

Quantitative traits for the tail suspension test: automation, optimization, and BXD RI mapping

Heena V. Lad · Lin Liu · José L. Payá-Cano ·
Cathy Fernandes · Leonard C. Schalkwyk

Received: 12 February 2007 / Accepted: 17 April 2007 / Published online: 7 July 2007
© Springer Science+Business Media, LLC 2007

Abstract Immobility in the tail suspension test (TST) is considered a model of despair in a stressful situation, and acute treatment with antidepressants reduces immobility. Inbred strains of mouse exhibit widely differing baseline levels of immobility in the TST and several quantitative trait loci (QTLs) have been nominated. The labor of manual scoring and various scoring criteria make obtaining robust data and comparisons across different laboratories problematic. Several studies have validated strain gauge and video analysis methods by comparison with manual scoring. We set out to find objective criteria for automated scoring parameters that maximize the biological information obtained, using a video tracking system on tapes of tail suspension tests of 24 lines of the BXD recombinant inbred panel and the progenitor strains C57BL/6J and DBA/2J. The maximum genetic effect size is captured using the highest time resolution and a low mobility threshold. Dissecting the trait further by comparing genetic association of multiple measures reveals good evidence for loci involved in immobility on chromosomes 4 and 15. These are best seen when using a high threshold for immobility, despite the overall better heritability at the lower threshold. A second trial of the test has greater duration of immobility and a completely

different genetic profile. Frequency of mobility is also an independent phenotype, with a distal chromosome 1 locus.

Introduction

The tail suspension test (TST) is used to screen for antidepressant activity in pharmacologic studies (Cryan et al. 2005; Steru et al. 1985, 1987). When a mouse is suspended by its tail, the initial response is to struggle, but this is followed by episodes of immobility, which is taken as an index of its depressive state, paralleling behavioral despair. Duration of immobility is found to be markedly reduced in mice administered antidepressants, showing predictive validity in the TST (Cryan et al. 2005; El Yacoubi et al. 2003; Vaugeois et al. 1996). Robust strain differences in baseline TST have been observed, although the strain rankings have varied (Liu and Gershenfeld 2001; Ripoll et al. 2003; Trullas et al. 1989). Studies such as these strongly point to an underlying genetic basis for immobility and so are potentially useful in investigating genes that are accountable for this phenotype. Baseline TST response and TST response to imipramine, along with measures from the open field test and the light-dark box have been explored across 12 inbred strains (Liu and Gershenfeld 2003) using factor analysis. It was found that baseline TST and imipramine response loaded independently and may thus be genetically independent, but this interpretation is complicated by the fact that the imipramine measurements were taken on a second TST trial, which they observed has greater immobility than the first trial.

Methods for automation of scoring using strain gauges or video analysis have been presented and validated by comparison with hand coding and detection of strain dif-

Electronic supplementary material The online version of this article (doi: 10.1007/s00335-007-9029-1) contains supplementary material, which is available to authorized users.

H. V. Lad · L. Liu · J. L. Payá-Cano · C. Fernandes ·
L. C. Schalkwyk (✉)
Social, Genetic and Developmental Psychiatry Centre (PO82),
Institute of Psychiatry, King's College London,
London SE5 8AF, United Kingdom
e-mail: spjglcs@iop.kcl.ac.uk

H. V. Lad
e-mail: Heena.Lad@iop.kcl.ac.uk

ferences and drug responses (Juszczak et al. 2006; Steru et al. 1987). Hand coding of course has an element of subjectivity and there are widely differing criteria in use. Automated methods are not entirely free of interrater subjectivity either because the settings for automated measurement of mobility are selected by the experimenter. Furthermore, the precise phenotypic measure may vary depending on the experimental setup which will affect details of video taping and, thus, analysis. Inconsistently measuring behavioral phenotypes in the TST within and across laboratories may reduce the validity of phenotyping data and, in turn, their value for genetic analyses. Most of these issues are not fully overcome with the use of an automated system where mobile behaviors are still subjectively predetermined; however, they offer the advantage of greater flexibility during analysis that enables us to better optimize and more accurately capture the phenotypes reflected in our behavioral data set. A broader series of data points may be evaluated because many of the settings can be altered during analysis, and so it is possible to look more specifically at a range of values that could be attributed to the immobility phenotype. These features should reduce interrater bias and avoid the need to use arbitrary cutoffs. Optimization criteria, other than correlation with hand-coding results, include magnitude of strain difference or drug responses (Juszczak et al. 2006). It is likely that different genetic loci affecting TST might differ in their behavioral profile, motivating us to couple scoring optimization with genetic dissection of the trait.

Several studies have investigated the loci that underlie the baseline immobility phenotype. In a cross of the DeFries High and Low open-field selection strains, Turri et al. (2001) detected loci on chromosomes 3, 5, 11, and 19 affecting their TST measure. Yoshikawa et al. (2002) looked at QTLs involved in immobility and single-QTL analysis demonstrated suggestive linkage for immobility on chromosomes 4, 8, and 14, with borderline significant linkage on chromosome 11 in a F₂ population of mice derived from C57BL/6 and C3H/He inbred strains. In a BALB/cJ × A/J cross, Crowley et al. (2006) detected loci on chromosomes 7, 12, and 19. Each of these studies used a different combination of inbred strains so it is not surprising that at least some different polymorphic loci are found in each case. Nonetheless, the populations from these four studies are all ultimately derived from combinations of strains C57BL/6, A, C3H, and BALB/c, and the lack of overlap is nearly complete, with only the chromosome 19 locus possibly coinciding in Turri et al. (2001) and Crowley et al. (2006). Using a cross between inbred strains NMRI and I29S6, chosen for maximal TST difference, Liu et al. (2006) detected loci on chromosomes 4, 5, 12, and 18 for basal-line TST. The chromosome 4 locus coincides with that from Yoshikawa et al. (2002). A different, more

proximal peak on chromosome 4 is seen with the related tail suspension-induced hyperthermia phenotype.

