

Genetic structure of the LXS panel of recombinant inbred mouse strains: a powerful resource for complex trait analysis

Robert W. Williams,^{4*} Beth Bennett,^{2*} Lu Lu,¹ Jing Gu,¹ John C. DeFries,^{2,3} Phyllis J. Carosone-Link,² Brad A. Rikke,² John K. Belknap,⁴ Thomas E. Johnson^{2,3}

¹Center of Genomics and Bioinformatics, Institute of Neuroscience, Department of Anatomy and Neurobiology, University of Tennessee Health Science Center, Memphis, Tennessee 38163, USA

²Institute for Behavioral Genetics, CB 447, University of Colorado, Boulder, Colorado 80309, USA

³Department of Psychology, CB 345, University of Colorado, Boulder, Colorado 80309, USA

⁴Portland Alcohol Research Center, Oregon Health Sciences University, Portland, Oregon 97239, USA

Received: 11 February 2004 / Accepted: 9 April 2004

Abstract

The set of LXS recombinant inbred (RI) strains is a new and exceptionally large mapping panel that is suitable for the analysis of complex traits with comparatively high power. This panel consists of 77 strains—more than twice the size of other RI sets—and will typically provide sufficient statistical power ($\beta = 0.8$) to map quantitative trait loci (QTLs) that account for $\sim 25\%$ of genetic variance with a genome-wide $p < 0.05$. To characterize the genetic architecture of this new set of RI strains, we genotyped 330 MIT microsatellite markers distributed on all autosomes and the X Chromosome and assembled error-checked meiotic recombination maps that have an average F_2 -adjusted marker spacing of ~ 4 cM. The LXS panel has a genetic structure consistent with random segregation and subsequent fixation of alleles, the expected $3\text{--}4 \times$ map expansion, a low level of nonsyntenic association among loci, and complete independence among all 77 strains. Although the parental inbred strains—Inbred Long-Sleep (ILS) and Inbred Short-Sleep (ISS)—were derived originally by selection from an 8-way heterogeneous stock selected for differential sensitivity to sedative effects of ethanol, the LXS panel is also segregating for many other traits. Thus, the LXS panel provides a powerful new resource for mapping complex traits across many systems and disciplines and should prove to be of great utility in modeling the genetics of complex diseases in human populations.

Recombinant inbred (RI) strains have been used in genetic mapping for approximately 30 years (Bailey 1981; Taylor 1978). They provide numerous advantages, one of the foremost being that strains retain their genetic identity for hundreds of generations but only need to be genotyped once. Furthermore, the additive genetic variance in a panel of RI strains is twice that of a corresponding F_2 intercross for the same trait, a feature that increases the effective power to map many types of subtle polygenic traits. (Dominance effects can be readily estimated by making sets of isogenic F_1 hybrids among RI strains.) Recombination rates in RI strains also average fourfold higher than in conventional test crosses (Haldane and Waddington 1931; Lynch and Walsh 1998; Williams et al. 2001), resulting in improved mapping resolution compared with backcrosses or intercrosses. It is possible to treat, test, and measure many isogenic individuals and to detect and reduce both environmental and technical sources of variance. This attribute is typically exploited to increase the precision of strain means, boosting the effective heritability of a trait. In fact, the effective heritability for genetically specified traits asymptotically approaches 100% with increased sample size (Belknap 1998). This is a crucial advantage for mapping fitness-related or behavioral traits and molecular phenotypes such as transcript and protein expression levels that often have low heritability and high environmental liability.

By phenotyping strains rather than individual mice, it is also possible to systematically explore a crucial and still uncharted area of mouse genetics, namely, gene-by-environment interactions. Sets of

*These authors contributed equally to this work.

Correspondence to: B. Bennett, E-mail: bennettb@colorado.edu

loci that modulate a trait under one set of conditions and environment may differ markedly from the set of loci that control a trait under a second set of conditions and environment. By using RI strains, it becomes practical to intentionally vary environmental factors and to study, for example, the effects of stress, environmental enrichment, and diet on behavioral traits or disease susceptibility.

Perhaps the most significant advantage of using RI strains is that they are a renewable resource for long-term interdisciplinary and collaborative research. Thousands of different phenotypes can be measured at different times and in different laboratories on the same genotypes, using the same strains. This makes it possible to explore genetic correlations, gene pleiotropy, and interactions among traits and the quantitative trait loci (QTLs) modulating those traits. The cumulative nature of such studies is well illustrated by the BXD RI strains, where over 50,000 traits have been collected in this one panel (Chesler et al. 2003; www.webqtl.org/search.html).

The most serious disadvantage of RI strains has been the modest numbers of strains available per set. While plant geneticists typically use 100–1000 RI lines (e.g., Stenchjem and Bruggemann 2001) to map traits, mouse and rat geneticists have been limited to sets of 15–30 strains. This has greatly diminished enthusiasm for the use of RI strains except among behavior geneticists. The power to detect a QTL increases with increasing number of RI strains (Belknap 1998; Belknap et al. 1996), and with only 26 strains (the size of the original BXD panel) a QTL must have an unusually large effect (56% of the additive genetic variance, V_A) to be detected with a power of 0.8 without subsequent confirmation (Belknap et al. 1996). The simple solution to this problem is to generate larger panels of RI strains from genetically diverse parental strains.

In this article, we provide an initial genetic characterization of the LXS panel of RI strains. This panel currently consists of 77 inbred strains that trace their descent from the inbred progeny of a multigeneration cross of eight diverse laboratory strains begun by Gerald McClearn and colleagues 45 years ago (McClearn et al. 1970). We estimate that the LXS RI panel collectively incorporates approximately 3600 recombination breakpoints, providing subcentimorgan precision for monogenic traits and 1–10-cM precision for QTLs modulating complex polygenic traits. We have genotyped the panel using a set of 330 well-characterized and physically mapped MIT microsatellite loci. Our maps provide good and highly usable, but not yet comprehensive, cov-

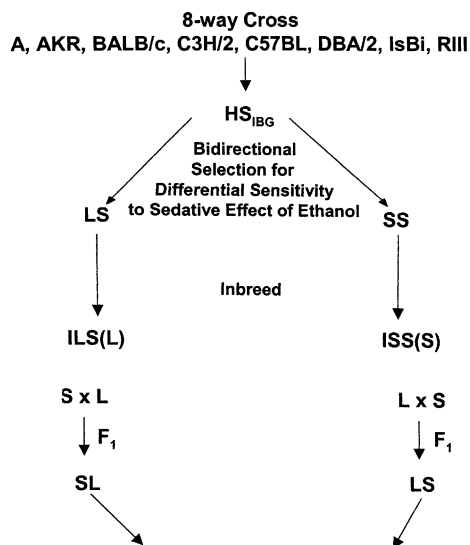
erage. Given the genetic diversity of the source strains, this RI panel should have widespread utility in dissecting polygenic traits and in modeling many complex human diseases.

