

Assessment of genetic susceptibility to ethanol intoxication in mice

Nathan R. Rustay*[†], Douglas Wahlsten[‡], and John C. Crabbe*

*Portland Alcohol Research Center, Department of Behavioral Neuroscience, Oregon Health & Science University, and Veterans Affairs Medical Center, 3710 Southwest U.S. Veterans Hospital Road (R&D 12), Portland, OR 97239; and [‡]Centre for Neuroscience and Department of Psychology, University of Alberta, Edmonton, AB, Canada T6G 2E9

Edited by Richard D. Palmiter, University of Washington School of Medicine, Seattle, WA, and approved January 2, 2003 (received for review November 29, 2002)

Increased use of gene manipulation in mice (e.g., targeted or random mutagenesis) has been accompanied by increased reliance on a very few rapid and simple behavioral assays, each of which aspires to model a human behavioral domain. Yet, each assay comprises multiple traits, influenced by multiple genetic factors. Motor incoordination (ataxia), a common characteristic of many neurological disorders, may reflect disordered balance, muscle strength, proprioception, and/or patterned gait. Impaired motor performance can confound interpretation of behavioral assays of learning and memory, exploration, motivation, and sensory competence. The rotarod is one of the most commonly used tests to measure coordination in mice. We show here that exactly how the rotarod test is performed can markedly alter the apparent patterns of genetic influence both in undrugged performance and sensitivity to ethanol intoxication. However, when tested with well chosen parameters, the accelerating rotarod test showed very high inter- and intralaboratory reliability. Depending on test conditions, ethanol can either disrupt or enhance performance in some strains. Genetic contribution to performance on the accelerating versus the fixed-speed rotarod assay can be completely dissociated under some test conditions, and multiple test parameters are needed to assess the range of genetic influence adequately.

The loss of motor coordination (ataxia) is a common characteristic of many neurological disorders and a frequent endpoint for studies of drug intoxication. Therefore, behavioral assays that model ataxia are of great importance to researchers who are interested in learning more about the mechanisms of drug action and disease. One of the most commonly used tests of motor incoordination is the rotarod (1–4), which has two variants: the accelerating rotarod (ARR) and the fixed-speed rotarod (FSRR). Studies of inbred strains (5–8), selected lines (9), and transgenic animals (10–12) have shown that rotarod performance is highly influenced by genetic background in mice. Genetically distinct mice often differ in their undrugged ability to perform, and differ in their sensitivities to ethanol and other drugs on the task.

Although the rotarod is widely used in biomedical research, there is little consensus on the ideal parameters and test schedules to produce optimal results. We have recently completed studies in genetically heterogeneous mice examining the influence of different rod diameters, rotation rate, and training regimens on rotarod performance, as well as their effects on sensitivity to ethanol intoxication. We obtained some expected results (e.g., training on the ARR improved performance, and performance was influenced by acceleration rate). More surprisingly, we found that higher acceleration rates suppressed sensitivity to ethanol intoxication. We also found that rod diameter did not markedly affect performance, provided that the diameter was large enough to prevent passive rotation on the rod (13).

Many researchers appear to design behavioral studies by simply adopting an apparatus and test strategy from the literature, without considering the potential effects of different apparatus and testing protocols. For genetic studies, this may be

a risky approach, because genotypes may perform well under some test conditions and poorly under others. Because researchers often do not know whether the specific test parameters they adopt are appropriate or ideal for their particular genotypes, it is difficult to interpret the results of a comparison between null mutant and wild-type mice that is restricted to a single test or condition. For example, mice with a null mutation for the serotonin 1B receptor subtype gene were less sensitive to ethanol than the 129 strain wild type when using two assays of intoxication (grid test and balance beam), but did not differ from wild type in sensitivity when using several other behavioral assays, including the ARR and FSRR (14). To assess the performance of genetically modified mice adequately, systematic data on a range of common background genotypes are needed, surveyed over different apparatus and test conditions. We studied inbred strains of mice on both the FSRR and the ARR and estimated their genetic codetermination. Because high intralaboratory reliability does not necessarily predict reliability across laboratories (15), we provide data on both intra- and interlaboratory reliability of ARR performance and report that apparent strain sensitivity to ethanol intoxication, and even the direction of the effects, depends markedly on how the tests are performed.