We performed a detailed analysis of immobility in the TST, which was the final test performed as part of a battery of nine behavioral tests. Taking advantage of naturally occurring behavioral variation in well-established inbred strains and genetic reference populations, the mice selected in this study included males from the BXD recombinant inbred (RI) panel which is derived from a cross between C57BL/6J and DBA/2J. An RI panel is essentially an F₂ cross immortalized by inbreeding. These allow convenient linkage mapping, using existing genotype data and multiple animals of each genotype. In particular, they are useful for investigating complex traits and preliminary mapping of quantitative trait loci (QTLs). Furthermore, direct estimates of (broad-sense) heritability can be made, which is very useful in optimizing trait measures for genetic mapping and genetic correlation, making them a key reagent for integrating diverse phenotypic data, including molecular phenotypes (Bystrykh et al. 2005; Chesler et al. 2005; Manly et al. 2005; Plomin et al. 2005).

Methods

Animals

Male C57BL/6J ($n = 16$), DBA/2J ($n = 9$), and BXD RI (24 lines, $n = 204$) mice were generated in the Comparative Biology Unit animal facilities at the Institute of Psychiatry using original stocks purchased from The Jackson Laboratory (Bar Harbor, ME, USA). Mice were weaned at 3 weeks of age and transferred at approximately 8 weeks of age to a separate housing facility where they were singly housed and habituated for 2 weeks before undergoing a battery of behavioral tests. Following the habituation period, the mean (\pm standard deviation [SD]) age of all mice was 79.7 ± 15.8 days. Animals were tested in four batches and all efforts were made to minimize the within- and between-batch variability of the RI lines in terms of age and numbers/strain tested. The mean (\pm SD) of the RI lines reported in the current study was 8.5 ± 4.24 mice per line.

Housing conditions

All mice were singly housed in standard cages measuring $30.5 \times 13 \times 11$ cm, with food (Rat & Mouse No. 1 Maintenance Diet, Special Diet Services, Essex, UK) and water available *ad libitum*. The housing room was maintained on a reversed 12:12 light cycle with white lights on from 20:00 to 8:00 hours and red light on in the dark cycle, and all behavioral tests were performed between 09:30 and 19:00 hours. Light intensity in the housing room was 400

lux during the lights-on period and less than 2 lux during the dark period. Four red cluster lights (LED cluster red light No. 310-6757; RS Components Northants, UK) of approximate wavelength 705 nm provided minimal red light during the dark phase, allowing experimenters to work with the mice during their dark phase. Ambient temperature in all rooms was maintained at $21 \pm 2^\circ\text{C}$ with 45% humidity level. Sawdust and nesting materials in each cage were changed once a week, but never on the day before or the day of testing to minimize the disruptive effect of cage cleaning on behavior. All housing and experimental procedures were performed in accordance with the UK Home Office Animals (Scientific Procedures) Act 1986.

Apparatus and experimental procedures

The behavioral tests were conducted in succession as follows: home cage activity, open field, novel object exploration, elevated plus maze, light/dark box, primary screen of SHIRPA, puzzle box, Morris water maze, and tail suspension test (partially described in Galsworthy et al. 2002, 2005; Mill et al. 2002). To circumvent the effect of inter-experimenter variability, the same experimenter performed TST in all mice. All mice were tested in a randomized order. Two separate trials were performed on each mouse at approximately the same time of day on consecutive days, with a minimum of 24 h between trials. Mice were moved to the behavioral suite adjacent to the housing room immediately before testing. Lighting under test conditions was set to 350 lux.

A cord (3 mm diameter) was extended and secured between two legs of an upturned chair, at least 30 cm height from the base. A cardboard cone was placed around the tail of each mouse immediately before the test, with its tail extending through the tip, to prevent tail climbing behaviors. The mouse was suspended at approximately one-third from the end of its tail, using soft padding around the area to protect the tail, and plastic clothes pegs were used to secure the mouse to the line. Each trial was 5 min long and recorded on videotape for further detailed analysis. Following each trial the mouse was returned to the housing room.

Analysis

Using the mobility detection module in EthoVision version 3.1 (Noldus Information Technology bv, Wageningen, The Netherlands; <http://www.noldus.com/site/doc200403002>), which is an automated tracking system, the video recordings collected for each trial of the TST were tracked and analyzed. The main variables for mobility detection in EthoVision are the difference in pixels between current and

previous samples detected, an averaging factor that is used for smoothing, and the thresholds assigned for mobility (Noldus Information Technology bv, 2005).

The subtraction method of object detection was used with a sample rate of five samples per second and an averaging factor of one. The threshold defined for an immobile posture to be assumed (percentage change in object area) was 12% and thus the mouse scored mobile above this threshold. During analysis, the settings used in tracking could be manipulated to look at different thresholds and running average values across the collected data set. To assess the validity of the automated method of scoring behaviors within the TST, one in five mice were simultaneously hand-coded, generating a smaller subset of data. A small subset of video footage (36 animals from batch 3) had been recorded at a different scale. For these the threshold parameter was adjusted by an empirical factor of two to make the results comparable with the remaining data.

Genotyping

Strain identity of all animals was verified by genotyping 11 unlinked SNP markers (rs13475902, rs13475988, rs13459051, rs13459052, rs13459060, rs13476554, rs13459069, rs13478483, rs13459109, rs3708840, rs13482131) across seven chromosomes that distinguish the BXD lines.