Materials and methods

Animals. The LXS RI panel was generated (Fig. 1) in the specific pathogen-free facility at the Institute for Behavioral Genetics, Boulder, CO, from crosses between the ILS (*Inbred Long Sleep*, abbreviated to L) and ISS (*Inbred Short Sleep*, abbreviated to S) strains of mice (Markel et al. 1995). These two strains are the product of bidirectional selection for long or short duration of loss-of-righting response following a high-dose, intraperitoneal injection of ethanol. The ILS and ISS and their progenitors have been widely used in alcohol research, and they have been cited in approximately 400 publications over a 30-year period. The LXS panel was initiated at the direction of one of the authors (JCD) and maintained and systematically inbred, until funding was obtained for their continued analysis and inbreeding by TEJ and BB.

The original derivation of the ILS and ISS strains has been described in detail by Bennett and colleagues (2002) but, in brief, these strains trace their descent from a heterogeneous stock (McClearn et al. 1970; McClearn and Kakihana 1981) derived from an eight-way cross of inbred strains maintained at that time by the Cancer Research Genetics Laboratory at the University of California at Berkeley (A, AKR, BALB/c, C3H/2, C57BL, DBA/2, Is/Bi, and RIII). To produce the LXS RIs, ILS and ISS mice were reciprocally intercrossed (three matings of each cross) and all four reciprocal matings (ten mating pairs per cross) of F_1 mice were used to produce the F_2 generation. Inbreeding at the F_2 generation began with 125 mating pairs derived in approximately equal numbers from the four matings used to produce the F_2 (Fig. 1). The F_2 crosses were designed to generate all possible combinations of Y Chromosome (L or S) and mitochondrial genomes (L or S) in approximately equal numbers of strains in all four combinations.

In subsequent generations, three brother–sister matings per strain were set up whenever possible. Matings were maintained until 2–3 litters had matured and the next generation of filial (F) matings had been initiated. A mating was considered unproductive if no litter was produced within 60 days, or if all mice died within a few days of birth. In the case of unproductive matings, progeny from the same line but previous generation were used to replace the original mating, if available; if not, males were



F ₂	Y chromosome	mitochondria	No. strains	
			N ₂	N ₂₂
SL x SL	L	S	23	16
SL x LS	S	S	31	17
LS x SL	L	L	37	22
LS x LS	S	L	34	22

Fig. 1. Derivation of the LXS RI strains showing the original 8-way cross, selection for differential ethanol sensitivity, and inbreeding. Approximately equal numbers of reciprocal F₁ crosses between ILS and ISS were started in 1996, and currently, after inbreeding is complete, approximately equal numbers remain. Each of the four F₂ crosses generated a unique pairing of Y Chromosome and maternal (mitochondrial) contribution.

switched between existing matings. The 20th generation of inbreeding was completed for the last of the LXS strains in 2002. Currently, four mating pairs are maintained per strain. The number of consecutive filial matings currently ranges from F₂₂ to F₂₅.

Genotyping. DNA purification and microsatellite marker screening were done as described by Williams et al. (2001), Lu et al. (2001), and Bennett et al. (2002). Genotyping was begun at F₁₄₋₁₈, a stage at which inbreeding was estimated to be ~94% complete. DNA was often pooled from two mice per strain to maximize the probability of detecting un-fixed loci. Such loci were genotyped again, in subsequent generations, though not always with pooled DNA samples. We also routinely re-genotyped markers associated with flanking proximal and distal recombination events in single strains (so-called double recombinants) to further reduce possible genotyping errors.

Databases and analytical tools. Marker genotypes were entered directly into a relational database system (FileMaker Pro) and exported to Microsoft Excel, MapManager QTX (Manly and Olson 1999; <http://mapmgr.roswellpark.org/mmQTX.html>), and R/qtl (Broman et al. 2003; <http://www.biostat.jhsph.edu/~kbroman/qtl/>) for subsequent analysis. We also assembled a database on the physical map location of just over 6300 MIT microsatellite loci by BLAT analysis of MIT sequence against the Mouse Genome Sequencing Consortium build of the mouse genome at www.genome.ucsc.edu. This made it possible to order markers using either the February 2002 release of the physical assembly (www.genome.ucsc.edu/cgi-bin/hgGateway) or the strain distribution patterns themselves. We generally followed the order of MGSCv3, but since this assembly is still incomplete, in some instances our genetic maps strongly support an alternate marker order. In most cases of contradictory marker orders, we weighted the physical information most strongly but looked for consensus involving that order and one which minimized apparent double-recombination events. Genotype tables for the LXS strains and additional analyses can be obtained as supplemental material at two sites: (1) WebQTL (www.webqtl.org) and (2) as a set of text, Excel, FileMaker, QTX, and R/qtl files at <ftp://atlas.utmcm.edu/public/lxs> [inactive]. These files include both the known or putative physical location of each marker, its centimorgan (cM) position in the final iteration of the Chromosome Committee Reports in 2001 (www.informatics.jax.org/ccr/searches/about.cgi?year=2001), and the cM marker positions calculated from the set of LXS strains themselves.

Results and discussion

Strains. Construction of the LXS panel was initiated by John DeFries and colleagues in 1996, with 125 F₂ mating pairs from reciprocal LXS and SXL F₁ matings. At the 14th generation 86 lines survived, and, currently, at the 22nd generation 77 strains survive. Compared to the loss of 5 of 12 strains of the first RI panel (Bailey 1971) or Green's (1966) estimate that only 1 in 6 strains survive 20 generations of inbreeding, the survival of 77 of 125 inbred strains was a better outcome than expected.