Materials and Methods

Rotarod. The AccuRotor Rota Rod (Accuscan Instruments, Columbus, OH) was used for both ARR and FSRR tests. The modified apparatus had a 63-cm fall height. Dowel surfaces were covered with 320 grit wet/dry sandpaper to provide a uniform surface and to reduce slipping. Starting at 0 rpm, the ARR was accelerated at a constant rate of 20–60 rpm/min (99.9 rpm maximum speed). The FSRR rotated at 3, 6.5, or 10 rpm.

ARR. Eight inbred strains (129S1/SvImJ, A/J, BALB/cByJ, BTBR T^{+tf/tf}, C3H/HeJ, C57BL/6J, DBA/2J, and FVB/NJ) were tested for the acquisition and maintenance of ARR performance and the intralaboratory reliability of genetic differences. For animal husbandry information, see *Supporting Text*, which is published as supporting information on the PNAS web site, www.pnas.org. The ARR accelerated at a rate of 20 rpm/min. On two successive days, mice were given 10 trials on the rod with a 30-sec intertrial interval (ITI). For ethanol and acceleration rate studies, male and female mice ($n = 6$ per dose per strain) were trained with 10 consecutive trials at 20 rpm/min. The next day, mice were then given one trial on the rotarod at 20 rpm/min to reacquaint the mice with the apparatus, followed by five baseline trials, one each at 20, 30, 40, 50, and 60 rpm/min consecutively. Mice were then injected with 0, 1.0, 1.25, 1.5, 1.75, or 2.0 g/kg ethanol and placed in individual holding cages. Thirty minutes later, mice were tested again at 20, 30, 40, 50, and 60 rpm/min consecutively. After 48 h, mice were retested in a

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: ARR, accelerating rotarod; FSRR, fixed-speed rotarod.

[†]To whom correspondence should be addressed. E-mail: rustayn@ohsu.edu.

Table 1. Peak performance of strains on both days of testing on the ARR

Strain	Asymptotic performance on ARR, sec	
	Day 1	Day 2
129S1/SvImJ	41.4 ± 2.9	51.0 ± 3.1
A/J	35.8 ± 2.2	42.4 ± 3.4
BALB/cByJ	45.7 ± 2.9	53.4 ± 2.3
BTBR T+ ^{tf/} tf	76.8 ± 8.6	103.1 ± 7.6
C3H/HeJ	35.9 ± 4.0	41.1 ± 3.6
C57BL/6J	53.6 ± 3.1	55.8 ± 4.2
DBA/2J	32.2 ± 2.4	40.8 ± 2.9
FVB/NJ	46.4 ± 3.1	55.3 ± 3.4

Values represent means ± SEM for the average latency to fall on the ninth and tenth trials for each day ($n = 12$ per strain). Bold text indicates values different from all other strains ($P < 0.01$).

similar manner, except that mice were administered 0, 2.0, 2.25, 2.5, 2.75, or 3.0 g/kg ethanol. Mice that received 0 or 2.0 g/kg on the first ethanol day were given the same dose on the second test day. All other mice were rerandomized to one of the four other dose groups.

For interlaboratory reliability studies, male and female mice from 21 inbred strains (129S1/SvImJ, A/J, AKR/J, BALB/cByJ, BTBR T+^{tf/}tf, C3H/HeJ, C57BL/6J, C57L/J, C58/J, CAST/Ei, DBA/2J, FVB/NJ, MOLF/Ei, NOD/LtJ, NZB/BINJ, PERA/Ei, PL/J, SJL/J, SM/J, SPRET/Ei, and SWR/J) were tested concurrently in Portland and Edmonton. Mice were trained with one day of 10 consecutive trials on the ARR at 20 rpm/min. The next day, mice were weighed and given three baseline trials, followed immediately by an injection of saline. After 30 min, mice were given three more tests and returned to the home cage. The next day, mice were treated identically, except that each mouse was injected with 2.0 g/kg (20% vol/vol) ethanol immediately after the three baseline trials, and tested 30 min later.

