Pain 126 (2006) 24–34

PAIN

www.elsevier.com/locate/pain

Screening for pain phenotypes: Analysis of three congenic mouse strains on a battery of nine nociceptive assays

Jeffrey S. Mogil [a,*], Jennifer Ritchie [a], Susana G. Sotocinal [a], Shad B. Smith [a], Sylvie Croteau [b], Daniel J. Levitin [a], Anna K. Naumova [b,c]

[a] *Department of Psychology and Centre for Research on Pain, McGill University, Montreal, Que., Canada*
[b] *Department of Obstetrics and Gynecology, Research Institute of McGill University Health Centre, Montreal, Que., Canada*
[c] *Department of Human Genetics, Research Institute of McGill University Health Centre, Montreal, Que., Canada*

Received 15 February 2006; received in revised form 24 May 2006; accepted 6 June 2006

Abstract

In an attempt to identify new genes responsible for variability in pain sensitivity, we tested three congenic mouse strains – in which a small portion of the genome of the MOLF/Ei strain has been placed on a C57BL/6 genetic background – on a battery of nine nociceptive assays, chosen to reflect those assays in most common use in the pain literature. Mice of both sexes were evaluated by two different testers at different points in time, allowing us to examine the relative importance of genotype, sex, tester and cohort effects on data from these assays. We find strong evidence for the existence of two quantitative trait loci (i.e., genomic regions containing variability-causing genes), one for thermal nociception on mouse chromosome (Chr) 17 (Chr 17; *Tpnr3*) and one for formalin test nociception on mouse Chr 12 (*Nociq3*). We note, however, that the nociceptive assays in this battery feature strong main effects and interactions of sex, tester, and cohort, which if not controlled or covaried can seriously confound interpretation of genetic experiments, including the comparison of transgenic knockout mice to their wild-type controls.

Keywords: Algesiometry; Genetics; Nociception; QTL; Sex difference; Tester effects

1. Introduction

Pain in humans and nociception in animals both display robust interindividual variability, which is now widely appreciated to have a significant genetic component (Mogil, 2004). The involvement of causal genetic variants in pain-related *variability* cannot be demonstrated with popular transgenic knockout models or microarray studies, but rather requires the application of genetic linkage mapping or association studies (Mogil and McCarson, 2000). Although many human association studies of pain are now being performed, they are still largely being performed one-gene-at-a-time, with only the most obvious candidate genes being investigated (Belfer et al., 2004). Truly novel genetic discoveries will require a more systematic, blinded approach (Risch, 2000), but for the time being these remain prohibitively expensive. In the meantime, the mouse genome will be intensively scrutinized for a potentially large number of nociception variability genes, which when found can be "translated" to humans (Mogil et al., 2003, 2005c; Fillingim et al., 2005).

An attractive strategy for identifying trait-relevant genes is the phenotypic characterization of congenic mouse strains (Bailey, 1981). Congenic strains are derived by repeated backcrossing over multiple generations while preserving (in the modern approach, via the DNA marker-assisted choice of breeder males) a

* Corresponding author. Tel.: +1 514 398 6085; fax: +1 514 398 4896.
E-mail address: jeffrey.mogil@mcgill.ca (J.S. Mogil).

0304-3959/$32.00
doi:10.1016/j.pain.2006.06.004

differential genomic segment (Bennett, 2000). One eventually produces an inbred strain genetically identical to the "recipient" strain that was backcrossed to, with the exception of a small genomic region from the "donor" strain. This technique is commonly used to reduce the genetic interval containing a quantitative trait locus (QTL) – a genomic region containing a trait variability gene or genes – but presents certain advantages for the initial detection of QTLs as well, since no molecular genotyping is required, and simple, robust statistical methods are appropriately applied.

One of us (A.K.N.) recently generated three congenic strains in which genetic segments (representing ≈1–2% of the genome each) from Chr 2, 12 and 17 from the MOLF/Ei strain had been placed on a C57BL/6 genetic background. The MOLF/Ei strain is derived from a different subspecies (*Mus musculus molossinus*) than C57BL/6, and is therefore likely to be phenotypically distinct (see Koide et al., 2000), although standard inbred strains are now known to be genomic mosaics including *M. m. molossinus* contributions (Wade et al., 2002). We tested these congenic strains systematically for nociceptive sensitivity, as compared to C57BL/6, using a battery of nociceptive tests that we had developed for use in testing transgenic knockout mice (Mogil et al., 2005a). Such test batteries are increasingly popular for phenotypic characterization of genetically altered mice (Crawley and Paylor, 1997), but to our knowledge no pain-specific battery has thus far been explicitly proposed. During the testing of these congenic strains, one of the experimenters left the laboratory and the remainder of the data was collected by another. Given that we have shown previously that tester is the largest source of tail-withdrawal test variability (Chesler et al., 2002a,b), we used the opportunity to evaluate the effect of tester on all nine assays of nociception considered. We also examined the influence of subject sex and cohort in these assays.

2. Materials and methods

2.1. Subjects

Subjects in these experiments were naïve, young adult (6–14 weeks old) C57BL/6J mice (The Jackson Laboratory, Bar Harbor, ME), and three congenic strains (see next section). In every assay, approximately equal numbers of male and female mice were tested simultaneously (Mogil and Chanda, 2005). All mice were weaned at 18–21 days, and housed in standard shoebox cages of 2–4 with same-sex littermates in a temperature-controlled (20 ± 1 °C) environment (14 h:10 h light/dark cycle; lights on at 07:00 h), and with *ad lib* access to food (Harlan Teklad 8604) and tap water.

Although not formally necessary, it would have been useful to test MOLF/Ei mice simultaneously to C57BL/6 and the MOLF-derived congenics. However, these mice were not available for purchase at the time these studies were conducted, due to very poor fecundity.

2.2. Generation of congenic strains

To generate congenic strains, (C57BL/6 × MOLF/Ei) F_1 hybrid mice were backcrossed to C57BL/6 mice for 10 generations. In each consecutive generation, the heterozygous mice carrying the desired chromosomal region from the wild-derived (*M. m. molossinus*) MOLF/Ei strain were selected and backcrossed to C57BL/6. In the case of the B6.MOLF2 strain, breeder mice were selected for heterozygosity at the *agouti* locus and microsatellite marker, *D2Mit213*. To generate the B6.MOLF17 strain, breeders were selected for heterozygosity at *D17Mit57*. The B6.MOLF12 strain was generated by selecting breeders for the imprinted gene, *Dlk1* (Croteau et al., 2005). The sizes of the donor MOLF/Ei segments were determined by genotyping microsatellite markers, and represent ≈30 cM from Chr 2 (B6.MOLF2), ≈13 cM from Chr 12 (B6.MOLF12), and ≈20 cM from Chr 17 (B6.MOLF17). The mouse genome is approximately 1400 cM, so these segments represent 0.9–2.1% of the total. Fig. 1 illustrates the approximate extent of the MOLF/Ei donor segments in each congenic strain. Congenic strains used in these experiments were bred in the first author's vivarium alongside C57BL/6 counterparts.