Statistics

Batch differences were removed by regression. The strain composition of the batches was not constant, so it is not guaranteed that the (true) batch means are equal. Nonetheless, regression is conservative because it will remove some strain differences when they are partially confounded with batch differences but it is unlikely to create spurious differences. Estimates of genetic effect size were calculated as $SS_{\text{strain}}/(SS_{\text{strain}} + SS_{\text{residual}})$, where SS are sums of squared deviations calculated using the *lm* and *anova* functions from the *stats* package of the R statistical environment (R Development Core Team 2006). Pearson's correlations were calculated using STATA version 9. Dependent-samples *t* test reported were calculated using the STATISTICA analysis tool.

WebQTL analysis

Strain means and variances were calculated across immobile measures for each trial. These were entered into WebQTL, which is a resource for analysis of RI data with databases of genotype, phenotype, and gene expression

data (Wang et al. 2003), for interval mapping and phenotypic correlation.

R/qtl analysis

The same data as were uploaded to WebQTL was also genetically mapped using R/qtl (Broman et al. 2003), with genotype data downloaded from Genenetwork.org, filtered to retain 801 markers with unique strain distribution patterns. The scanone function was used, with default settings (method = ‘em’, model = ‘normal’). The resulting table of LOD scores was plotted using the image function (R graphics package), to produce Fig. 4 and Supplementary Fig. 1, and the persp function used to generate Supplementary Fig. 2.

Results

Measures of immobility

The mobility module in EthoVision generates data for frequency and duration measures in the analysis profile. Trait values for duration and frequency of immobility demonstrate the variation within and between strains (Fig. 1) and transgression in some lines. We investigated the effects of changing the mobility threshold from 12% to

20% and looked at these across a range of running average intervals (1, 5, and 10). Figure 2 illustrates the effect of the threshold and averaging parameters on frequency and duration of immobility in the population as a whole. The mean frequency and the variance in frequency depend strongly on the averaging interval (decreasing as the average interval increases) and less so on the threshold. In contrast, mean duration of immobility depends on the threshold rather than the running average interval, and its variance increases with increasing averaging interval. Similar patterns were apparent in both trials for each quantitative measure.

Automated versus manual scoring

A fifth of the data set was scored manually concurrently with the automated system to determine how closely the automated scores reflect manually observed behaviors. For duration of immobility, Pearson’s correlations were positive and significant at $p < 0.001$ (Supplementary Table 1). When the threshold of immobility was set at 12%, the correlation values were between 0.58 and 0.78 in trials 1 and 2. Trial 2 measures of immobile duration were generally better correlated with hand coding than those of trial 1. Increasing the threshold to 20% included mobility that would be disregarded by manual scoring, which indeed was reflected in the lower correlations across the two trials with a range of

Fig. 1 Strain means across BXD lines and the progenitor strains for duration (top panel) and frequency (bottom panel) of immobility in trials 1 (dark bars) and 2 (light bars). Error bars represent the standard error of the mean

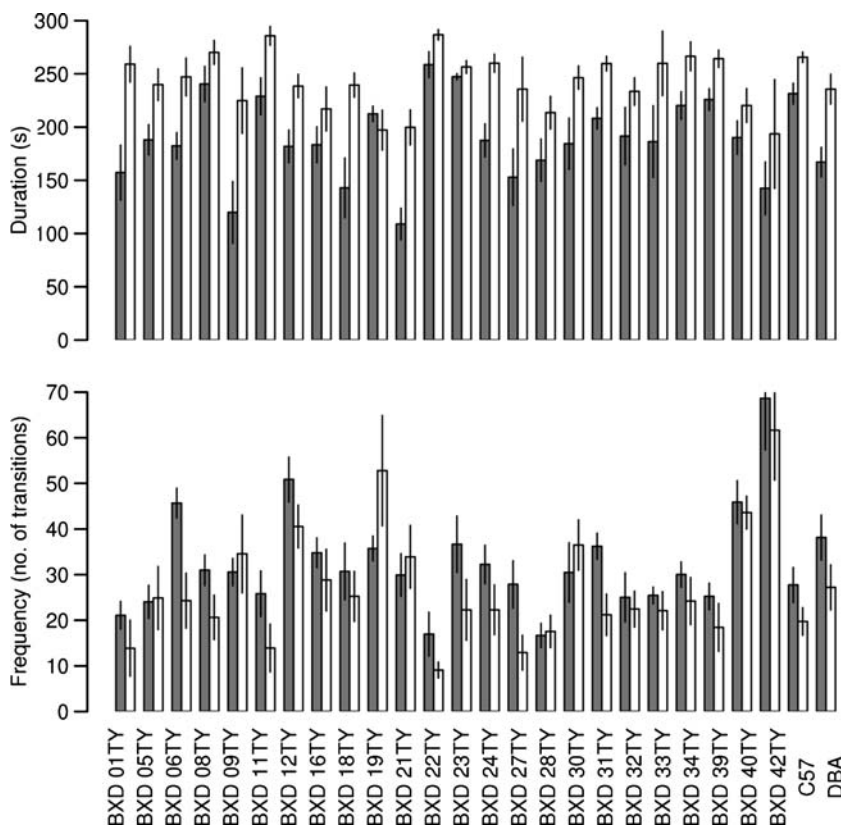
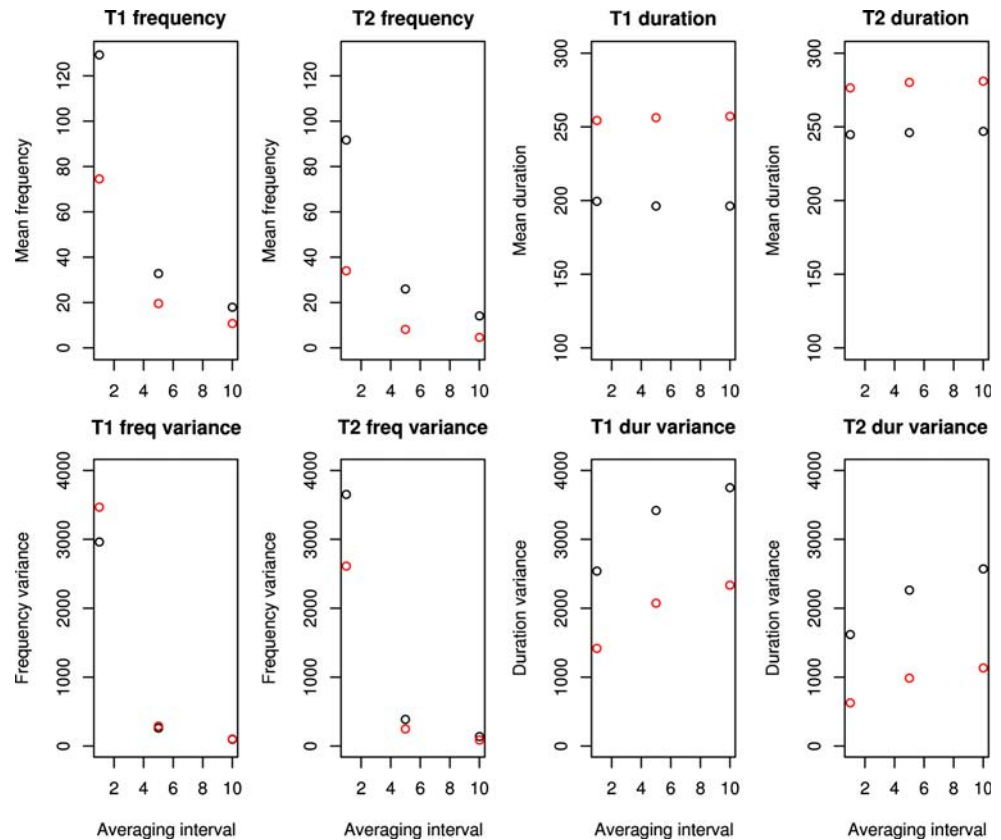