Genotyping. We initially screened 813 MIT microsatellite markers for sequence length polymorphisms between the parental strains ISS and ILS. Of this set, 361 (44%) were polymorphic using methods that consistently detect a difference of 4 bp

or greater and that rarely detect differences of as little as a single repeat (2 bp). This level of genetic difference is similar to that observed in other inbred strains. We subsequently genotyped 330 of these markers, which were comparatively evenly distributed across the genome, in all of the RI strains (Fig. 2). A total of 172 strain-generation combinations were typed and retyped, approximately 80,000 genotypes in total. Here we report on the final set of genotypes for these 330 MIT markers, in a final set of 73 strains. Four strains that were apparently not fully inbred (see below) are still being genotyped. If the 330 markers were spaced evenly across the mouse genome, this would be equivalent to mean separation of ~ 4.1 cM [$1450 \text{ cM} / (330 + n \text{ of chromosomes})$]. The median and mean distance between markers (excluding the coverage gap at the ends of the chromosomes) was similarly 4–5 cM (median ~ 4.0 cM, mean ~ 4.7 cM, Fig. 2). Although the maximal gap on each chromosome is still large and quite variable (7–23 cM; Table 1), a randomly selected gene will typically be well under 4 Mb from the closest flanking marker. Thus, the map, while not final, is nonetheless quite usable for the analysis of both Mendelian and complex traits.

Comparison of genotypes between DNA samples collected at F_{14-18} and at F_{18-20} identified a number of apparently unfixed or miscalled loci. Most of these genotypes were retested at F_{22-23} . Four strains showed a large number of inconsistent genotypes over the multiple generations of testing. These strains (LXS60, LXS102, LXS107, and LXS115) are still unfixed in a number of intervals and are therefore undergoing additional genotyping and inbreeding. These strains have been excluded from the following results and discussion.

Marker fixation. The percent of markers segregating decreases by approximately 19.1% in each successive generation of full sib mating (Green 1981). By F_{20} , 1.8% of loci (6 of 330) are expected to be unfixed (Fig. 3A). In fact, only one strain greatly exceeded this value, with 17 segregating markers. Most strains had 1%–2% of markers still segregating (Fig. 4B: mean = 1.1% at F_{20} versus the expectation of 1.8%), and in 22 strains we found no unfixed markers (Fig. 3B). Since we did not invariably type DNA samples from both parents that gave rise to the next generation of inbreeding, the percentage of unfixed alleles will tend to be underestimated slightly. But even if we double the number of loci segregating (the most conservative estimate of total residual unfixed loci is $2\times$ the number of detected unfixed loci), then we have still achieved 98% fixation at the

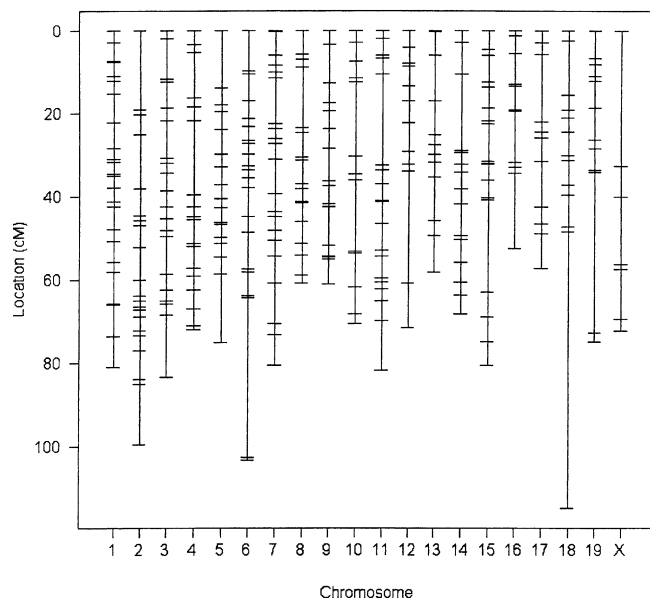


Fig. 2. Location of markers (horizontal bars) used to generate the LXS genetic map, consisting of 330 polymorphic microsatellite markers. The largest single gap on each chromosome is described in more detail in Table 1. These marker positions are based on the LXS genetic map; the proximal end of the chromosome is shown at the 0 position; the location of the distal end of the chromosome was determined by using the MGD distance between the last distal marker genotyped and the last marker position on that chromosome.

F_{20} . The RI strains are now 2–5 generations beyond the F_{20} and thus are estimated to have fewer than 1 in 100 unfixed loci.

Segregation distortion. Segregation distortion in RI sets can give rise to serious problems in mapping QTLs. In particular, an imbalance of *L* and *S* alleles at unlinked markers on different chromosomes generated by epistatic interactions that affect fecundity or fitness could lead to spurious nonsynthetic associations (Williams et al. 2001). It is straightforward to test for deviations of allele segregation. At any given marker, the expectation is that half of the LXS strains should carry an ILS allele and the other half should carry an ISS allele. In the absence of selection or breeding errors, deviations around this 50:50 expectation should have a binomial distribution. The observed frequency of these parental alleles was very close to expectation. The modal value, as shown in Figure 4A, is 50% (mean \pm SD = $49.0 \pm 0.6\%$). This distribution also did not deviate significantly from the binomial prediction (observed $\chi^2 = 12.62$; $\chi^2_{0.95} = 14.07$, $p \sim 0.085$; Fig. 4). The extremes were LXS49 and LXS10 with 27.6% and 63.6% ILS marker genotypes, respectively.

Table 1. Largest gaps between adjacent markers on each chromosome in the LXS RIs

Chr	Length		Number of markers	Largest gap size ^a			Number of markers screened in gap region ^e	
	MGD ^b	LXS ^c		MGD (cM) ^b	LXS (cM) ^c	Physical size (Mb) ^d	Nonpolymorphic	Total markers screened
1	112	260	25	4.6	7.2	7.0	3	5
2	114	281	24	8.5	19.2	15.7	12	16
3	95	227	20	9.7	15	19.4	3	5
4	84	228	18	22.7	17.8	41.3 ^f	20	25
5	92	228	18	10.0	16.5	13.1	11	13
6	75	250	21	8.0	38.4	9.3	8	10
7	74	247	22	0.8	10.8	5.6	2	6
8	82	188	17	13.0	14.4	18.1	16	20
9	74	197	15	10.0	9.4	14.1	16	21
10	77	198	13	12.0	17.9	36.2	10	15
11	80	234	20	7.0	21.9	20.7	28	32
12	61	176	12	14.0	26.9	24.9	17	20
13	80	168	13	14.0	10.9	9.8	8	13
14	69	193	16	6.5	18.3	18.3	12	14
15	74.6	218	18	6.6	20.6	17.9	28	35
16	72	152	13	12.2	18	21.3	14	16
17 ^g	58	143	12	8.2	16.1	9.7 ^g	6	11
18	59	218	13	5.0	66.2	12.1	0	2
19	57	157	12	8.7	38.6	12.1	11	15
X	77	151	7	17	32.7	39.6 ^f	10	15

^aLargest gap between markers on each chromosome identified using the LXS map.