2.3. Testers

Two testers, both full-time, professional Research Associates in the first author's laboratory, collected all behavioral data (including all scoring of associated digital videos; see below): J.R. and S.G.S. J.R. has substantial experience testing rodents on standard nociceptive assays, including more than 3 years of experience with these assays in mice at the start of the study. She tested all of the B6.MOLF17 mice and half of the B6.MOLF12 mice, and then left the laboratory. S.G.S. was trained on these assays directly by J.R., and tested all of the

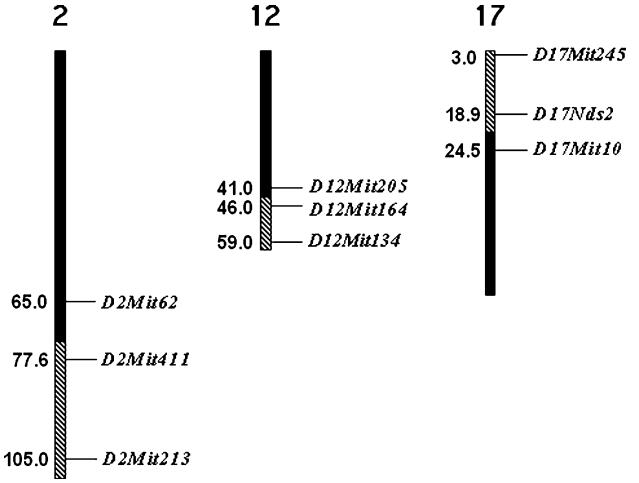

Fig. 1. Differential genomic segments of congenic strains B6.MOLF2, B6.MOLF12 and B6.MOLF17 on mouse chromosomes 2, 12 and 17, respectively. C57BL/6 (B6)-derived genome is shown in black; MOLF/Ei genome is shown in cross-hatching. On the right, the markers used to determine the boundaries of the congenic regions are shown; their position in cM is on the left. Exact starting and ending positions of the markers in Mb can be found at http://www.ensembl.org/Mus_musculus/index.html.

B6.MOLF2 mice and half of the B6.MOLF12 mice (the latter approximately 1 year after J.R. had tested a separate cohort of them).

In all experiments, C57BL/6 mice were tested concurrently to congenic mice (and in approximately equal numbers), and thus were tested by both J.R. and S.G.S. on two separate occasions each.

One potential source of variability among testers in chemical/inflammatory assays is a varying "definition" of an abdominal constriction/writhe or a hindpaw lick/bite (see Sections 2.4.1 and 2.4.3). We have investigated this issue in our laboratory (unpublished observations), and found moderate-to-high inter-rater reliability of abdominal constriction/writhing count and formalin test licking duration data (r's = 0.6–0.8). The inter-rater reliability is greatly improved by using sampling strategies in both tests (r's > 0.9). A subset of the abdominal constriction data reported herein were rescored using sampling by a third observer, and genotypes were found still not to differ in any case.

2.4. Algesiometric assays

Nine of the 10 most common (see Section 4.3 and Fig. 5a) algesiometric assays were chosen in an attempt to cover most of the pain "domain" as defined by previous work in our laboratory (Mogil et al., 1999b; Lariviere et al., 2002). Given the very high genetic correlations between strain-dependent hypersensitivity induced by inflammatory vs. neuropathic injury (Mogil et al., 1999b; S.E. Crager, S.G. Sotocinal and J.S. Mogil, unpublished data), for our present purposes it was not necessary to include a neuropathic hypersensitivity model.

All assays have been fully described on a number of occasions (Mogil et al., 1999a, 2005a; Lariviere et al., 2002), and are described only briefly here (in alphabetical order by abbreviation), below. Note that the *same* mice were tested on all acute nociceptive assays (hot-plate test, paw-withdrawal test, tail clip test, tail-withdrawal test, von Frey test; in random order across experiments), for reasons of practicality and ethics. At least 48 h separated successive tests. Our experience suggests that previous exposure to any of these acute assays does not greatly affect responses on any other, although we do not have systematic data proving this fact. It has been shown that mice tested on a battery of behavioral assays before being tested on the hot-plate test show reduced latencies compared to naïve mice (McIlwain et al., 2001; Voikar et al., 2004), but the relevance of this finding is unclear since in those studies the hot-plate test was the only algesiometric assay. For the tonic assays (abdominal constriction test, formalin test, and zymosan hypersensitivity tests), naïve mice were used in one test only. All testing occurred between 08:00 and 16:00 h, in the animals resting phase.

2.4.1. Abdominal constriction test (AC)

Mice were habituated for at least 30 min to four individual Plexiglas observation chambers (15 cm diameter; 22.5 cm high), placed atop a glass surface suspended over four black/white, high-resolution video cameras. Mice were injected intraperitoneally (10 ml/kg) with 0.9% acetic acid, and videotaped digitally for 30 min after the injection. The videotapes were later viewed by the same experimenter who gave the injections, and the number of lengthwise constrictions of the abdominal musculature ("writhes") was counted. Sample sizes were 7–8/genotype.

2.4.2. Hot-plate test (HP)

Mice were placed within a transparent Plexiglas cylinder (15 cm diameter; 22.5 cm high) on a metal surface (Columbus Instruments Hotplate Analgesia Meter) maintained at 50.0 or 53.0 °C (±0.2 °C) (HP_{50} and HP_{53}, respectively). The latency to respond with a hindpaw lick or shake/flutter, whichever came first, was measured to the nearest 0.1 s with a stopwatch. This test was only performed once, since repeated testing leads to systematic latency alterations (Wilson and Mogil, 2001). Sample sizes were 12–16/genotype/temperature.

2.4.3. Formalin test (F)

Mice were habituated for at least 30 min to four individual Plexiglas cylinders (as above) placed atop a glass surface suspended over four black/white, high-resolution video cameras. All subjects were then given a subcutaneous injection of 5% formalin into the plantar right hindpaw (20 μl volume), and videotaped digitally for 60 min following the formalin injection. Videotape observations were later sampled for 5 s at 1-min intervals by the same experimenter who gave the injections, and the presence or absence of right hindpaw licking/biting in that 5-s period was scored. The early/acute phase of the formalin test (F_{early}) was defined as 0–10 min post-injection, and the tonic/late phase (F_{late}) as 10–60 min post-injection. Data are presented as the percentage of samples in each phase in which licking/biting was detected. Sample sizes were 8–12/genotype.

2.4.4. Paw-withdrawal test (PW)

Mice were placed on a 3/16th-in. thick glass floor within small (9 × 5 × 5 cm high) Plexiglas cubicles, and a focused high-intensity projector lamp beam was shone from below onto the mid-plantar surface of the hindpaw (Hargreaves et al., 1988). The commercial device (IITC Model 336) was set to either 15% active intensity (PW_{15}) or 20% active intensity (PW_{20}). Latency to withdraw from the stimulus was measured to the nearest 0.1 s. Mice were tested at four time points separated by at least 30 min; at each time point, each hindpaw was tested twice, separated by at least 30 s. Data presented are means of the 16 separate latency determinations, since no repeated measures effects were found with repeated measures ANOVA (data not shown). Sample sizes were 12–16/genotype/intensity.

2.4.5. Tail clip test (TC)

All mice were lightly restrained in a cloth/cardboard holder, and an alligator clip with rubber cuffs around the teeth of each jaw (exerting ≈730 g of force) was applied to the tail 1 cm from the base. The mouse was immediately removed from the holder and the latency to lick, bite, or grab the clip was measured with a stopwatch to the nearest 0.1 s, after which the clip was immediately removed. This test was only performed once, since repeated testing leads to systematic latency alterations (S.B. Smith and J.S. Mogil, unpublished data). Sample sizes were 12–16/genotype.

2.4.6. Tail-withdrawal test (TW)

While lightly restrained in a cloth/cardboard holder, the distal half of the mouse's tail was dipped into a bath of water thermostatically controlled at 47.0 or 49.0 °C (±0.1 °C) (TW_{47} and TW_{49}, respectively). Latency to respond to the heat stimulus by vigorous flexion of the tail was measured. Mice were tested at four time points separated by 30 min; at each time point, two latency determinations (separated by ≈20 s) were made and averaged. Data presented are means of the eight separate latency determinations, since no repeated measures effects were noted. Sample sizes were 12–16/genotype/temperature.