FSRR. Eight inbred strains were tested for performance on the FSRR by using parameters determined from studies conducted with genetically heterogeneous mice (13). All mice were tested at 3, 6.5, and 10 rpm consecutively, with 72 h between tests. For each speed, mice were given three practice trials (30 sec maximum, 30 sec ITI) before being given a 3-min criterion test. As soon as a mouse passed one criterion test (or was tested a maximum of 10 times), it was injected with 1.0, 2.0, or 3.0 g/kg ethanol and immediately placed back on the rod. If mice were able to stay on the rod for 3 min after ethanol, they were removed and returned to their home cages.

To compare with the ARR results another way, mice from the same eight inbred strains were tested on the FSRR 30 min after ethanol treatment. Mice were trained to perform on the rotarod at 3 and 6.5 rpm one day and 10 rpm the following day. For each speed, mice were given three practice trials followed immediately by a 3-min criterion test. As soon as a mouse passed the criterion test, it was returned to its home cage. Five days after training at 10 rpm, mice were tested again. On this day, mice were given three trials at 6.5 rpm, immediately injected with 1.25 or 1.75 g/kg ethanol, and tested 30 min later at 3, 6.5, and 10 rpm consecutively. Pilot studies showed that 2.0 g/kg was too high a dose for the majority of mice to be able to perform 30 min later. If mice were able to stay on the rod for 3 min after ethanol, they were given a score of “pass” and taken off the rod.

Results

Sexes did not differ significantly, so all reported analyses are collapsed across sex. Testing on the ARR at an acceleration rate

Table 2. Strain performance on the ARR

	Saline	1 g/kg	2 g/kg
Strain			
129S1/SvImJ	0.3 ± 6.4	11.2 ± 9.6	-22.8 ± 7.7
A/J	13.1 ± 3.8	22.4 ± 10.2	-12.5 ± 2.6
BALB/cByJ	2.1 ± 3.8	17.9 ± 8.3	11.0 ± 6.1
BTBR T+ ^{tf/} tf	-20.7 ± 7.8	-3.7 ± 9.5	-48.0 ± 18.4
C3H/HeJ	21.8 ± 10.2	14.4 ± 11.3	-9.6 ± 8.1
C57BL/6J	4.5 ± 7.5	14.7 ± 5.0	-14.3 ± 5.6
DBA/2J	15.0 ± 2.3	31.6 ± 2.5	1.7 ± 8.2
FVB/NJ	5.2 ± 3.9	13.2 ± 8.6	15.4 ± 5.1
Strain effect size, ω^2	0.434	0.201	0.502
Test-retest reliability (Pearson's r)	0.761	*	0.890

Values represent mean ± SEM for change from baseline latency to fall (sec) after saline or ethanol treatment ($n =$ five to six per strain per cell). Positive values represent increased performance, whereas negative values represent decreased performance compared to baseline. Bold text indicates values different from saline performance in each strain ($P < 0.05$).

*Only the saline and 2 g/kg ethanol groups received the same dose on successive tests (see text).

of 20 rpm/min revealed a significant learned component to the task (Table 1). Collectively, all strains improved over trials [$F(9,711) = 59.8, P < 0.001$] and days [$F(1,79) = 65.0, P < 0.001$]. Strains differed in their asymptotic performance on the ARR