^bMarker positions are based on the Mouse Genome Database (<http://www.informatics.jax.org/>).

^cMarker positions are based on the mapping results using LXS data corrected for map expansion.

^dPhysical size in megabases from NCBI (http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=10090/).

^eNot all markers in this region were tested.

^fThese correspond to the two largest gaps shown in Figure 5; see text for further discussion.

^gThis was the only chromosome for which there is a larger gap (of 10.9 Mb, between the distal marker and the end of the chromosome, identified from NCBI by the terminal marker in their database).

Eleven markers in three chromosomal regions showed more ILS alleles than expected ($p = 0.017$, pointwise test), although none of these deviations was statistically significant following Bonferroni

correction. As Bonferroni corrections are probably overly conservative (see below), these regions are discussed briefly here, as they may represent areas of selection. In addition, any nonsynthetic associations

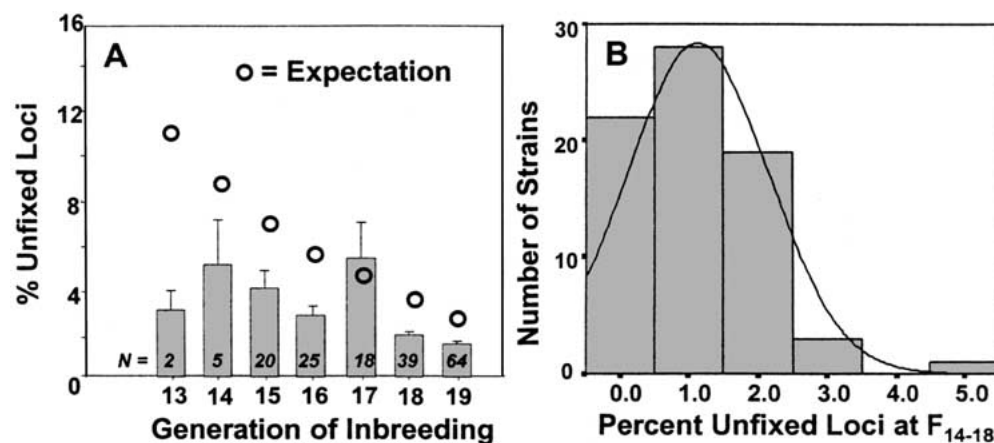


Fig. 3. Unfixed loci. A. Percent of loci (\pm SE) segregating at each generation of sampling. The expectation for the percent of unfixed loci resulting from full-sib crosses is indicated by the open circles along with the observed value for all generations for which DNA was tested. Fewer strains were sampled at the earlier generations (e.g., 13 and 14), so these values are not representative. **B.** Percent loci segregating at the second genotyping, F_{14-18} . There were 43 strains that had significantly ($p < 0.05$) fewer than the expected 17 (5%) unfixed marker loci.

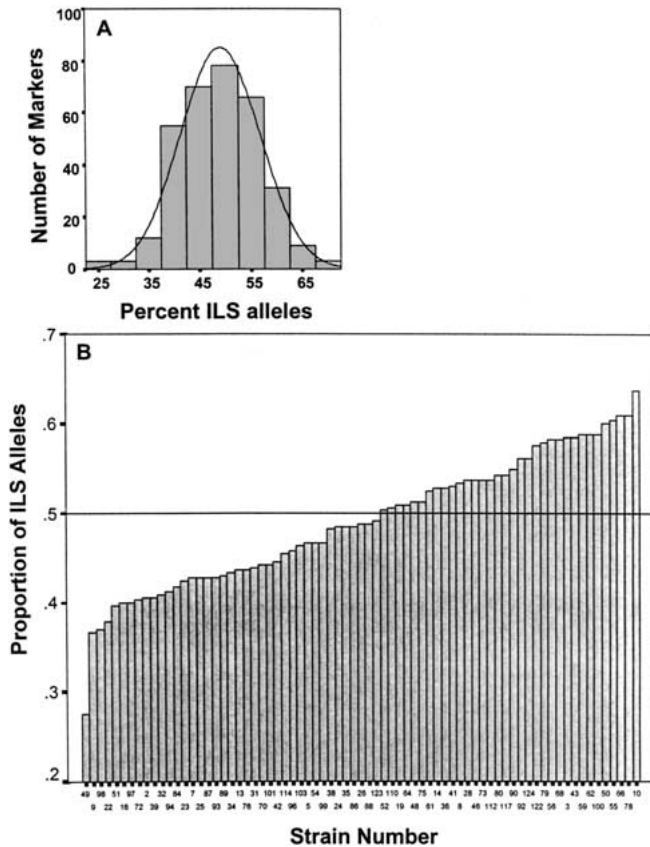


Fig. 4. The distribution of the ILS allele frequency computed for (A) each of the 330 markers genotyped in the LXS panel with normal curve superimposed, and (B) each individual strain. For any given marker or strain, 50% of the alleles in the RI panel were expected to be from ILS and 50% from ISS (horizontal line).

that arise as a result of segregation distortion can affect subsequent mapping. On Chr 11, eight markers between 46 and 66 cM had an average of 68% ILS alleles per marker, differing from expectation at $p = 0.025$ (pointwise). On Chr 6, three markers from the proximal end to 15.6 cM also showed an excess of ILS alleles, averaging 67% ($p < 0.025$). One of these markers, *D6Mit341* at 7.5 cM, showed a nonsyntenic correlation with a marker on Chr 2, possibly because of the excess of ILS alleles, although the Chr 2 marker, *D2Mit81*, did not show an equivalent distortion. On Chr 7, six of eight markers between 41 and 60 cM showed an excess of ISS alleles. The Chr 7 region also is characterized by a significant excess of the albino genotype (55 of 77 strains; $\chi_1^2 = 14.1$, $p = 0.002$). (This excess is not significant with a Bonferroni correction for 250 or more tests.) As 330 markers are not 330 independent linkage groups, a more realistic, though still conservative, Bonferroni correction might be 100, over twice the average number of breakpoints per strain (2.14). This correction suggests that the bias on Chr 7 may be statistically significant and may repre-