2.4.7. von Frey test (VF)

Mice were tested on the von Frey test using the up-down staircase method of Dixon (Chaplan et al., 1994). Mice were placed on a metal mesh floor within small Plexiglas cubicles (9 × 5 × 5 cm high), and a set of eight calibrated von Frey fibers (ranging from 0.007 to 1.40 g of force) were applied to the plantar surface of the hindpaw until they bowed. The threshold force required to elicit withdrawal of the paw (median 50% paw withdrawal) was determined on two separate days. Data presented are from both left and right hindpaws averaged together since no laterality effects were noted, and from both days averaged together since no repeated measures effects were noted. Sample sizes were 12–16/genotype.

2.4.8. Zymosan thermal hypersensitivity (ZYM_{PW})

Immediately following determination of baseline thermal sensitivity using the radiant heat paw-withdrawal test (active intensity setting = 15%; four baselines on each hindpaw separated by 15 min; see Section 2.4.4), all mice were injected, subcutaneously into the right hindpaw, with a 0.25 mg/ml solution of zymosan (20 μl injection volume) (Meller and Gebhart, 1997). Mice were retested for thermal sensitivity of both hindpaws (two determinations per hindpaw at each time point) at 60, 120, 180, 240, 300 and 360 min post-injection. Data presented are percentages of the maximum possible thermal hypersensitivity over the 6-h testing period calculated with respect to area over the time–effect curve (using the trapezoidal rule). No genotype differences were noted in the time course of zymosan's effects on thermal sensitivity. Sample sizes were 6–10/genotype.

2.4.9. Zymosan mechanical hypersensitivity (ZYM_{VF})

Immediately following determination of baseline mechanical sensitivity using von Frey fibers as described in Section 2.4.7, all mice were injected, subcutaneously into the right hindpaw, with a 0.25 mg/ml solution of zymosan (20 μl injection volume). Mice were retested for mechanical sensitivity of both hindpaws at 60, 120, 180, 240, 300 and 360 min post-injection. Data presented are percentages of the maximum possible mechanical hypersensitivity (i.e., allodynia) over the 6-h testing period calculated with respect to area over the time–effect curve (using the trapezoidal rule). No genotype differences were noted in the time course of zymosan's effects on mechanical sensitivity. Sample sizes were 9–12/genotype.

2.5. Data analysis

Data from 10 mice were excluded after they were identified as statistical outliers (Studentized residual > 3.0). In all assays, genotype differences were evaluated using Student's *t*-test (two-way). The interacting influence of sex, tester, genotype and cohort (i.e., the same strain tested on separate occasions, >6 months apart) was examined in two ways. First, data from C57BL/6 mice – tested on two separate occasions by both testers – were analyzed by three-way ANOVA (sex × tester × cohort). Second, data from the B6.MOLF12 experiment – in which congenics and associated controls were tested separately by both testers – were analyzed by three-way ANOVA (sex × tester × genotype). For all statistical analyses, a criterion of $\alpha = 0.05$ was adopted. To correct for multiple testing the false discovery rate (FDR) method was used (Benjamini and Hochberg, 1995).

From the three-way ANOVAs, we also calculated partial η_p^2, a measure of effect size (strength of association) of each main effect and interaction. It is calculated simply as: $\eta_p^2 = SS_{effect}/(SS_{effect} + SS_{error})$. We used this statistic to calculate the average effect size of sex, tester, cohort and genotype (B6.MOLF12 vs. C57BL/6) across the 13 dependent measures evaluated (see Tables 1 and 2).

3. Results

3.1. Sex differences

As can be seen in Table 1, a main effect of sex was observed in the sensitivity of C57BL/6 mice on the PW_{20}, TW_{49} and VF assays. Strong trends towards significance were seen in the F_{late} and PW_{15} assays, and a weaker trend in the AC test. In every case except for F_{late}, female mice were more sensitive than male mice of this strain, in broad agreement with previous findings from our laboratory using this strain (Mogil, 2003). In three assays (HP_{53}, TW_{47} and VF), sex interacted significantly with tester, such that sex differences were seen when S.G.S. tested but not when J.R. did. One highly significant three-way interaction involving sex was seen in the PW_{20} assay, but in this case the interaction was caused simply by the sex difference being larger in one study than the others.

Similar sex differences were seen when analyzing C57BL/6 vs. congenic strain data sets (see, e.g., Table 2 for B6.MOLF12). Of greater importance, though, are potential sex × genotype interactions, since they would suggest the existence of sex-specific genetic linkages (Mogil, 2003). As can be seen in Table 2, only one such interaction achieved significance in the B6.MOLF12 data set, for HP_{53}, reflecting a much larger difference due to genotype in male vs. female mice. In the B6.MOLF2 data set, three significant (all $p = 0.01$–0.03) sex × genotype interactions were observed, for HP_{53}, TW_{49} and VF; in all cases the genotype difference was larger in males. In the B6.MOLF17 data set, no sex × genotype interactions were significant.

We note that with respect to the major genetic findings of this study – altered F_{late} sensitivity in B6.MOLF12 mice and altered thermal sensitivity in B6.MOLF17 mice (see Section 3.2) – sex appears not to be an important factor.

Table 1
Effects of subject sex, tester and cohort across nociceptive assays using C57BL/6 mice[a]

Assay[b]	n[c]	Sex		Tester		Cohort		S × T		S × C		T × C		S × T × C	
		p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2
AC	23	0.15	0.13	0.49	0.03	0.98	0.00	0.25	0.09	0.72	0.01	0.78	0.01	0.85	0.00
F_{early}	44	0.49	0.01	0.85	0.00	0.85	0.00	0.78	0.00	0.89	0.00	0.52	0.01	0.17	0.05
F_{late}	44	*0.08*	0.08	0.50	0.01	0.56	0.01	0.38	0.02	0.32	0.03	0.80	0.00	0.96	0.00
HP_{50}	43	0.95	0.00	0.32	0.03	0.21	0.04	0.87	0.00	0.28	0.03	0.77	0.00	0.38	0.02
HP_{53}	43	0.40	0.02	**0.02**	0.14	0.20	0.05	**0.03**	0.13	0.11	0.07	**0.01**	0.16	0.89	0.00
PW_{15}	43	*0.08*	0.09	0.43	0.02	0.93	0.00	0.13	0.07	0.56	0.01	0.26	0.04	0.49	0.02
PW_{20}	43	<**0.001**	0.31	0.69	0.01	*0.10*	0.07	0.60	0.01	0.90	0.00	<**0.001**	0.31	<**0.001**	0.40
TC	41	0.52	0.01	0.91	0.00	0.27	0.04	0.19	0.05	0.83	0.00	0.25	0.04	0.68	0.01
TW_{47}	43	0.29	0.03	<**0.001**	0.54	0.68	0.01	**0.05**	0.11	0.99	0.00	0.69	0.01	0.51	0.01
TW_{49}	43	**0.05**	0.11	0.11	0.07	**0.03**	0.13	0.65	0.01	**0.02**	0.15	**0.003**	0.22	0.93	0.00
VF	43	**0.05**	0.11	*0.06*	0.10	**0.04**	0.11	**0.04**	0.12	0.92	0.00	**0.001**	0.28	0.92	0.00
ZYM_{PW}	21	0.30	0.08	0.21	0.12	**0.05**	0.26	0.43	0.05	*0.08*	0.21	*0.08*	0.22	*0.06*	0.24
ZYM_{VF}	32	0.86	0.00	**0.02**	0.21	0.56	0.01	0.98	0.00	0.54	0.02	0.22	0.06	0.29	0.05
Average	39		0.08		0.10		0.06		0.05		0.04		0.10		0.06

[a] Values represent p-values and η_p^2 (partial eta-squared) values from a three-way ANOVA (sex × tester × cohort) performed on data from C57BL/6 mice in all studies. Significant (uncorrected) p-values are in bold; p-values less than or equal to 0.10 but greater than 0.05 are italicized. Cohort levels were the B6.MOLF12 and *not* B6.MOLF12 studies (i.e., B6.MOLF2 study for S.G.S., B6.MOLF17 study for J.R.).