Fig. 2 Population mean and variance for frequency and duration of immobility in trial 1 (T1) and trial 2 (T2) of the tail suspension test. These are calculated using the complete set of 226 animals in the study, from 26 inbred strains (24 BXD RI strains, C57BL/6J, and DBA/2J). Mobility threshold: 12% black points, 20% red points



0.44–0.61. Frequency measures were poorly correlated to the manual scores, probably because it is difficult to consistently record frequency information by hand. The analysis here focuses on duration of immobility (the measure conventionally used), but there may be a completely different phenotype available from the frequency scores.

TST phenotypic response

In performing two trials, we questioned whether trials 1 and 2 elicited a different or a similar behavioral response in the TST. Figure 1 shows that there was a quantitative difference and dependent-samples *t*-test calculations confirmed that this was significant at the $p < 0.001$ level (Supplementary Table 2), with an increase in the mean time spent immobile during the second trial.

Genetic effect size

Figure 3 illustrates the fraction of variance attributable to genetic differences (strains) over the 12 combinations of threshold and averaging time. This demonstrates the same trend for both trials and in both frequency and duration, which is that the genetic effect size (eta squared) is highest at the lowest averaging interval and 12% threshold. Although the differences observed were modest, duration of immobility was more heritable in trial 1 than in trial 2. The majority

of all heritability estimates for trial 1 are over 0.3, whereas only one estimate in trial 2 met this criterion (Table 1).

Interval mapping analyses

We calculated whole-genome scans for all 12 combinations of scoring parameters. LOD scores are plotted side by side in Fig. 4 for each trial across these measures. The strongest QTL peaks for trial 1 immobility duration are seen on chromosome 4 (peak LOD score = 4.56, rs13477796) and 15 (peak LOD score = 3.41, rs13459176), and contrary to the expectation from heritability estimates, these are both best seen using the 20% mobility threshold. Linkage is not affected much by the averaging interval. For trial 2, the strongest duration QTLs are on chromosomes 11 (LOD = 3.59, rs13481087) and 18 (LOD = 2.19, gnf18.027.000), and there was a pattern of weaker signals that do not overlap with those for trial 1. These indicate that there are different QTL regions influencing behavioral measures across trait scores for the two trials. Frequency of mobility gives quite a different picture, with a distal chromosome 1 locus for trial 1 (LOD = 2.55, rs6202860), and a stronger signal for trial 2 on chromosome 18, corresponding to that for duration (Supplementary Fig. 1).

Comparable results were obtained with interval mapping using WebQTL. Whole-genome LRS score plots for trial 1,

Fig. 3 Genetic effect sizes. These are estimates of the fraction of variance attributable to strain, calculated using within- and between-strain sums of squared deviances for the set of 26 inbred strains (24 BXD RI strains, C57BL/6J, and DBA/2J). Mobility threshold: 12% black points, 20% red points