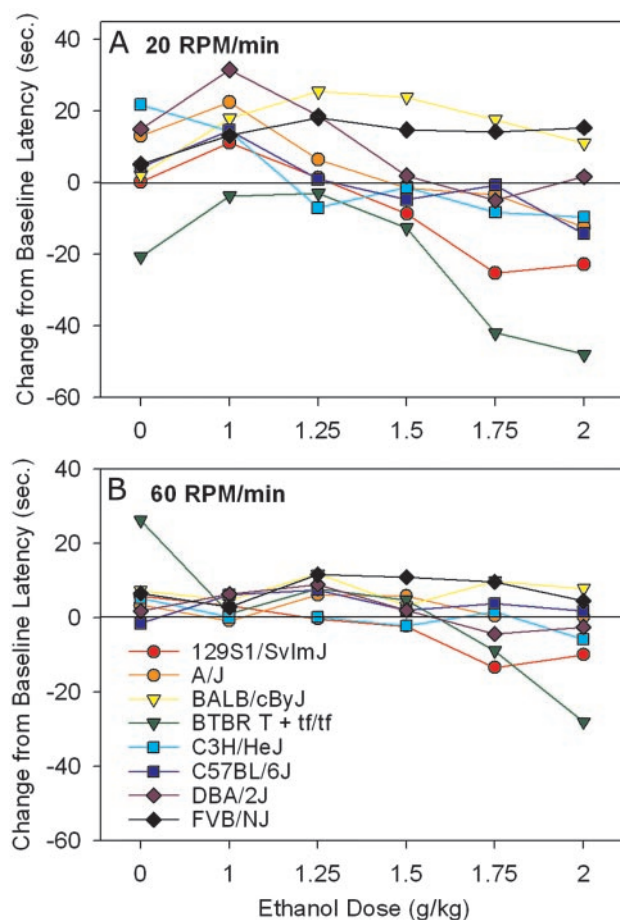


Fig. 1. Effects of ethanol dose and acceleration rate on ARR performance in inbred strains at 20 (A) or 60 (B) rpm/min. Values represent the mean change from baseline latency to fall (sec) for five to six mice per dose per strain.