[b] AC, abdominal constriction test; F_{early}, formalin test, early/acute phase; F_{late}, formalin test, late/tonic phase; HP_{50}, 50 °C hot-plate test; HP_{53}, 53 °C hot-plate test; PW_{15}, paw-withdrawal test, 15% maximum intensity; PW_{20}, paw-withdrawal test, 20% maximum intensity; TC, tail clip test; TW_{47}, 47 °C tail-withdrawal test; TW_{49}, 49 °C tail-withdrawal test; VF, von Frey test; ZYM_{PW}, zymosan-induced thermal hypersensitivity (measured on PW_{15} test); ZYM_{VF}, zymosan-induced mechanical hypersensitivity (measured on VF test).

[c] Total C57BL/6 sample size, including roughly equal numbers of males and females combined.

3.2. Genotype effects

Genotype effects were assessed via two-way Student's t-test in data collapsed by sex (and, for B6.MOLF12, collapsed by tester); this is statistically warranted since the data set is balanced with respect to both sex and tester. Significant differences between C57BL/6 and a congenic strain were revealed in the following assays:

Table 2
Effects of subject sex, tester (and/or cohort) and genotype across nociceptive assays: comparison of C57BL/6 and B6.MOLF12 strains[a]

Assay[b]	n[c]	Sex		Tester/cohort		Genotype		S × T/C		S × G		T/C × G		S × T/C × G	
		p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2	p	η_p^2
AC	16	0.47	0.07	0.35	0.11	0.89	0.00	0.71	0.02	0.86	0.00	0.54	0.05	0.47	0. 07
F_{early}	46	0.93	0.00	0.78	0.00	*0.06*	0.09	0.86	0.00	0.60	0.01	0.81	0.00	0.28	0.03
F_{late}	46	0.58	0.01	0.68	0.00	<**0.001**	0.31	0.80	0.00	0.75	0.00	0.58	0.01	0.54	0.01
HP_{50}	32	0.70	0.01	**0.002**	0.33	*0.05*	0.15	0.98	0.00	0.51	0.02	**0.05**	0.15	0.30	0.04
HP_{53}	32	0.11	0.10	0.40	0.03	**0.002**	0.33	0.22	0.06	**0.02**	0.19	0.48	0.02	0.35	0.04
PW_{15}	32	**0.02**	0.21	0.39	0.03	0.12	0.10	0.21	0.06	0.51	0.02	0.25	0.06	**0.005**	0.28
PW_{20}	31	<**0.001**	0.45	**0.01**	0.24	0.38	0.03	0.23	0.06	0.85	0.00	0.65	0.01	0.42	0.03
TC	32	0.55	0.02	**0.003**	0.30	*0.06*	0.14	0.36	0.03	0.30	0.05	**0.04**	0.16	0.96	0.00
TW_{47}	32	0.25	0.05	**0.002**	0.34	0.47	0.02	0.34	0.04	0.91	0.00	**0.007**	0.27	0.12	0.10
TW_{49}	32	**0.003**	0.31	**0.03**	0.18	0.54	0.02	0.11	0.10	0.38	0.03	**0.04**	0.16	0.20	0.07
VF	32	0.54	0.02	0.81	0.00	**0.03**	0.18	**0.02**	0.22	*0.05*	0.15	0.19	0.07	< **0.001**	0.41
ZYM_{PW}	16	0.25	0.16	**0.01**	0.59	0.27	0.15	0.38	0.10	0.39	0.09	0.83	0.01	0.40	0.09
ZYM_{VF}	22	0.88	0.00	**0.05**	0.25	0.99	0.00	0.52	0.03	0.54	0.03	0.93	0.00	0.99	0.00
Average	31		0.11		0.18		0.12		0.06		0.05		0.07		0.09

[a] Values represent p-values and η_p^2 (partial eta-squared) values from a three-way ANOVA (sex × tester × genotype) performed on data from the B6.MOLF12 study only. Significant (uncorrected) p-values are in bold; p-values less than or equal to 0.10 but greater than 0.05 are italicized.

[b] AC, abdominal constriction test; F_{early}, formalin test, early/acute phase; F_{late}, formalin test, late/tonic phase; HP_{50}, 50 °C hot-plate test; HP_{53}, 53 °C hot-plate test; PW_{15}, paw-withdrawal test, 15% maximum intensity; PW_{20}, paw-withdrawal test, 20% maximum intensity; TC, tail clip test; TW_{47}, 47 °C tail-withdrawal test; TW_{49}, 49 °C tail-withdrawal test; VF, von Frey test; ZYM_{PW}, zymosan-induced thermal hypersensitivity (measured on PW_{15} test); ZYM_{VF}, zymosan-induced mechanical hypersensitivity (measured on VF test).

[c] Total sample size (both genotypes combined and both sexes combined).