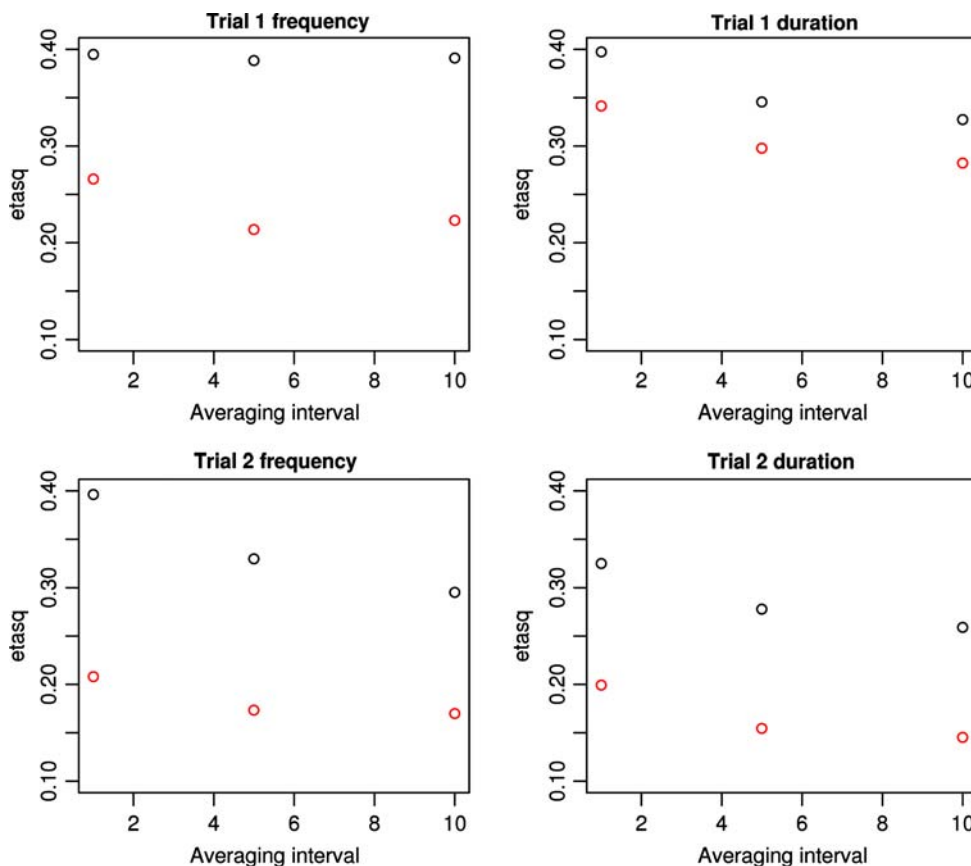


Table 1 Genetic effect sizes calculated for immobile duration using η^2

Threshold % Averaging factor	<12			<20		
	1	5	10	1	5	10
Trial 1	0.40	0.35	0.33	0.34	0.30	0.28
Trial 2	0.32	0.28	0.26	0.20	0.16	0.15

immobile duration, are shown in Fig. 5. Of particular interest were the linkage differences observed within trials between the two thresholds of immobility and across the averaging intervals. The effect of increasing the threshold altered the QTL peak profile. At a threshold of immobility of 12%, the significance of chromosome 15 was highest with a borderline suggestive significant peak on chromosome 4, while at a 20% threshold the chromosome 4 peak was significant with a reduced effect of chromosome 15. Averaging interval does not have as much of an effect as the threshold, but we do see the best LOD scores for both the chromosome 4 and 15 loci at an average interval of 5, contrary to what one would expect from heritabilities (Table 1). The QTL profile of trial 2 is completely different, but also shows a large effect of threshold. Supplementary Fig. 2 shows the same data in a different graphical form.

The chromosome 4 interval is approximately 20 Mb and contains a few genes that could be potentially interesting: *Ptprd*, *Tyrp1*, and *Mpdz*. Trait correlations using hippocampus consortium gene expression data (*M430v2 BXD Dec05* RMA) indicated that *Ptprd* is positively correlated with our TST scores for immobile duration ($0.76, p = 4.77e-06$). A much smaller QTL region (<13 Mb) on chromosome 15 included the candidate genes *Prkaa1*, *Gdnf*, *Slc1a3*, and *Pdzd2*. The most exciting gene whose expression data correlated with our trait scores in trial 2 was *Slc1a2* ($0.72, p = 3.09e-05$), which is a glial cell transporter that works together with *Slc1a3* to reduce cytotoxicity of glutamate (Lehre et al. 1995).

Discussion

The use of genetic reference populations such as recombinant inbred panels and their potential for accumulating data across time and between laboratories greatly facilitates understanding complex biological systems (Chesler et al. 2003). The availability of large-scale molecular phenotype data, chiefly from gene expression studies, is the source of much excitement and new advances. Their usefulness is ultimately determined by the link to phenotypes, and identifying informative measures in animal models to charac-

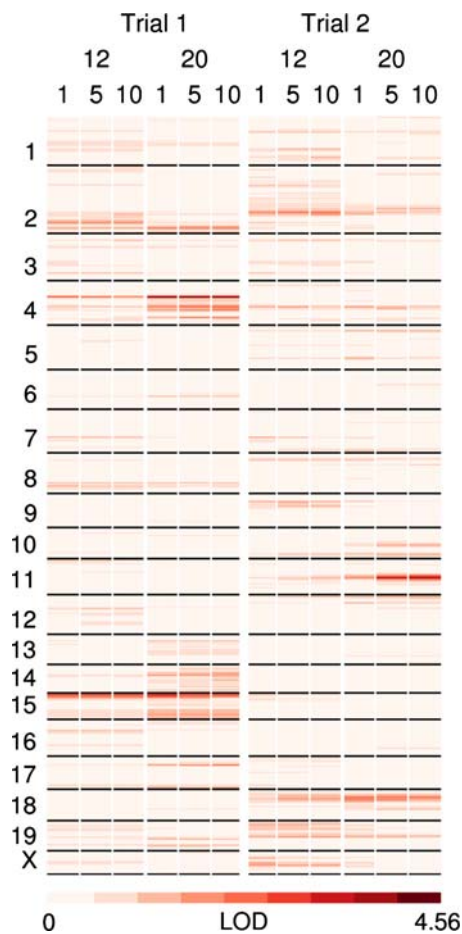


Fig. 4 Plot of LOD scores by position for 12 whole-genome scans of duration of immobility. Each vertical strip of the multimap corresponds to a genome scan that would conventionally be represented by a line plot. LOD scores are represented by color intensity. The genome scan was done with 801 markers with unique strain distribution patterns, and Y distance is proportional to number of markers, only an approximation of physical position. The centromere is at the top for each chromosome. The same data is presented as a perspective (“surface”) plot in Supplementary Fig. 2

terize them and translating these findings to other species, including human, remain as challenging as ever.