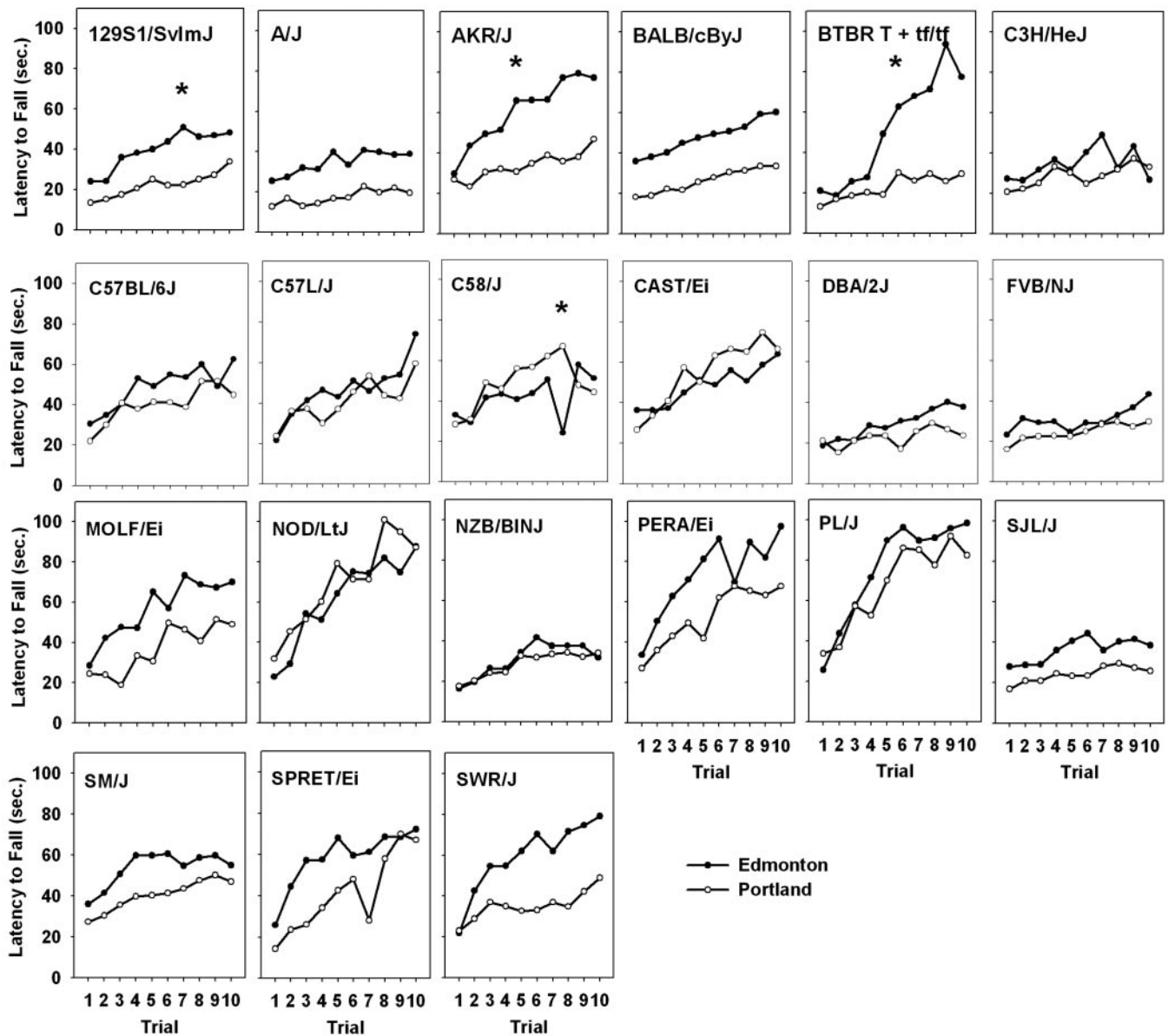


Fig. 2. Acquisition of ARR performance in Edmonton and Portland. Values represent mean latency to fall during the 10 acquisition trials at 20 rpm/min ($n = 6-12$ per strain per site). *, Strains that showed significantly different acquisition between sites ($P < 0.05$).

[$F(7,79) = 31.8, P < 0.001$], with BTBR T+^{tf/tf} performing markedly better than all of the rest. Strains also differed in their performance across trials [$F(63,711) = 4.6, P < 0.001$] and days [$F(7,79) = 13.5, P < 0.001$]. Genotypic test-retest reliability was determined by comparing strain means for the average of the last five trials on day 1 with the average of the last five trials on day 2. These strain means were correlated with a Pearson's $r = 0.97$, showing that day 1 performance was highly predictive of performance on day 2.

We next assessed the effect of ethanol dose and acceleration rate on ARR. Strains differed significantly in their ability to perform on the ARR at the end of training [$F(7,263) = 28.2, P < 0.001$]. Therefore, we sought an index of postethanol performance that would reflect both basal ability and sensitivity to ethanol. Both the change from baseline latency to fall after ethanol and the percent of baseline performance produced large and reliable genetic differences. Genotypic reliabilities were slightly higher when using the change from baseline, as were the

effect sizes for strain when each dose and acceleration rate was analyzed. We therefore report ARR data as the change from baseline latency to fall after treatment (Table 2).

It was clear from this experiment that the most robust strain and ethanol dose effects were seen at the slowest acceleration rate and that the fastest acceleration rate suppressed these differences (Fig. 1). There was a wide range of performance among strains when tested at 20 rpm/min; however, at 60 rpm/min, there was less variation among strains. Importantly, testing at the slower acceleration rate allowed detection of ethanol's dose-related effects, and low doses of ethanol improved performance over baseline in some strains. Higher doses of ethanol tended to impair performance across strains; however, there were a few strains (BALB/cByJ, FVB/NJ, and DBA/2J) that were not affected by the 2 g/kg dose. When tested at higher doses of ethanol (up to 3 g/kg) 48 h later, these more resistant strains also showed impairment (data not shown). Strain differences in sensitivity to ethanol-enhanced performance were all