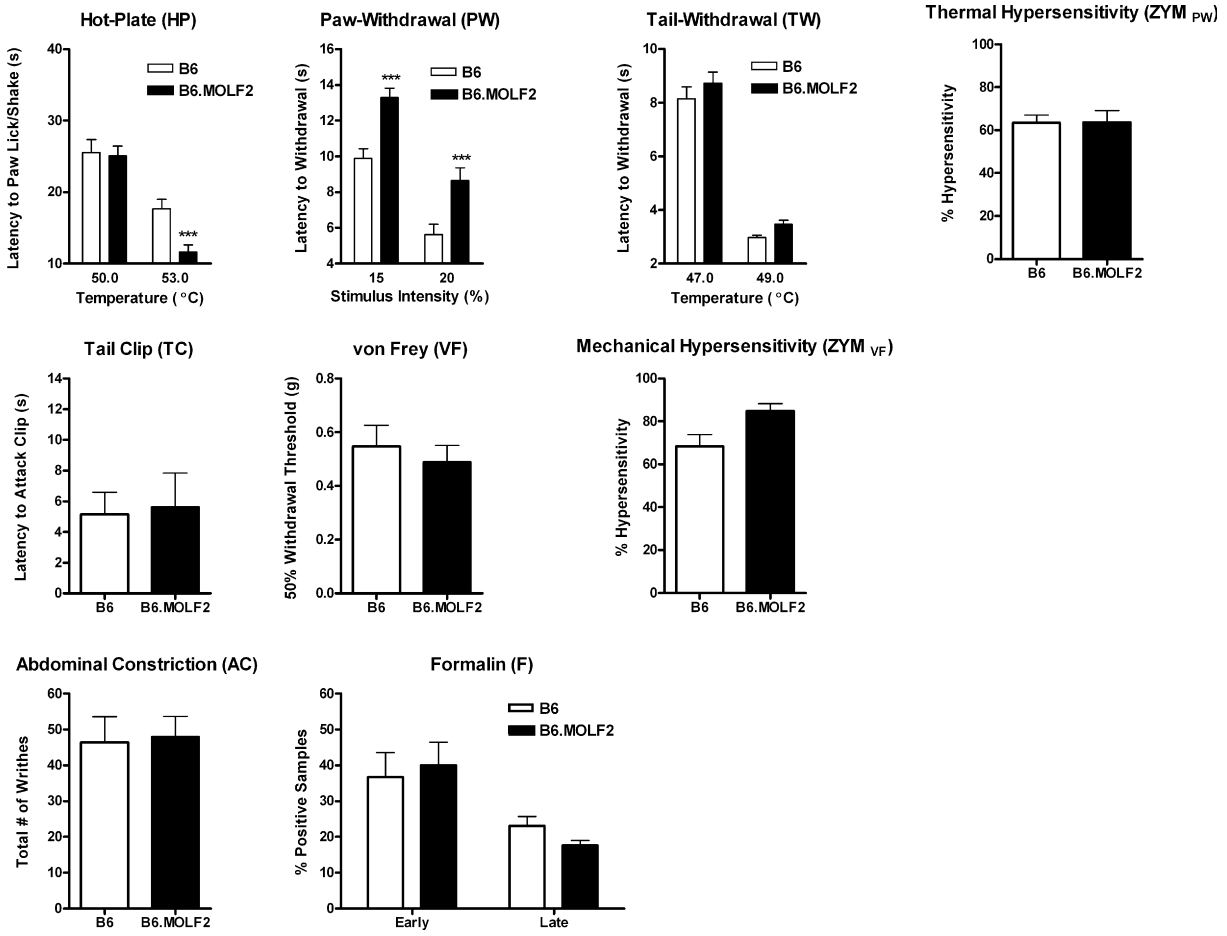

Fig. 2. Sensitivity of C57BL/6 (B6) and B6.MOLF2 mice on nine algesiometric assays. Symbols and bars represent mean ± SEM. Thermal assays (and a test of thermal hypersensitivity) are shown on the top level, mechanical assays (and a test of mechanical hypersensitivity) on the middle level, and chemical/inflammatory assays on the bottom level. All mice were tested concurrently by S.G.S. ***$p < 0.001$ compared to B6.

C57BL/6 vs. B6.MOLF2 (see Fig. 2): HP_{53}, PW_{15}, PW_{20}
C57BL/6 vs. B6.MOLF12 (see Fig. 3): *HP_{53}*, F_{late}
C57BL/6 vs. B6.MOLF17 (see Fig. 4): *HP_{50}*, HP_{53}, PW_{20}, TW_{47}, *TW_{49}*

Differences that survived FDR correction at $p < 0.05$ are noted above in regular type; those that do not are shown in italics.

3.3. Tester effects

Previous work from our laboratory revealed tester as the largest source of trait variability in the TW_{49} assay (Chesler et al., 2002a,b). Inspection of both tables reveals that TW_{49} is hardly unique in this respect, with main effects of tester and tester × genotype interactions observed in multiple assays. In fact, tester featured the largest η_p^2 statistic of any main effect or interaction in both tables. However, the multiple instances of tester × cohort interactions in Table 1 suggest that main effects of tester in Table 2 may at least partially represent cohort effects. With the present design we cannot separate these two possibilities.

3.4. Higher-order interactions

Tables 1 and 2 reveal a number of two-way and even three-way interactions not specifically mentioned above. Their existence is intriguing, but such higher-order interactions should be treated with great caution and are generally non-interpretable.

4. Discussion

4.1. Identification of novel QTLs using congenic strains

The finding of a significant difference in a phenotypic mean between C57BL/6 and a congenic strain formed on a C57BL/6 background, assuming all mice are bred and tested concurrently and identically, represents *prima facie* evidence of the donor genomic region as a QTL for the phenotype in question. This, in turn, implies the

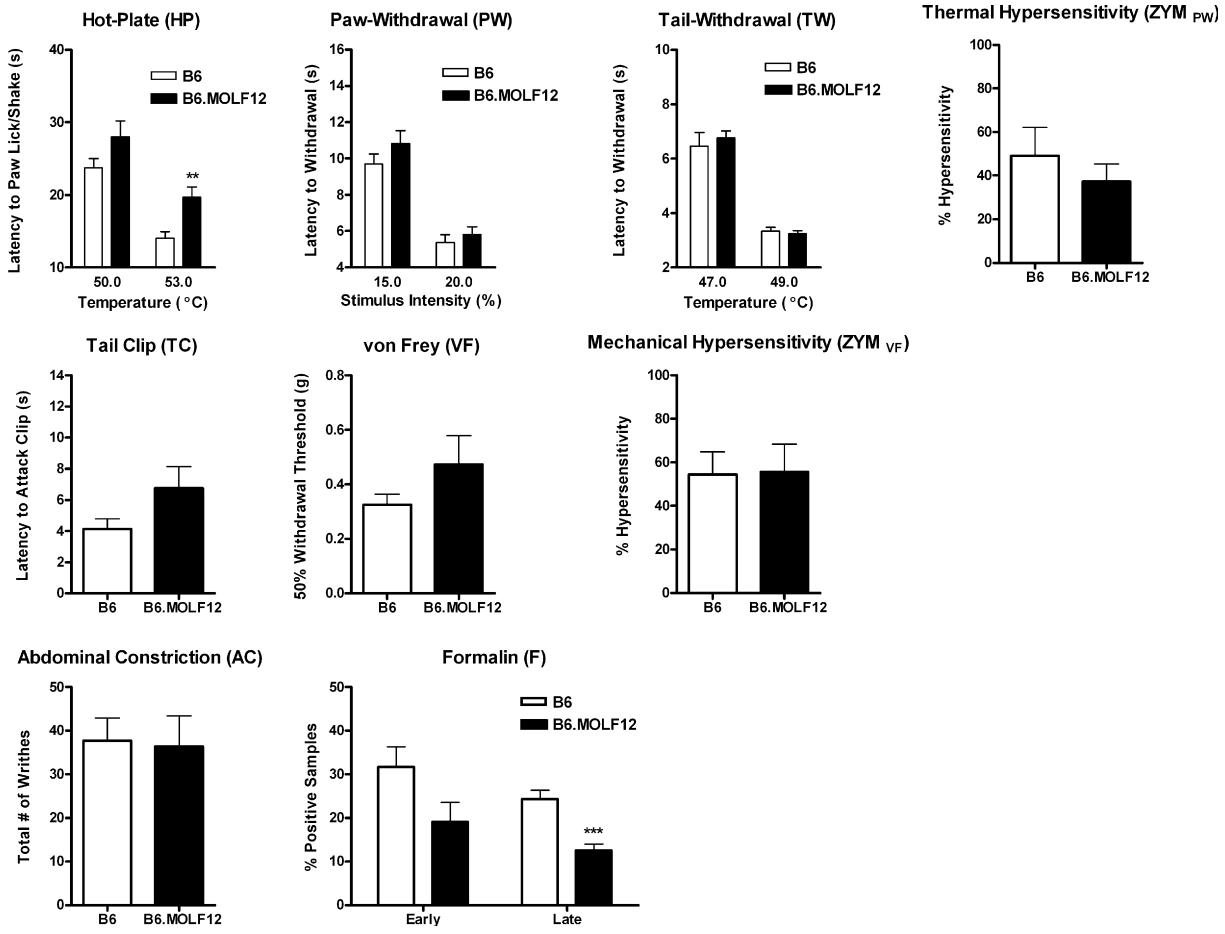

Fig. 3. Sensitivity of C57BL/6 (B6) and B6.MOLF12 mice on nine algesiometric assays. Symbols and bars represent mean ± SEM. Thermal assays (and a test of thermal hypersensitivity) are shown on the top level, mechanical assays (and a test of mechanical hypersensitivity) on the middle level, and chemical/inflammatory assays on the bottom level. Data shown are combined means of mice tested in separate cohorts by J.R. and S.G.S. (see Table 2). $^{**}p < 0.01$ compared to B6 (not significant after FDR correction). $^{***}p < 0.001$ compared to B6.

existence of a gene or genes within this genomic region whose allelic DNA sequence variants are causal to the strain difference. The congenic strategy offers analytical transparency, since significance can be evaluated by *t*-test instead of complex interval mapping methodologies. Also, with this strategy the influence of epistasis (i.e., gene–gene interactions) is minimized or abolished, and the size of the genetic effect is obvious. Finally, the congenic strains themselves serve as the first step towards positional cloning of the responsible gene(s).