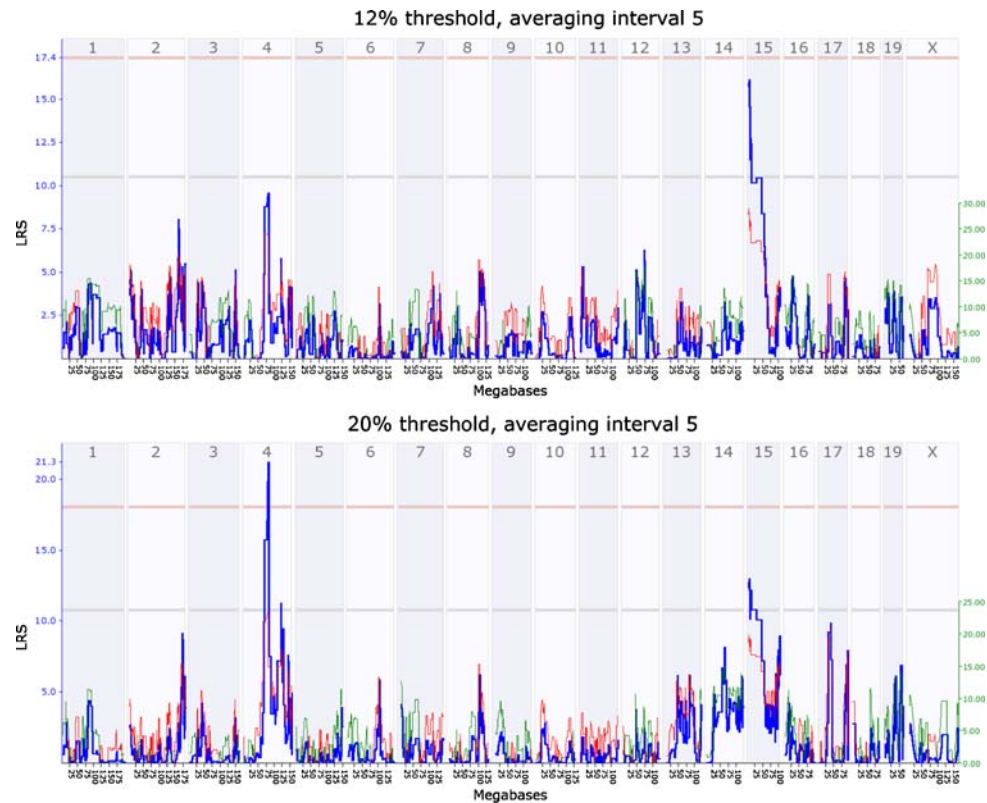
Recombinant inbred lines offer key advantages in understanding complex phenotypes such as behavioral measures. Measurement of the phenotype on multiple replicates of the same genotype is possible, allowing direct estimation of environmental variance and a well-characterized series of genotypes, giving access to genetic variance. The proportion of overall variance attributable to strain (broad sense heritability or genetic effect size) in general will be lower than those seen from human populations, such as from twin studies, because the degree of polymorphism present in recombinant inbred panels is limited. One obvious use of heritability estimates is for optimizing phenotypic measures to extract the maximum genetic information from, for example, a behavioral test.

This could be done with any panel of inbred strains, but a recombinant inbred or other genetically informative panel allows the analysis to go a step further and optimize the measure for detection of association with specific loci. This dissection in our study led to the slightly counterintuitive finding that optimizing on overall heritability would have reduced our ability to detect several loci of relatively large effect. We are not testing such a large number of methods that this effect can be attributed to a survey for chance associations (i.e., multiple testing). Rather, it is likely that other genetic effects are present, most likely due to numerous loci of small effect size, that we do not have the statistical power to detect, and these have a different profile. It is clear that the loci that we do detect respond differently to the analysis parameters that we have looked at here.

While it is not the standard procedure to perform two trials of the TST, we wanted to find out if a different response was elicited between trials. Dependent-samples t tests confirmed that there was a significant intertrial mean difference and that in the second trial the mean duration of immobility was greater. These are consistent with findings by Liu and Gershenfeld (2003) where similar differences in the duration of immobility were found between the first and second trial in the TST. It is possible that repeated exposure to the test removes novelty and decreases the anxiogenic response to the TST, altering the underlying phenotypic response. Initial exposure to the TST could have exacerbated the anxiety and struggling behaviors, while the second trial may be measuring a learned depressive-like behavioral response or behaviors similar to chronic mild stress. Liu and Gershenfeld (2003) argue on the basis of factor analysis that baseline strain differences in TST and imipramine response are distinct phenotypes. Although their finding was confounded by testing the imipramine response in trial 2 only, our results do support that trial 2 is at least genetically independent of trial 1 in the TST.

Interval mapping of the trait means and variances in WebQTL generated some linkage results that depended on the parameters defined. These were determined according to our experimental setup and video recording of the TST. The effect of significance seen for QTLs on chromosomes 4 and 15 in trial 1 was opposite at an altered threshold and across a range of averaging factors. Interesting QTL results found in trial 1 include an approximately 20-Mb region on chromosome 4 and a region centromeric on chromosome 15 (<13 Mb). A positive correlation was found with our trait values and hippocampal gene expression data for *Ptprd*, which is located within the QTL on chromosome 4. This is potentially interesting because *Ptprd* has been implicated in promoting cell growth and differentiation and neurite growth and is involved in cell signaling (Pulido