Table 3. Strain performance by site on the ARR

Strain	Baseline		2.0 g/kg ethanol	
	Edmonton	Portland	Edmonton	Portland
129S1/SvImJ	44.9 ± 3.8	28.2 ± 2.9	-20.5 ± 2.0	-12.6 ± 2.7
A/J	43.1 ± 4.6	19.5 ± 1.4	-12.6 ± 5.2	-5.5 ± 3.2
AKR/J	58.7 ± 6.2	47.4 ± 4.3	5.2 ± 5.4	-1.9 ± 3.2
BALB/cByJ	60.4 ± 6.0	36.0 ± 4.9	6.7 ± 6.0	5.5 ± 3.9
BTBR T+ ^{tf/} tf	90.5 ± 10.9	39.4 ± 7.2	-58.1 ± 8.7	-28.0 ± 7.2
C3H/HeJ	43.8 ± 3.3	33.4 ± 4.2	-18.0 ± 5.8	-6.8 ± 6.5
C57BL/6J	51.7 ± 3.8	39.1 ± 4.4	-8.6 ± 5.0	0.0 ± 5.8
C57L/J	54.3 ± 6.9	38.8 ± 3.7	-15.6 ± 10.4	-13.5 ± 3.7
C58/J	58.1 ± 13.4	73.2 ± 9.9	-7.2 ± 13.4	-23.3 ± 13.5
CAST/Ei	99.6 ± 10.1	84.0 ± 21.3	-36.3 ± 17.7	-0.8 ± 9.3
DBA/2J	42.6 ± 4.4	23.5 ± 2.4	1.1 ± 5.8	4.0 ± 2.9
FVB/NJ	35.9 ± 4.5	28.0 ± 3.1	16.6 ± 5.9	6.0 ± 2.9
MOLF/Ei	56.5 ± 6.8	65.0 ± 8.8	26.9 ± 15.4	-27.8 ± 14.0
NOD/LtJ	99.3 ± 6.0	86.1 ± 7.2	-42.8 ± 12.2	-17.5 ± 7.9
NZB/BINJ	35.8 ± 2.6	29.5 ± 2.1	5.8 ± 5.5	0.2 ± 6.4
PERA/Ei	112.1 ± 18.5	73.4 ± 9.1	-44.1 ± 18.7	-33.3 ± 13.1
PL/J	82.5 ± 7.2	68.5 ± 8.3	-33.8 ± 10.8	-45.6 ± 11.1
SJL/J	43.8 ± 4.3	28.8 ± 2.1	-0.1 ± 2.5	-0.5 ± 4.1
SM/J	71.7 ± 6.3	59.0 ± 7.1	-13.8 ± 6.7	-29.5 ± 6.8
SPRET/Ei	72.4 ± 11.5	53.5	-24.1 ± 7.4	-9.0
SWR/J	113.9 ± 10.7	68.1 ± 7.1	-24.7 ± 9.4	-18.3 ± 5.7

Values represent means ± SEM for the baseline latency (sec) or change from baseline latency to fall (sec) after 2 g/kg ethanol. Only four SPRET/Ei mice in Edmonton and one in Portland completed testing. All other means represent 6–12 mice per strain per site. For change from baseline scores, positive values represent increased performance, whereas negative values represent decreased performance compared to baseline. Bold values represent significant differences in ethanol sensitivity between sites ($P < 0.05$).

but abolished at 60 rpm/min, as was the tendency for higher doses to impair performance strain-specifically. Blood ethanol concentrations (BEC) from the eight strains ruled out a pharmacokinetic explanation of strain sensitivity differences. Al-

