, E. W. Hsieh, A. C. Cervino *et al.*, 2005 A genome-wide set of congenic mouse strains derived from DBA/2J on a C57BL/6J background. Genomics **86:** 259–270.

Demant, P., and A. A. M. Hart, 1986 Recombinant congenic strains - a new tool for analyzing genetic traits determined by more than one gene. Immunogenetics **24:** 416–422.

Elliott, R. W., D. R. Miller, R. S. Pearsall, C. Hohman, Y. Zhang *et al.*, 2001 Genetic analysis of testis weight and fertility in an interspecies hybrid congenic strain for chromosome X. Mamm. Genome **12:** 45–51.

Flint, J., W. Valdar, S. Shifman and R. Mott, 2005 Strategies for mapping and cloning quantitative trait genes in rodents. Nat. Rev. Genet. **6:** 271–286.

Fortin, A., L. R. Cardon, M. Tam, E. Skamene, M. M. Stevenson *et al.*, 2001a Identification of a new malaria susceptibility locus (Char4) in recombinant congenic strains of mice. Proc. Natl. Acad. Sci. USA **98:** 10793–10798.

Fortin, A., E. Diez, D. Rochefort, L. Laroche, D. Malo *et al.*, 2001b Recombinant congenic strains derived from A/J and C57BL/6J: a tool for genetic dissection of complex traits. Genomics **74:** 21–35.

Gouya, L., F. Muzeau, A. M. Robreau, P. Letteron, E. Couchi *et al.*, 2007 Genetic study of variation in normal mouse iron homeostasis reveals ceruloplasmin as an HFE-hemochromatosis modifier gene. Gastroenterology **132:** 679–686.

Grisel, J. E., J. K. Belknap, L. A. O'Toole, M. L. Helms, C. D. Wenger *et al.*, 1997 Quantitative trait loci affecting methamphetamine responses in BXD recombinant inbred mouse strains. J. Neurosci. **17:** 745–754.

Groot, P. C., C. J. Moen, W. Dietrich, J. P. Stoye, E. S. Lander *et al.*, 1992 The recombinant congenic strains for analysis of multigenic traits: genetic composition. FASEB J. **6:** 2826–2835.

Guénet, J. L., and F. Bonhomme, 2003 Wild mice: an ever-increasing contribution to a popular mammalian model. Trends Genet. **19:** 24–31.

Guénet, J. L., and X. Montagutelli, 1994 The contribution of wild specimens to the establishment of the mouse genetic map, pp. 285–298 in *Genetics in Wild Mice*, edited by K. Moriwaki and R. Kominami. Japan Scientific Society Press, Tokyo.

Guénet, J. L., C. Nagamine, D. Simon-Chazottes, X. Montagutelli and F. Bonhomme, 1990 *Hst-3*: an X-linked hybrid sterility gene. Genet. Res. **56:** 163–165.

Haldane, J. B. S., 1922 Sex ratio and unisexual sterility in hybrid animals. J. Genet. **12:** 101–109.

Iakoubova, O. A., C. L. Olsson, K. M. Dains, D. A. Ross, A. Andalibi *et al.*, 2001 Genome-tagged mice (GTM): two sets of genome-wide congenic strains. Genomics **74:** 89–104.

Ideraabdullah, F. Y., E. de la Casa-Esperon, T. A. Bell, D. A. Detwiler, T. Magnuson *et al.*, 2004 Genetic and haplotype diversity among wild-derived mouse inbred strains. Genome Res. **14:** 1880–1887.

Jansa, P., P. Divina and J. Forejt, 2005 Construction and characterization of a genomic BAC library for the Mus m. musculus mouse subspecies (PWD/Ph inbred strain). BMC Genomics **6:** 161.

Lee, P. D., B. Ge, C. M. Greenwood, D. Sinnett, Y. Fortin *et al.*, 2006 Mapping cis-acting regulatory variation in recombinant congenic strains. Physiol. Genomics **25:** 294–302.

Lemay, A. M., and C. K. Haston, 2005 Bleomycin-induced pulmonary fibrosis susceptibility genes in AcB/BcA recombinant congenic mice. Physiol. Genomics **23:** 54–61.

L'Hôte, D., C. Serres, P. Laissue, A. Oulmouden, C. Rogel-Gaillard *et al.*, 2007 Centimorgan-range one-step mapping of fertility traits using interspecific recombinant congenic mice. Genetics **176:** 1907–1921.

Martin, B., G. Chapouthier and R. Motta, 1992a Analysis of B10.D2 recombinant congenic mouse strains shows that audiogenic and beta-CCM-induced seizures depend on different genetic mechanisms. Epilepsia **33:** 11–13.

Martin, B., C. Marchaland, J. Phillips, G. Chapouthier, C. Spach *et al.*, 1992b Recombinant congenic strains of mice from B10.D2 and DBA/2: their contribution to behavior genetic research and application to audiogenic seizures. Behav. Genet. **22:** 685–701.

Mashimo, T., M. Lucas, D. Simon-Chazottes, M.-P. Frenkiel, X. Montagutelli *et al.*, 2002 A nonsense mutation in the gene encoding 2′-5′-oligoadenylate synthetase/L1 isoform is associated with West Nile virus susceptibility in laboratory mice. Proc. Natl. Acad. Sci. USA **99:** 11311–11316.

Melanitou, E., F. Joly, M. Lathrop, C. Boitard and P. Avner, 1998 Evidence for the presence of insulin-dependent diabetes-associated alleles on the distal part of mouse chromosome 6. Genome Res. **8:** 608–620.