Fig. 5 Whole-genome scans of immobile duration measures in trial 1, comparing 12% and 20% mobility thresholds. These are similar to the second and fifth columns of Fig. 3 but prepared using WebQTL. These demonstrate the different QTL profiles for chromosomes 4 and 15, when the thresholds are set at 12% and 20% with an averaging factor of 5. Significant LRS thresholds are indicated by the upper pink line and suggestive by the gray line (LRS ~10)



et al. 1995). The peak LRS marker on chromosome 4, rs3708061, is located within an intron of *Mpdz*, a gene known to have multiple nonsynonymous polymorphisms between C57BL/6J and DBA/2J and has been identified as a quantitative trait gene for strain differences in sensitivity to seizures from withdrawal of alcohol or benzodiazepines and some chemical convulsants, most strongly for those that act with glutamergic signaling (Fehr et al. 2004; Shirley et al. 2004). This might be of particular interest in connection with a gene in the chromosome 15 region, *Slc1a3*, which belongs to solute carrier family 1, considered to be involved in high-affinity glial transport of glutamate. Few studies have documented the role of *Slc1a3* in neurobehavioral disorders; however, a recent report has reported its dysregulation in depressed individuals. Choudary et al. (2005) found that *SLC1A2* and *SLC1A3* were significantly downregulated in cortical areas of depressed individuals with concurrent upregulation of AMPA (α -amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid) kainate receptor genes. The resultant effect of these genes could potentially cause apoptosis, possibly contributing to hippocampal shrinkage seen in depression (Czeh and Lucassen 2007). Although the trait correlation with *Slc1a3* proved to be modestly significant, these expression correlation results are quite interesting and could be investigated further along with the finding that *Slc1a2* gene expression data correlated with our trait scores in trial 2. Furthermore, there could be a possible interacting effect between these

two loci that is not detectable from the current sample. The absence of the same QTL effects in trial 2 with very few peaks suggesting linkage could in part be supported by the QTL effects observed from the anxiety measures found by Turri et al. (2001). When they dissociated QTLs found for measures of anxiety, prior exposure to the test apparatus diminished the effect of chromosome 15 QTLs. Vöikar et al. (2004) also revealed reduced emotionality when repeatedly testing mice through a behavioral screen. The chromosome 15 QTL for measures of anxiety in Henderson et al. (2004) is close to but may not coincide with our chromosome 15 locus, which could be linked. Bolivar and Flaherty (2003) reported a QTL peak for intersession habituation on chromosome 15, which is much further distal.

From these results we conclude that for our experimental setup, exploring a range of average intervals at two different thresholds, the best genetic profile was given at a threshold of 20% and an averaging interval of 5. These results are contrary to the expectation that broad sense heritability estimates are an informative way of dissecting the phenotypic variance attributable to genes. Using EthoVision, Juszczak et al. (2006) investigated mobility in the TST at 2.5% and 3% thresholds but used a higher sample rate (12.5 video frames/second) and did not alter the averaging interval across scores. This highlights that the optimal analysis parameters depend on the details of video recording: resolution, lighting, color, and background.

In the present study TST was performed at the end of a battery of behavioral tests, which was designed to dissect and correlate phenotypes from a range of behaviors with overlapping pathways across a panel of BXD strains. It is possible that the genetic profile obtained in our data set would vary if experimentally naïve mice underwent the TST. However, considering the noninvasive nature of the preceding tests, with the most stressful test performed at the end of the battery, we suspect the differences would be marginal. McIlwain et al. (2001) demonstrated that battery tested and naïve mice displayed task-dependent differences but displayed similar levels of anxiety-related behaviors. Mice were housed in individual cages in our study; therefore, results may differ in comparison to group-housed mice results because individual housing effects in these strains have been studied and shown to have altered behavior in comparison to group housing (Võikar et al. 2005); however, TST was not included in their test battery. Paradoxically, handling in tests prior to the TST may have reduced habituation to the experimental environment (Võikar et al. 2004).

Our results show that scoring of behaviors in the TST using an automated system such as EthoVision and linking these to genetic analyses are complementary to extending further the genetic underpinnings of immobility.

Acknowledgments Funding for this work was provided by the Medical Research Council (G0000170). The authors thank S. Whatley for supplying the animals.