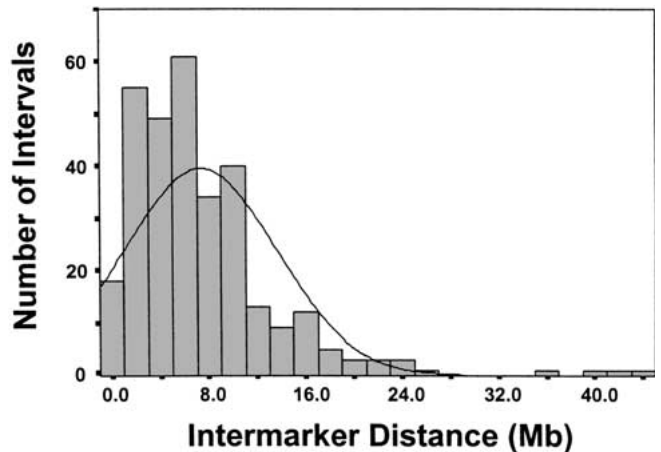


Fig. 5. Histogram of interval length between microsatellite markers in Mb, based on the marker positions from the NCBI database (http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=10090). The two largest gaps have a superscript "f" in Table 1 and are discussed in the text.

sent either an inadvertent bias in selecting white pups for mating during the generation of the LXS panel or, more likely, a fitness advantage of an *S* allele that is linked to the albino *c* allele at the *Tyr* locus carried by ISS. Thus, the Chr 7 region, and perhaps the other regions, represent areas of selection, either inadvertent or fitness-related.

Marker distribution. Despite our best attempt, the spacing of genetic markers was not uniform across the genome (Fig. 2). The average distance between markers is 7.4 Mb (this figure was calculated by dividing the total physical distance from the proximal marker to the distal marker on each chromosome, summed over all chromosomes, divided by 330 markers) or roughly 4 cM. Although the variability is large, as shown by the right hand tail in Figure 5, which shows the distribution of intermarker gaps, there are relatively few gaps larger than 16 Mb. The largest gap on each chromosome is given in Table 1; these gaps were identified using the LXS map, based on recombinations observed in the LXS RI, since the utility of this panel for mapping is based in part on this recombination frequency. These gaps are described using two additional sources in Table 1: (1) MGD positions based on consensus marker locations from the Mouse Genome Informatics Web Site, The Jackson Laboratory, Bar Harbor, ME, December 2003 (URL: www.informatics.jax.org); and (2) the physical distance between the flanking markers in megabases (Mb) based on the UCSC Genome Browser and NCBI database (http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=10090). The largest regions between markers on each chromosome range

from 9.3 to 43.4 Mb (Table 1). The two largest gaps, on Chromosomes 4 and X, are indicated by superscript "f" Table 1 and are shown in the histogram of gap sizes in Figure 5. In only one case (Chromosome 17) is there a larger gap than indicated in Table 1, between the last marker shown in Figure 2 and the distal end of the chromosome.

Some of these gaps may be recombinational hotspots. For example, on Chr 6, the LXS map distance between *D6Mit291* and *D6Mit15* is 38.3 cM but the physical distance is only 9.3 Mb (Fig. 2). Similarly, on Chromosome 18, the LXS map distance between *D18Mit7* and *D18Mit128* is 66.2 cM, while the physical distance is 12.1 Mb (Fig. 2). Other gaps are probably due to mapping discrepancies in the Mouse Genome Database (<http://www.jax.org/>). For example, the genomic sequence (NCBI) suggests that the markers mapping to and defining the distal end of Chr 2 (Fig. 2) are actually proximal to our most distal marker in the LXS panel, indicating that the large terminal gap shown in Figure 2 is not real. The majority of the gaps (e.g., the two largest gaps on Chr 4 and Chr X; Fig. 5 and Table 1) appear to be real, as their physical sizes (41.3 and 43.4 Mb, respectively) approximate the genetic map distance reported by MGD and observed in the LXS.

Other regions may simply not be polymorphic due to shared ancestry in the mouse strains used to generate the common inbred strains (Wade et al. 2002). Moreover, the ILS and ISS parental strains share chromosomal-regions that are expected to be identical by descent (IBD) at multiple sites throughout their genome. Given that eight inbred strains (six of which were commonly used strains) contributed to the heterogeneous stock (HS) from which the LS and SS selected lines were derived (McClearn and Kakihana 1981; Fig. 1), one-eighth of the genome should be IBD, assuming random fixation. In fact, this figure is probably an underestimate due to fixation of alleles in the HS progenitor strains. In 15 of the gap regions (Table 1, last 2 columns), 70% or more of the markers tested were not polymorphic, consistent with the hypothesis that some or all of these regions are IBD. Although in most gaps exhaustive sampling of MIT markers was not done, on Chromosomes 3, 8, and 11, we genotyped 83%, 56%, and 61%, respectively, of all microsatellite markers listed at MGD. We also typed all available MIT microsatellites in two intervals which appear to be IBD. On Chr 2, none of 21 MIT markers between 147.0 and 160.3 Mb was polymorphic (Bennett et al. 2002), and on Chr 15, none of 19 markers between 70.0 and 72.9 Mb was polymorphic (unpublished).

The order of markers (Fig. 2) was generally consistent with the MGSCv3 physical map. In five

cases, on four chromosomes, the LXS order appeared to reverse the physical map order. (There is some uncertainty as to the position of *D15Mit93* on the physical map so this reversal may not be real.) In all of these cases, there was some ambiguity in the genetic maps, with some supporting the physical map and others supporting the LXS genetic map. In the most extreme example, no genetic map supports the marker order switch involving *D14Mit257* and *D14Mit101*.

Map expansion. The mean uncorrected recombination map length of the LXS RI strains is 41.5 Morgans (M) compared with the 13–15-M lengths of single-generation meiotic maps. This represents a map expansion of about 3.4-fold, with a 95% confidence interval (CI) of 3.1–3.7. This expansion was calculated as the ratio of the expanded map distance, between proximal and distal markers, to the unexpanded map distance, from MGD, summed over all chromosomes. Map expansion in F_2 -derived RI strains is expected to average four-fold in RI strains mapped with infinite marker density (Haldane and Waddington 1931), 3.7-fold with 1-cM marker density, and two-fold with 16-cM marker density (Williams et al. 2001). Therefore, the 3.4-fold expansion we observed based on a 4–5-cM map density is in line with expectation. The map expansion among strains ranged from a low of 2.2-fold in LXS3 to a high of 4.4-fold in LXS35 (Fig. 6).