We have herein identified at least two new and robust QTLs, one for thermal nociception on Chr 17 (dubbed *Tpnr3*, since two other "*t*hermal *p*ain *r*esponse" QTLs have been described; Mogil et al., 1997; Mogil et al., 2005b) and one for formalin test sensitivity on Chr 12 (dubbed *Nociq3*, again the third "*noc*iception, *i*nflammatory *Q*TL" so described; Wilson et al., 2002). We are less sure of the existence of a QTL for thermal nociception on Chr 2, as discussed below. Efforts are well underway towards the positional cloning of *Tpnr3* and *Nociq3*. We have constructed subcongenic mice, in which the congenic regions of B6.MOLF17 and B6.MOLF12, respectively, have been further reduced by the detection and breeding of appropriate recombinant individuals. This effort has allowed the genomic region containing the responsible gene(s) to be reduced to only several million basepairs (data not shown), containing as few as 17 known or predicted genes. Currently, we are using sequencing, microarray and quantitative RT-PCR approaches to decide among the remaining candidates. We note, though, that in both cases there are no genes in the linked region that have ever been proposed as being relevant to pain processing or modulation. This fact bolsters our contention that "blind" linkage mapping is an excellent heuristic strategy for pain research. Furthermore, it suggests there may be many more "pain genes" in the genome than those already suspected by virtue of their protein's known involvement in pain processing.

In theory, this type of analysis could be performed all over the genome using existing, commercially available congenic strains and transgenic knockout mice, which

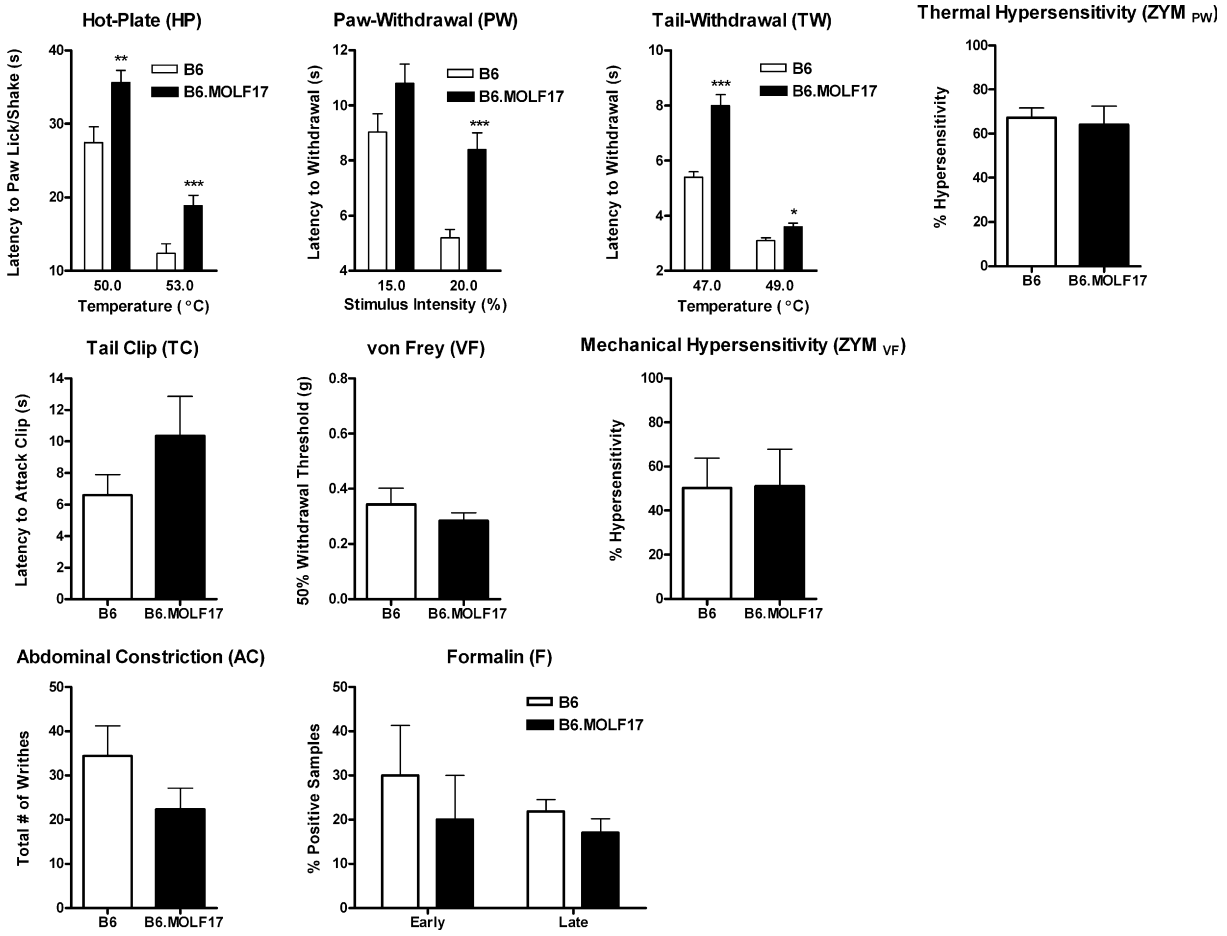

Fig. 4. Sensitivity of C57BL/6 (B6) and B6.MOLF17 mice on nine algesiometric assays. Symbols and bars represent mean ± SEM. Thermal assays (and a test of thermal hypersensitivity) are shown on the top level, mechanical assays (and a test of mechanical hypersensitivity) on the middle level, and chemical/inflammatory assays on the bottom level. All mice were tested concurrently by J.R. *$p < 0.05$ compared to B6 (not significant after FDR correction). **$p < 0.01$ compared to B6 (not significant after FDR correction). ***$p < 0.001$ compared to B6.

are virtually all *de facto* congenic strains with a 129-derived donor segment surrounding the null mutation placed on a C57BL/6 background (Bolivar et al., 2001). Bolivar and colleagues (2001) have described a simple method for distinguishing whether phenotypic differences between knockout mice and their wild-type controls are due to the targeted gene or nearby "hitchhiking" genes.

4.2. The tester/cohort effect in basic pain research

The present study provides evidence for robust and widespread tester and/or cohort effects on algesiometric assays, in a manner similar to that we described previously for the TW_{49} test (Chesler et al., 2002a,b). These effects are perhaps an instantiation of the phenomenon demonstrated by Crabbe and colleagues (1999), who after going to heroic lengths to standardize mouse acquisition, husbandry, testing equipment and testing procedures still observed considerable genotype × laboratory environment interactions on a range of behavioral tests. In our case, however, the variability is entirely *within*-lab.