References

- Bolivar V, Flaherty L (2003) A region on chromosome 15 controls inter-session habituation in mice. *J Neurosci* 23:9435–9438
- Broman KW, Wu H, Sen S, Churchill GA (2003) R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19:889–890
- Bystrykh L, Weersing E, Dontje B, Sutton S, Pletcher MT, et al. (2005) Uncovering regulatory pathways that affect hematopoietic stem cell function using ‘genetical genomics.’ *Nat Genet* 37:225–232
- Chesler EJ, Wang J, Lu L, Qu Y, Manly KF, et al. (2003) Genetic correlates of gene expression in recombinant inbred strains: a relational model system to explore neurobehavioral phenotypes. *Neuroinformatics* 1:343–357
- Chesler EJ, Lu L, Shou S, Qu Y, Gu J, et al. (2005) Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. *Nat Genet* 37:233–242
- Choudary PV, Molnar M, Evans SJ, Tomita H, Li JZ, et al. (2005) Altered cortical glutamatergic and GABAergic signal transmission with glial involvement in depression. *Proc Natl Acad Sci U S A* 102(43):15653–15658
- Crowley JJ, Brodtkin ES, Blendy JA, Berrettini WH, Lucki I (2006) Pharmacogenomic evaluation of the antidepressant citalopram in the mouse tail suspension test. *Neuropsychopharmacology* 31(11):2433–2442
- Cryan JF, Mombereau C, Vassout A (2005) The tail suspension test as a model for assessing antidepressant activity: review of pharmacological and genetic studies in mice. *Neurosci Biobehav Rev* 29:571–625
- Czeh B, Lucassen PJ (2007) What causes the hippocampal volume decrease in depression? Are neurogenesis, glial changes and apoptosis implicated? *Eur Arch Psychiatry Clin Neurosci* doi: 10.1007/s00406-007-0728-0
- El Yacoubi M, Bouali S, Popa D, Naudon L, Leroux-Nicollet I, et al. (2003) Behavioral, neurochemical, and electrophysiological characterization of a genetic mouse model of depression. *Proc Natl Acad Sci U S A* 100:6227–6232
- Fehr C, Shirley RL, Metten P, Kosobud AE, Belknap JK, et al. (2004) Potential pleiotropic effects of Mpdz on vulnerability to seizures. *Genes Brain Behav* 3:8–19
- Flint J (2003) Animal models of anxiety In *Behavioural Genetics in the Post-Genomic Era* (Washington, DC: American Psychological Association)
- Galsworthy MJ, Paya-Cano JL, Monleon S, Plomin R (2002) Evidence for general cognitive ability (g) in heterogeneous stock mice and an analysis of potential confounds. *Genes Brain Behav* 1(2):88–95
- Galsworthy MJ, Paya-Cano JL, Liu L, Monleon S, Gregoryan G, et al. (2005). Assessing reliability, heritability and general cognitive ability in a battery of cognitive tasks for laboratory mice. *Behav Genet* 35(5):675–692
- Henderson ND, Turri MG, DeFries JC, Flint J (2004) QTL analysis of multiple behavioral measures of anxiety in mice. *Behav Genet* 34:267–293
- Juszczak GR, Sliwa AT, Wolak P, Tymosiak-Zielinska A, Lisowski P, et al. (2006) The usage of video analysis system for detection of immobility in the tail suspension test in mice. *Pharmacol Biochem Behav* 85:332–338
- Lehre KP, Levy LM, Ottersen OP, Storm-Mathisen J, Danbolt NC (1995) Differential expression of two glial glutamate transporters in the rat brain: quantitative and immunocytochemical observations. *J Neurosci* 15(3):1835–1853
- Liu X, Gershenfeld HK (2001) Genetic differences in the tail suspension test and its relationship to imipramine response among 11 inbred strains of mice. *Soc Biol Psychiatry* 49:575–58
- Liu X, Gershenfeld HK (2003) An exploratory factor analysis of the Tail Suspension Test in 12 inbred strains of mice and an F2 intercross. *Brain Res Bull* 60:223–231
- Liu X, Stancliffe D, Lee S, Mathur S, Gershenfeld HK (2006) Genetic dissection of the tail suspension test: a mouse model of stress vulnerability and antidepressant response. *Biol Psychiatry* doi: 10.1016/j.biopsych.2006.08.017
- Manly KF, Wang J, Williams RW (2005) Weighting by heritability for detection of quantitative trait loci with microarray estimates of gene expression. *Genome Biol* 6:R27
- McIlwain KL, Merriweather MY, Yuva-Paylor LA, Paylor R (2001) The use of behavioral test batteries: effects of training history. *Physiol Behav* 73:705–717
- Mill J, Galsworthy MJ, Paya-Cano JL, Sluyter F, Schalkwyk LC, et al. (2002). Home-cage activity in heterogeneous stock (HS) mice as a model of baseline activity. *Genes Brain Behav* 1(3):166–173
- Noldus Information Technology bv (2005) *EthoVision* 3.1. Wageningen, The Netherlands
- Plomin R, DeFries JD, McClearn GE, McGuffin P (2005) *Behavioural Genetics*, 4th ed. (San Francisco: Worth Publishers and W. H. Freeman)
- Pulido R, Krueger NX, Serra-Pages C, Saito H, Streuli M (1995) Molecular characterization of the human transmembrane pro-

- tein-tyrosine phosphatase delta. Evidence for tissue-specific expression of alternative human transmembrane protein-tyrosine phosphatase delta isoforms. *J Biol Chem* 270:6722–6728
- R Development Core Team (2006) *R: A language and environment for statistical computing* (Vienna, Austria: R Foundation for Statistical Computing), ISBN 3-900051-07-0, URL <http://www.R-project.org>
- Ripoll N, David DJ, Dailly E, Hascoet M, Bourin M (2003) Antidepressant-like effect in various mice strains in the tail suspension test. *Behav Brain Res* 143:193–200
- Shirley RL, Walter NA, Reilly MT, Fehr C, Buck KJ (2004) Mpdz is a quantitative trait gene for drug withdrawal seizures. *Nat Neurosci* 7:699–700
- Steru L, Chermat R, Thierry B, Mico JA, Lenegre A, et al. (1987) The automated Tail Suspension Test: a computerized device which differentiates psychotropic drugs. *Prog Neuropsychopharmacol Biol Psychiatry* 11:659–671
- Steru L, Chermat R, Thierry B, Simon P (1985) The tail suspension test: a new method for screening antidepressants in mice. *Psychopharmacology (Berl)* 85:367–370
- Trullas R, Barrington J, Skolnick P (1989) Genetic differences in a tail suspension test for evaluating antidepressant activity. *Psychopharmacology* 99:287–288
- Turri MG, Datta SR, DeFries J, Henderson ND, Flint J (2001) QTL analysis identifies multiple behavioral dimensions in ethological tests of anxiety in laboratory mice. *Curr Biol* 11:725–734
- Vaugeois JM, Odievre C, Loisel L, Costentin J (1996) A genetic mouse model of helplessness sensitive to imipramine. *Eur J Pharmacol* 316:R1–R2
- Võikar V, Vasar E, Rauvala H (2004) Behavioral alterations induced by repeated testing in C57BL/6J and 129S2/Sv mice: implications for phenotyping screens. *Genes Brain Behav* 3:27–38
- Võikar V, Polus A, Vasar E, Rauvala H (2005) Long-term individual housing in C57BL/6J and DBA/2 mice: assessment of behavioral consequences. *Genes Brain Behav* 4:240–252
- Wang J, Williams RW, Manly KF (2003) WebQTL: web-based complex trait analysis. *Neuroinformatics* 1:299–308
- Yoshikawa T, Watanabe A, Ishitsuka Y, Nakaya A, Nakatani N (2002) Identification of multiple genetic loci linked to the propensity for “behavioural despair” in mice. *Genome Res* 12(3):357–366