As expected, all chromosomes showed expansion in total length (Table 1), as determined by summing the distances (based on MGD and recombination frequency for LXS) between the extreme centromeric and telomeric markers. Thus, each value in Table 1 slightly underestimates the chromosome length. The average distance between pairs of adjacent markers based on the LXS map is 13.2 cM; the equivalent distance on the MGD map is 4.3 cM. Most intervals (94%) between marker pairs (311) are expanded in LXS relative to MGD. However, 35 intervals were contracted, and one showed no change. Most of the apparent contractions are quite small, less than 3 cM, and only two of these intervals fall outside of a two-SD cutoff from the mean size. These two intervals, one on Chr 16 at approximately 50 cM, and the second on the X Chr at approximately 58 cM, are not supported by the physical map and thus are probably not real.

Independence of strain genotypes. If RI strains share large regions with similar or identical genotypes across a number of chromosomes, this is generally caused by breeding errors. Nonindependence of strain genotypes reduces the information content

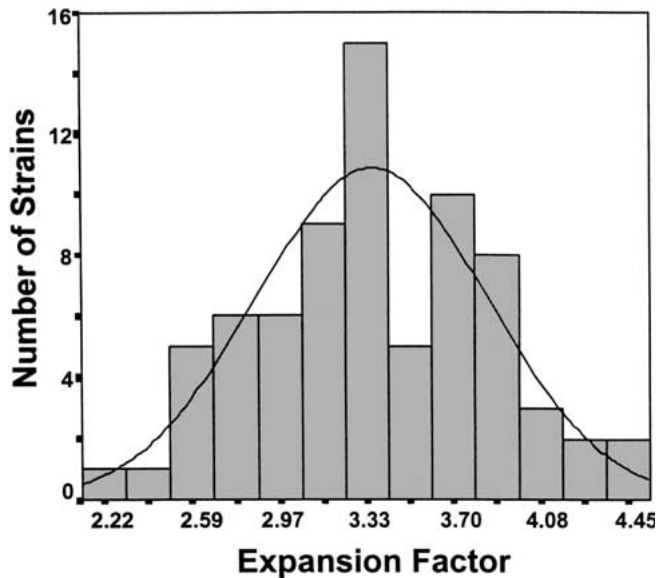


Fig. 6. Mean map expansion in LXS RI strains (3.39 ± 0.15), expressed as mean number of recombinations per 100 cM. This factor was calculated by dividing the total number of recombinations per strain by the total map distance between the most proximal and most distal markers, summed over all chromosomes, in the 73 strains used for mapping.

of an RI set and distorts genetic maps. To identify and exclude such nonindependent strains, we computed all pairwise correlations of genotypes of LXS strains. Figure 7 shows the genetic similarity of the RI strains to each other based on the percentage of identical genotypes at 330 markers. This analysis revealed that two of the LXS strains were 97% identical, indicating that they were not independent. One of these two strains has been removed and is not counted as one of the 77 strains comprising the LXS panel. The normal distribution of genotypic similarity for the 77 remaining strains indicates that the percent identities are randomly distributed as expected, and thus all 77 strains are informative for genetic mapping. The slight skew gave a mean of 50.5% identity, but this deviation was not statistically significant. The distribution range also matches closely with the BXN RI strains characterized for strain independence by Williams et al. 2001) and the HXB/BXH rat RI panel characterized by Jirout et al. (2003).

Nonsyntenic correlations. Two markers that are tightly linked will have nearly the same strain distribution pattern (SDP) and a high positive correlation. In contrast, unlinked markers and markers on two different chromosomes in general should not have significant positive or negative correlations beyond that expected by chance. However, non-

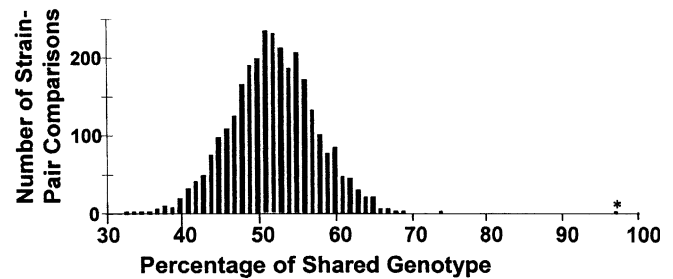


Fig. 7. Genetic identity among LXS RI strains. The percentage of identical genotypes was determined for all pairwise combinations of 78 strains from a matrix of strain genotypes. One pair of strains showed 97% identity (*) and one of those strains was removed from the panel.

syntenic associations arise when regions on different chromosomes show statistically significant positive or negative correlations. This can be due to chance fixation of the same alleles, but in some cases it is likely due to selective advantage of epistatic combinations of alleles on two or more different chromosomes (Williams et al. 2001). Nonsyntenic associations are revealed by searching a correlation matrix of marker genotypes for all pairs of markers with high correlations. A number of nonsyntenic loci are effectively linked, albeit weakly. One such correlation ($r = 0.33$) between *D6Mit341* and *D2Mit81* results from a previously mentioned disproportionately high representation of ILS alleles at both markers. Given the very large number of correlations we examined ($330 \times 329/2$), we expect some correlations with statistically significant absolute values by chance. The number of effective tests is much less than 54,285 as a result of linkage, but even if we assume the effective number of tests is approximately 10,000, then a correlation of 0.3 is still not significant at $p < 0.05$. Nevertheless, even though this effect is not statistically significant, it will still generate weak spurious correlations. In general, the large size of the LXS panel makes it far less sensitive to nonsyntenic association than the BXD RI set ($n = 26$) and even smaller panels such as the CXB set ($n = 13$).

Genetic mapping in the LXS. We are presently using the LXS RI panel to map a variety of alcohol-related traits, including ethanol-induced loss of righting, psychomotor activation, hypothermia, blood ethanol concentration at awakening, and acute functional tolerance, and numerous CNS morphometric traits. These ethanol-related traits in particular are all known to differ significantly between the ILS and ISS parental strains (Markel et al. 1995; Owens et al. 2002). Many other traits have been shown to differ between the selected, noninbred LS

and SS strains and have been mapped in a previous, smaller RI panel (LSXSS; DeFries et al. 1989). A brief sampling of these traits includes the alcohol-related phenotypes listed above, as well as ethanol preference (Church et al. 1979), corticosterone release in response to acute (Wand 1990) or chronic ethanol treatments (Wand 1989), a variety of behavioral stressors (Wand 1990), structural responses in the hippocampus in response to acute (Poelchen et al. 2000) or chronic ethanol (Lang et al. 1997; Markham et al. 1987; Huang and McArdle 1993), open field activity and defecation (Swanberg et al. 1979), and the anxiolytic effects of ethanol measured in several different tests (Allan and Isaacson 1985; Cao et al. 1993; Lapin et al. 1996; Stinchcomb et al. 1989). In addition, differential sensitivity to numerous other anesthetic agents, alcohols, and drugs of abuse have been found in the ILS and ISS (Simpson et al. 1998). We therefore anticipate that the large LXS panel will find extensive uses in neurogenetic, neuropharmacological, and behavioral studies (cf. Deitrich and Pawloski 1990).