Although a main effect of tester or cohort is not necessarily worrisome as long as appropriate controls are included in experimental designs, tester/cohort × genotype interactions present a more serious problem. Whether or not a genotype "effect" is uncovered (e.g., wild-type vs. knockout) may depend on particularities of the tester and/or laboratory environment at the time of testing (e.g., season/humidity, personnel changes, noise levels). In the present study, the marginally significant genotypic differences between B6.MOLF12 and C57BL/6 in the HP_{50}, TC and even VF assays (see Table 2 and Fig. 3) were likely confounded by interactions with tester, and thus cannot be considered particularly robust. Since rather low numbers of subjects are typically tested (see below), and these subjects may derive from one or two breeding pairs at most, the cohort effect may actually represent a so-called litter effect (Blizard, 1992), in which variation in maternal care, for example, could influence adult anxiety levels (Francis et al., 1999). The

only real solution to the problem is to increase sample sizes and test more extensively, but one must trade off sample size within-test with the number of different tests performed. A partial solution, as we have used here, is the concurrent testing of a "standard strain" (in this case, C57BL/6) (Blizard et al., 2005). More generally, the obvious solution to tester/cohort confounds is to have a large number of mice, from many different litters, tested at one time by a single person. Given that this will often be practically impossible, we recommend strongly the analysis of these sorts of data sets using tester and cohort as statistical covariates.

What might explain the large effect of tester on algesiometric data collected presently? One obvious possibility is tester experience: J.R. had far more experience using these assays than S.G.S. However, an examination of the coefficients of variation (CVs) of the C57BL/6 vs. B6.MOLF12 data sets collected by both testers revealed that S.G.S. actually produced "tighter" data (i.e., with lower CVs) than J.R. in eight of 13 measures. It is interesting to note that tester effects and interactions were *not* limited to those assays involving direct handling of mice during or immediately prior to data collection itself (HP, TC, and TW).

4.3. Composition and power of nociceptive test batteries

The present data make obvious the value of a systematic approach to phenotyping mice for pain. Although congenic strains were evaluated here, we expect that these protocols will be mostly utilized for the testing of transgenic knockout mice vis-à-vis wild-type controls. The assays in the test battery used here represent nine of the top 10 assays by frequency in transgenic studies of pain (see Fig. 5a). We believe the full battery is useful in that each of three fundamental nociceptive modalities (thermal, mechanical, and chemical/inflammatory) is represented by at least two assays (HP/PW/TW, TC/VF, AC/F, respectively). Although for our purposes the use of an inflammatory stimulus to induce mechanical and thermal hypersensitivity (i.e., allodynia or hyperalgesia) was appropriate, a neuropathic injury could replace or supplement our use of zymosan. The use of inflammatory assays features the advantage that the degree of hindpaw edema can be quantified and used as a covariate and to identify "bad" injections.

The redundancy built into the battery can be useful for interpreting findings. The consistently reduced sensitivity of B6.MOLF17 across three thermal assays and multiple stimulus intensities gives us very high confidence that a gene (or genes) in the MOLF-derived region of mouse Chr 17 play an important role in thermal nociception generally. By contrast, consider the data obtained in thermal assays in the B6.MOLF2 congenic strain. If only PW was used, it would be confidently concluded that B6.MOLF2 mice had reduced

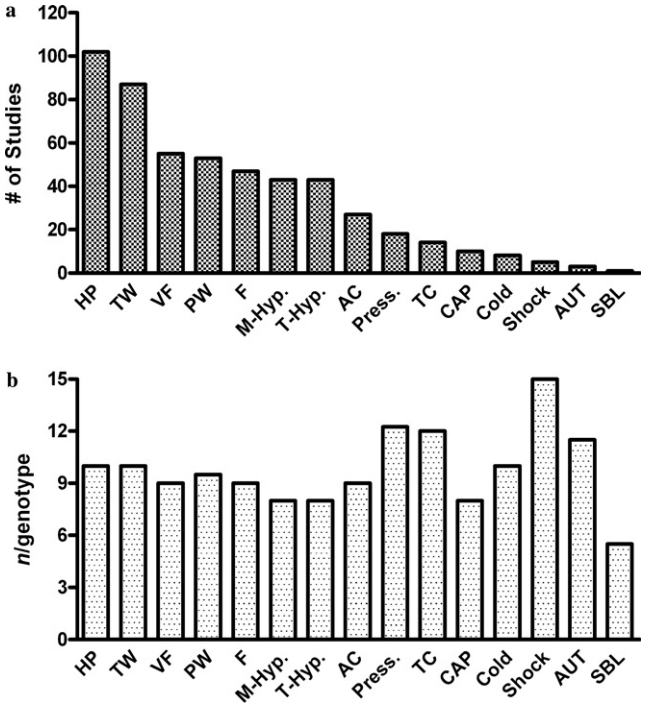

Fig. 5. Assay and sample size choices in the transgenic knockout mouse pain literature. Data were obtained by careful inspection of the Methods and results sections of 225 published papers (1996–2005) in which awake, behaving transgenic knockout mice were compared (at least) to wild-type controls on an algesiometric assay. Papers, the titles of which are provided in an Appendix on-line, were part of the hardcopy collection of the first author (as of August 4, 2005), and likely represent a large majority of available papers on this topic at that date. Graph (a) illustrates the frequency with which each assay was utilized (total: 516, an average of 2.4 assays/paper). Graph (b) illustrates the median sample size (*n*) per genotype in each assay, averaged across genotype. Assay abbreviations are as follows, from left to right in graph (a): HP, hot-plate test; TW, tail-withdrawal test (from radiant heat or hot water); VF, von Frey test; PW, paw-withdrawal test from radiant heat; F, formalin test; M-Hyp., mechanical hypersensitivity (i.e., allodynia, induced by neuropathy or inflammation, typically measured on the VF test); T-Hyp., thermal hypersensitivity (i.e., hyperalgesia, induced by neuropathy or inflammation, typically measured on the PW test); AC, abdominal constriction test (to any intraperitoneal irritant); Press., threshold to withdrawal of foot or tail from pressure (Randall–Selitto test); TC, tail-clip/pinch test (i.e., latency to respond to a suprathreshold mechanical stimulus); CAP, capsaicin-induced licking; Cold, response to cold stimulus (acetone, cold water or cold plate); Shock, threshold response (vocalization, flinch or jump) from electric footshock; AUT, autotomy behavior; SBL, scratching–biting–licking to intrathecally injected algogen.

sensitivity. However, no significant differences were observed on the TW_{47}, TW_{49} and HP_{50} assays, and in the HP_{53} assay this congenic strain was significantly *more* sensitive. Either genes exist in the MOLF-derived region of mouse Chr 2 that play a complex role in thermal nociception – perhaps only in supraspinal pathways or at higher stimulus intensities – or these significant differences were obtained by chance. The single significant phenotypic difference of B6.MOLF12 congenics from C57BL/6 in the thermal assays, in HP_{53}, might similarly be disregarded.

Especially in the "single test" assays (AC, HP, F, TC, ZYM_{PW}, an ZYM_{VF}), one might wonder whether sufficient statistical power exists to detect differences between genotypes when they exist. We restricted ourselves presently to sample sizes that are reflective of those actually used in the current transgenic mouse literature, $n = 6$–16/ genotype (see Fig. 5b). The median sample size in most of these assays (except for suprathreshold mechanical assays, which feature higher variability) appears to be $n = 9$/genotype. Given the large number of barely significant genotypic differences and trends towards differences observed in these studies, we wonder whether such sample sizes are high enough.

5. Conclusions

There is no doubt that transgenic knockout and genetic linkage mapping studies are of considerable utility to the furthering of knowledge of the basic science of pain; hopefully, they will be of value clinically at some point in the near future. This particular study pointed to the existence and genomic location of at least two more genes affecting variable nociceptive sensitivity in mice. It is also clear, however, that the usefulness of mouse molecular genetics depends critically on the strength of the behavioral testing, a topic that has been given far less attention in pain research than it deserves.