Moen, C. J., P. C. Groot, A. A. Hart, M. Snoek and P. Demant, 1996 Fine mapping of colon tumor susceptibility (Scc) genes in the mouse, different from the genes known to be somatically mutated in colon cancer. Proc. Natl. Acad. Sci. USA **93:** 1082–1086.

Moen, C. J., M. Snoek, A. A. Hart and P. Demant, 1992 Scc-1, a novel colon cancer susceptibility gene in the mouse: linkage to CD44 (Ly-24, Pgp-1) on chromosome 2. Oncogene **7:** 563–566.

Moen, C. J., H. J. Stoffers, A. A. Hart, H. V. Westerhoff and P. Demant, 1997 Simulation of the distribution of parental strains' genomes in RC strains of mice. Mamm. Genome **8:** 884–889.

Moen, C. J., M. A. van der Walk, M. Snoek, B. F. van Zutphen, O. von Deimling *et al.*, 1991 The recombinant congenic strains - a novel genetic tool applied to the study of colon tumor development in the mouse. Mamm. Genome **1:** 217–227.

Montagutelli, X., R. Turner and J. H. Nadeau, 1996 Epistatic control of non-Mendelian inheritance in mouse interspecific crosses. Genetics **143:** 1739–1752.

Morel, L., B. P. Croker, K. R. Blenman, C. Mohan, G. Huang *et al.*, 2000 Genetic reconstitution of systemic lupus erythematosus immunopathology with polycongenic murine strains. Proc. Natl. Acad. Sci. USA **97:** 6670–6675.

Mott, R., and J. Flint, 2002 Simultaneous detection and fine mapping of quantitative trait loci in mice using heterogeneous stocks. Genetics **160:** 1609–1618.

Newman, T. L., E. Tuzun, V. A. Morrison, K. E. Hayden, M. Ventura *et al.*, 2005 A genome-wide survey of structural variation between human and chimpanzee. Genome Res. **15:** 1344–1356.

Oka, A., T. Aoto, Y. Totsuka, R. Takahashi, M. Ueda *et al.*, 2007 Disruption of genetic interaction between two autosomal regions and the X chromosome causes reproductive isolation between mouse strains derived from different subspecies. Genetics **175:** 185–197.

Pilder, S. H., 1997 Identification and linkage mapping of Hst7, a new M. spretus/M. m. domesticus chromosome 17 hybrid sterility locus. Mamm. Genome **8:** 290–291.

Rhodes, M., R. Straw, S. Fernando, E. Evans, T. Lacey *et al.*, 1998 A high resolution microsatellite map of the mouse genome. Genome Res. **8:** 531–542.

Roy, M. F., N. Riendeau, J. C. Loredo-Osti and D. Malo, 2006 Complexity in the host response to Salmonella Typhimurium infection in AcB and BcA recombinant congenic strains. Genes Immun. **7:** 655–666.

Sebastiani, G., V. Blais, V. Sancho, S. N. Vogel, M. M. Stevenson *et al.*, 2002 Host immune response to Salmonella enterica serovar Typhimurium infection in mice derived from wild strains. Infect. Immun. **70:** 1997–2009.

Serreze, D. V., M. Prochazka, P. C. Reifsnyder, M. M. Bridgett and E. H. Leiter, 1994 Use of recombinant congenic and congenic strains of NOD mice to identify a new insulin-dependant diabetes resistance gene. J. Exp. Med. **180:** 1553–1558.

Singer, J. B., A. E. Hill, L. C. Burrage, K. R. Olszens, J. Song *et al.*, 2004 Genetic dissection of complex traits with chromosome substitution strains of mice. Science **304:** 445–448.

Staelens, J., B. Wielockx, L. Puimege, F. Van Roy, J. L. Guénet *et al.*, 2002 Hyporesponsiveness of SPRET/Ei mice to lethal shock

induced by tumor necrosis factor and implications for a TNF-based antitumor therapy. Proc. Natl. Acad. Sci. USA **99:** 9340–9345.

Stassen, A. P., P. C. Groot, J. T. Eppig and P. Demant, 1996 Genetic composition of the recombinant congenic strains. Mamm. Genome **7:** 55–58.

Taylor, B. A., 1978 Recombinant inbred mice: use in gene mapping, pp. 423–438 in *Origin of Inbred Mice*, edited by H. C. Morse. Academic Press, New York.

Turcotte, K., J. C. Loredo-Osti, P. Fortin, E. Schurr, K. Morgan *et al.*, 2006 Complex genetic control of susceptibility to Mycobacterium bovis (Bacille Calmette-Guerin) infection in wild-derived Mus spretus mice. Genes Immun. **7:** 684–687.

Valdar, W., L. C. Solberg, D. Gauguier, S. Burnett, P. Klenerman *et al.*, 2006 Genome-wide genetic association of complex traits in heterogeneous stock mice. Nat. Genet. **38:** 879–887.

van Wezel, T., A. P. Stassen, C. J. Moen, A. A. Hart, M. A. van der Valk *et al.*, 1996 Gene interaction and single gene effects in colon tumour susceptibility in mice. Nat. Genet. **14:** 468–470.

Wang, J., R. W. Williams and K. F. Manly, 2003 WebQTL: web-based complex trait analysis. Neuroinformatics **1:** 299–308.

Yi, N., A. Diament, S. Chiu, K. Kim, D. B. Allison *et al.*, 2004 Characterization of epistasis influencing complex spontaneous obesity in the BSB model. Genetics **167:** 399–409.

Communicating editor: L. Siracusa