Other phenotypes, such as body weight, and the response of body temperature to dietary restriction do not differ between the parental strains but nonetheless show significant variability and heritability across the LXS strains and, thus, should also be particularly amenable to QTL analysis (Rikke et al. 2003, 2004; Bennett et al. unpublished). We believe this panel will thus provide a valuable new resource to the mouse community for genetic mapping. In a panel of 77 RIs, QTLs accounting for approximately 25% of V_A can be detected at a statistically significant p -value (<0.0001) with a power of 80%, by testing replicate mice within each strain. The exact within-strain sample size depends on the heritability of the trait (Belknap 1998). In a simple illustration of the increased power in the LXS panel, the tyrosinase locus (*Tyr*) was mapped in this panel, and a single major peak at *D7Mit350* (73.3 Mb) near the *Tyr* locus (75.9 Mb) was identified with a LOD of 29. In an earlier RI panel (LSXSS) of 27 strains, derived from the noninbred, selected LS and SS lines (DeFries et al. 1989), *Tyr* mapped to the same position with a LOD of 12. The albino LOD-score analysis was extended to 19 other loci as well, and a similar improvement of 2.8-fold in LOD score was obtained on average when using 73 strains versus 27 strains. In general, LOD scores are directly proportional to the number of strains (n): $\text{LOD} \approx (n \times r^2)/4.61$, where r^2 is the QTL effect size, with 1 df. Several of the alcohol-related phenotypes that we are currently investigating in the LXS panel have previously been assessed in the LSXSS (Markel et al. 1996; Erwin et al. 1997; Gehle and Erwin 2000); we are currently

mapping these traits in the LXS panel and anticipate similar improvements in significance of QTLs.

Conclusions

We have generated and characterized the LXS panel of 77 RI strains that will be a useful mapping resource for the mouse genetics community. Currently, 330 microsatellite marker genotypes at an average marker spacing of 4–5 cM have been assessed in 73 of 77 strains. We are adding additional markers to the SDP to achieve a smaller intermarker distance and are completing genotypic assessment of the four strains with inconsistent genotypes in the earlier screens. The number of strains in this panel provides 80% power, without subsequent confirmation, to map QTLs that explain as little as 23% of V_A at a statistically significant level. Even at the present marker density, the mapping precision will be improved by approximately threefold over a panel of 27 strains.

Acknowledgments

We thank Jerome Salazar and Jean Wah for careful and reliable animal care and breeding and Eugene Thomas for the initial breeding. We also thank Nancy Phares-Zook for her research contributions and for her outstanding job managing our many diverse databases. Linda Lou Wessman, Justin Springett, and Jessica Hall cut many tails and extracted DNA. Lena Gordon ran numerous screening and confirmatory gels. We also thank our colleagues at IBG and elsewhere who gave advice throughout the study and in reading the many versions of the manuscript. We thank Shuhua Qi for excellent support in genotyping the LXS strains and Yanhua Qu for helping with the BLAT analysis of microsatellite sequences. Dr. Elissa Chesler patiently and on numerous occasions gave BB very thorough instruction on application and interpretation of R/qtl. This work was supported by grants from the NIH [RO1 AA11984-(TEJ) and KO1 AA00195 (TEJ) and P50 AA13755 (RWW)], by the Ellison Foundation for Medical Research (TEJ), and by funds from the University of Colorado. Genotyping costs were supported in part by P20-MH 62009 (a Human Brain Project to RWW) and the JNIA Genotyping Core (U24AA13513 to LL and RWW).

References

1. Allan A, Isaacson R (1985) Ethanol-induced grooming in mice selectively bred for differential sensitivity to ethanol. *Behav Neural Biol* 44, 386–392