Acknowledgements

Thanks to Dr. Elissa J. Chesler and the anonymous reviewers for helpful comments on the manuscript. This work was carried out with the support of NIH Grant DA15191, the Canada Foundation for Innovation and Canada Research Chair Programs (J.S.M.), and a CIHR operating grant and New Investigator Award (A.K.N.). A.K.N. was supported by the John R. and Clara M. Fraser Memorial Award of the McGill University Faculty of Medicine.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.pain.2006.06.004.

References

Bailey DW. Strategic uses of recombinant inbred and congenic strains in behavior genetics research. In: Gershon ES, Matthysse S, Breakefield XO, Ciaranello RD, editors. Genetic research strategies for psychobiology and psychiatry. Pacific Grove, CA: Boxwood Press; 1981. p. 189–98.

Belfer I, Wu T, Kingman A, Krishnaraju RK, Goldman D, Max MB. Candidate gene studies of human pain mechanisms: a method for optimizing choice of polymorphisms and sample size. Anesthesiology 2004;100:1562–72.

Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B 1995;57:289–300.

Bennett B. Congenic strains developed for alcohol- and drug-related phenotypes. Pharmacol Biochem Behav 2000;67:671–81.

Blizard DA. Recombinant-inbred strains: general methodological considerations relevant to the study of complex characters. Behav Genet 1992;22:621–33.

Blizard DA, Wada Y, Onuki Y, Kato K, Mori T, Taniuchi T, et al. Use of a standard strain for external calibration in behavioral phenotyping. Behav Genet 2005;35:323–32.

Bolivar VJ, Cook MN, Flaherty L. Mapping of quantitative trait loci with knockout/congenic strains. Genome Res 2001;11:1549–52.

Chaplan SR, Bach FW, Pogrel JW, Chung JM, Yaksh TL. Quantitative assessment of tactile allodynia evoked by unilateral ligation of the fifth and sixth lumbar nerves in the rat. J Neurosci Methods 1994;53:55–63.

Chesler EJ, Wilson SG, Lariviere WR, Rodriguez-Zas SL, Mogil JS. Identification and ranking of genetic and laboratory environment factors influencing a behavioral trait, thermal nociception, via computational analysis of a large data archive. Neurosci Biobehav Rev 2002a;26:907–23.

Chesler EJ, Wilson SG, Lariviere WR, Rodriguez-Zas SL, Mogil JS. Influences of laboratory environment on behavior. Nat Neurosci 2002b;5:1101–2.

Crabbe JC, Wahlsten D, Dudek BC. Genetics of mouse behavior: interactions with laboratory environment. Science 1999;284:1670–2.

Crawley JN, Paylor R. A proposed test battery and constellations of specific behavioral paradigms to investigate the behavioral phenotypes of transgenic and knockout mice. Horm Behav 1997;31:197–211.

Croteau S, Roquis D, Charron MC, Frappier D, Yavin D, Loredo-Osti JC, et al. Increased plasticity of genomic imprinting of *Dlk1* in brain is due to genetic and epigenetic factors. Mamm Genome 2005;16:127–35.

Fillingim RB, Kaplan L, Staud R, Ness TJ, Glover TL, Campbell CM, et al. The A118G single nucleotide polymorphism of the μ-opioid receptor gene (OPRM1) is associated with pressure pain sensitivity in humans. J Pain 2005;6:159–67.

Francis D, Diorio J, Liu D, Meaney MJ. Nongenomic transmission across generations of maternal behavior and stress responses in the rat. Science 1999;286:1155–8.

Hargreaves K, Dubner R, Brown F, Flores C, Joris J. A new and sensitive method for measuring thermal nociception in cutaneous hyperalgesia. Pain 1988;32:77–88.

Koide T, Moriwaki K, Ikeda K, Niki H, Shiroishi T. Multi-phenotype behavioral characterization of inbred strains derived from wild stocks of *Mus musculus*. Mamm Genome 2000;11:664–70.

Lariviere WR, Wilson SG, Laughlin TM, Kokayeff A, West EE, Adhikari SM, et al. Heritability of nociception. III. Genetic relationships among commonly used assays of nociception and hypersensitivity. Pain 2002;97:75–86.

McIlwain KL, Merriweather MY, Yuva-Paylor LA, Paylor R. The use of behavioral test batteries: effects of training history. Physiol Behav 2001;73:703–17.

Meller ST, Gebhart GF. Intraplantar zymosan as a reliable, quantifiable model of thermal and mechanical hyperalgesia in the rat. Eur J Pain 1997;1:43–52.

Mogil JS. Interaction between sex and genotype in the mediation of pain and pain inhibition. Semin Pain Med 2003;1:197–205.

Mogil JS. editor. The genetics of pain. Progress in pain research and management, Vol. 28. IASP Press, Seattle, WA, 2004, 349 pp.

Mogil JS, Breese NM, Witty M-F, Ritchie J, Rainville M-L, Ase A, et al. Transgenic expression of a dominant-negative ASIC3 subunit leads to increased sensitivity to mechanical and inflammatory stimuli. J Neurosci 2005a;25:9893–901.

Mogil JS, Meirmeister F, Seifert F, Strasburg K, Zimmermann K, Reinold H, et al. Variable sensitivity to noxious heat is mediated by differential expression of the CGRP gene. Proc Natl Acad Sci USA 2005b;102:12938–43.

Mogil JS, Ritchie J, Smith SB, Strasburg K, Kaplan L, Wallace MR, et al. Melanocortin-1 receptor gene variants affect pain and μ-opioid analgesia in mice and humans. J Med Genet 2005c;42:583–7.

Mogil JS, Chanda ML. The case for the inclusion of female subjects in basic science studies of pain. Pain 2005;117:1–5.

Mogil JS, McCarson KE. Finding pain genes: bottom-up and top-down approaches. J Pain 2000;1(Suppl. 1):66–80.

Mogil JS, Richards SP, O'Toole LA, Helms ML, Mitchell SR, Belknap JK. Genetic sensitivity to hot-plate nociception in DBA/2J and C57BL/6J inbred mouse strains: possible sex-specific mediation by δ_2-opioid receptors. Pain 1997;70:267–77.

Mogil JS, Wilson SG, Bon K, Lee SE, Chung K, Raber P, et al. Heritability of nociception. I. Responses of eleven inbred mouse strains on twelve measures of nociception. Pain 1999a;80:67–82.

Mogil JS, Wilson SG, Bon K, Lee SE, Chung K, Raber P, et al. Heritability of nociception. II. "Types" of nociception revealed by genetic correlation analysis. Pain 1999b;80:83–93.

Mogil JS, Wilson SG, Chesler EJ, Rankin AL, Nemmani KVS, Lariviere WR, et al. The melanocortin-1 receptor gene mediates female-specific mechanisms of analgesia in mice and humans. Proc Natl Acad Sci USA 2003;100:4867–72.

Risch NJ. Searching for genetic determinants in the new millennium. Nature 2000;405:847–56.

Voikar V, Vasar E, Rauvala H. Behavioral alterations induced by repeated testing in C57BL/6J and 129S2/Sv mice: implications for phenotyping screens. Genes Brain Behav 2004;3:27–38.

Wade CM, Kulbokas Jr EJ, Kirby AW, Zody MC, Mullikin JC, Lander ES, et al. The mosaic structure of variation in the laboratory mouse genome. Nature 2002;420:574–8.

Wilson SG, Chesler EJ, Hain HS, Rankin AL, Schwarz JZ, Call SB, et al. Identification of quantitative trait loci for chemical/inflammatory nociception in mice. Pain 2002;96:385–91.

Wilson SG, Mogil JS. Measuring pain in the (knockout) mouse: big challenges in a small mammal. Behav Brain Res 2001;125:65–73.