2. Bailey DW (1971) Recombinant inbred strains. *Transplantation* 11, 325–327
3. Bailey DW (1981) Recombinant inbred strains and bilineal congenic strains. In: *The Mouse in Biomedical Research*, Foster HL, Small JD, Fox JG eds. New York: Academic Press, pp 223–239
4. Belknap JK (1998) Effect of within-strain sample size on QTL detection and mapping using recombinant inbred mouse strains. *Behav Genet* 28, 29–38
5. Belknap JK, Mitchell SR, O'Toole LA, Helms ML, Crabbe JC (1996) Type 1 and type 2 error rates for quantitative trait loci (QTL) mapping studies using recombinant inbred mouse strains. *Behav Genet* 26, 149–160
6. Bennett B, Beeson M, Gordon L, Johnson TE (2002) Reciprocal congenics defining individual quantitative trait loci for sedative/hypnotic sensitivity to ethanol. *Alcohol Clin Exp Res* 26, 149–157
7. Broman KW, Rowe LB, Churchill GA, Paigen K (2002) Crossover interference in the mouse. *Genetics* 160, 1123–1131
8. Cao W, Burkholder T, Wilkins L, Collins AC (1993) A genetic comparison of behavioral actions of ethanol and nicotine in the mirrored chamber. *Pharmacol Biochem Behav* 45, 803–809
9. Chesler EJ, Wang JB, Lu L, Qu Y, Manly KE, et al. (2003) Genetic correlates of gene expression in recombinant inbred strains: A relational model system to explore neurobehavioral phenotypes. *Neuroinformatics* 1, 343–357
10. Church AC, Fuller JL, Dann L (1979) Alcohol intake in selected lines of mice: Importance of sex and genotype. *J Comp Physiol Psychol* 93, 242–246
11. DeFries JC, Wilson JR, Erwin VG, Petersen DR (1989) LS × SS recombinant inbred strains of mice: Initial characterization. *Alcohol Clin Exp Res* 13, 196–200
12. Deitrich RA, Pawlowski AA (1990) Initial sensitivity to alcohol. In: *A Workshop on Alcohol Intoxication* (Keystone, CO: U.S. Department of Health and Human Services, Public Health Service)
13. Erwin VG, Markel PD, Johnson TE, Gehle VM, Jones BC (1997) Common quantitative trait loci for alcohol-related behaviors and central nervous system neurotensin measures: Hypnotic and hypothermic effects. *J Pharmacol Exp Ther* 280, 911–918
14. Gehle VM, Erwin VG (2000) The genetics of acute functional tolerance and initial sensitivity to ethanol for an ataxia test in the LS × SS RI strains. *Alcohol Clin Exp Res* 24, 579–587
15. Green EL (1981) *Genetics and Probability in Animal Breeding Experiments*. New York: Oxford University Press
16. Haldane JBS, Waddington CH (1931) Inbreeding and linkage. *Genetics* 16, 357–374
17. Huang GJ, McArdle JJ (1993) Chronic ingestion of ethanol increases the number of Ca²⁺ channels of hippocampal neurons of long-sleep but not short-sleep mice. *Brain Res* 615, 328–330
18. Jirout M, Krenova D, Kren V, Breen L, Pravenec M, et al. (2003) A new framework marker-based linkage map and SDPs for the rat HXB/BXH strain set. *Mamm Genome* 14, 537–546
19. Lang D, Beno M, Fifkova E, Eason H (1997) Fine structure of hippocampal dendrites in the dentate fascia of LS/SS mice after chronic ethanol treatment. *Prog Neuropsychopharmacol Biol Psychiatry* 21, 1031–1042
20. Lapin IP, Mirzaev S (1996) The contrary effects of ethanol on the behavior of short-and long-sleep C57BL/6 mice in a dark-light chamber. *Eksp Klin Farmakol* 59, 50–52
21. Lu L, Airey DC, Williams RW (2001) Complex trait analysis of the hippocampus: Mapping and biometric analysis of two novel gene loci with specific effects on hippocampal structure in mice. *J Neurosci* 21, 3503–3514
22. Lynch M, Walsh B (1998) *Genetics and Analysis of Quantitative Traits*. (Sunderland, MA: Sinauer Associates, Inc.)
23. Manly KF, Olson JM (1999) Overview of QTL mapping software and introduction to Map Manager QT. *Mamm Genome* 10, 327–334
24. Markel PD, DeFries JC, Johnson TE (1995) Use of repeated measures in ethanol-induced loss of righting reflex in inbred long-sleep and short-sleep mice. *Alcohol Clin Exp Res* 19, 299–304
25. Markel PD, Fulker DW, Bennett B, Corley RP, DeFries JC, et al. (1996) Quantitative trait loci for ethanol sensitivity in the LS × SS recombinant inbred strains: Interval mapping. *Behav Genet* 26, 447–458
26. Markham J, Fifkova E, Scheetz A (1987) Effect of chronic ethanol consumption on the fine structure of the dentate gyrus in long-sleep and short-sleep mice. *Exp Neurol* 95, 290–302
27. McClearn GE, Kakihana R (1981) Selective breeding for ethanol sensitivity: Short-sleep and long-sleep mice. In: *Development of Animal Models as Pharmacogenetic Tools*, McClearn GE, Deitrich RA, Erwin VG eds. (Washington, DC: U.S. Government Printing Office), pp 81–113
28. McClearn GE, Wilson JR, Meredith W (1970) The use of isogenic and heterogenic mouse stocks in behavioral research. In: *Contributions to Behavior-Genetic Analysis: The Mouse As a Prototype*, Lindzey G, Thiessen DD eds. (New York: Appleton-Century-Crofts), pp 3–22
29. Owens J, Bennett B, Johnson TE (2002) Possible pleiotropic effects of genes specifying sedative/hypnotic sensitivity to ethanol on other alcohol-related traits. *Alcohol Clin Exp Res* 26, 1461–1467
30. Poelchen W, Proctor WR, Dunwiddie TV (2000) The *in vitro* ethanol sensitivity of hippocampal synaptic gamma-aminobutyric acid (a) responses differs in lines of mice and rats genetically selected for behavioral sensitivity or insensitivity to ethanol. *J Pharmacol Exp Ther* 295, 741–746
31. Rikke BA, Yerg JE, Battaglia ME, Nagy TR, Allison DB, et al. (2003) Strain variation in the response of body temperature to dietary restriction. *Mech Ageing Dev* 124, 663–678
32. Rikke BA, Yerg JE, Battaglia ME, Nagy TR, Allison DB, et al. (2004) Quantitative trait loci specifying the response of body temperature to dietary restriction. *J Gerontol* 59A, 11

33. Stenehjem S, Bruggemann E (2001) Glume bar phenotypes in a b73 × mo17 recombinant inbred population reveal the epistatic interaction between p11 and b1. In: Marize Genetic Conference, 42nd Annual Marize Genetic Conference, Coeur d'Alene, ID (<http://www.maizegdb.org/maize-meeting/2001/03abstracts.pdf>, p 80)
34. Stinchcomb A, Bowers BJ, Wehner JM (1989) The effects of ethanol and ro 15-4513 on elevated plus-maze and rotarod performance in long-sleep and short-sleep mice. *Alcohol* 5, 369–376
35. Swanberg KM, Wilson JR, Kalisker A (1979) Developmental and genotypic effects on pituitary–adrenal function and alcohol tolerance in mice. *Dev Psychobiol* 12, 201–210
36. Taylor BA (1978) Recombinant inbred strains: Use in gene mapping. In: *Origins of Inbred Mice*, Morse HC ed. New York: Academic Press, pp 423–435
37. Wade CM, Kulbokas EJ, Kirby AW, Zody MC, Mullikin JC, et al. (2002) The mosaic structure of variation in the laboratory mouse genome. *Nature* 420, 574–578
38. Wand GS (1989) Ethanol differentially regulates proadrenocorticotropin/endorphin production and corticosterone secretion in LS and SS lines of mice. *Endocrinology* 124, 518–526
39. Wand GS (1990) Differential regulation of anterior pituitary corticotrope function is observed *in vivo* but not *in vitro* in two lines of ethanol-sensitive mice. *Alcohol Clin Exp Res* 14, 100–106
40. Williams RW, Gu J, Qi S, Lu L (2001) The genetic structure of recombinant inbred mice: High-resolution consensus maps for complex trait analysis. *Genome Biol* 2(11), RESEARCH0046