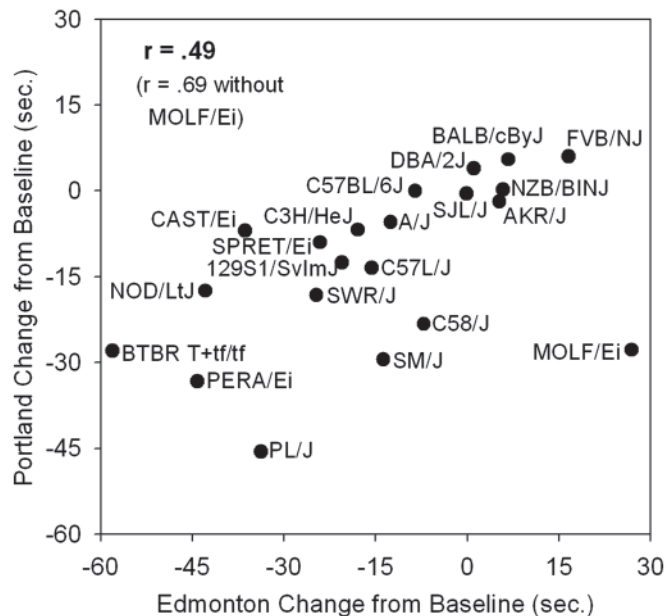


Fig. 3. Sensitivity to ethanol intoxication across sites. Symbols represent strain means for the change from baseline latency to fall (sec) 30 min after 2 g/kg ethanol. Positive values represent increased performance, whereas negative values represent decreased performance compared with baseline.

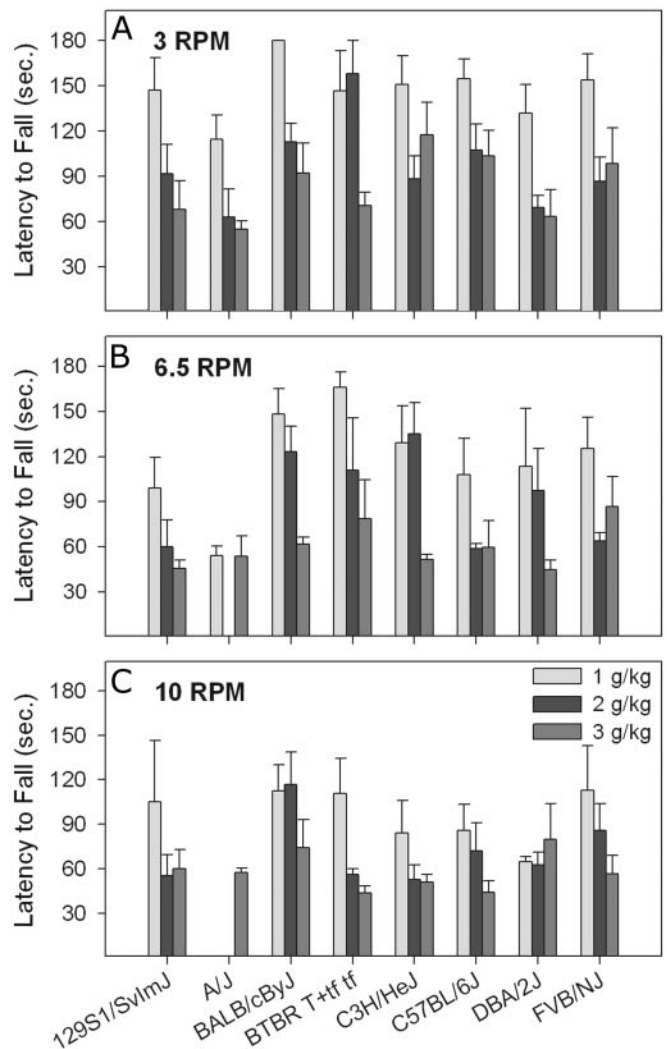


Fig. 4. Performance on the FSRR at 3 (A), 6.5 (B), or 10 (C) rpm immediately after ethanol. Values represent means ± SEM for two to eight mice per strain per dose. Bars are missing because only mice that were able to pass a 3-min criterion test at the given rpm before ethanol administration are included.

though strains differed in BEC [$F(7,86) = 7.4, P < 0.001$], strain means for BEC did not correlate with postethanol performance on the ARR ($r = 0.19$, not significant, data not shown).

Using the information gained from the ethanol dose and acceleration rate experiment, 21 strains of mice were tested in two separate laboratories to assess inter- and intralaboratory genetic reliability of the ARR data. Strain performances were highly consistent both between and within laboratories (Fig. 2, Table 3). However, there were a few strains (BTBR T+^{tf/}tf, AKR/J, and 129S1/SvImJ) that showed different acquisition of rotarod performance in Portland and Edmonton. These strains tended to acquire the task better in Edmonton. Even with these differences, interlaboratory correlations of strain means for peak acquisition performance after initial training (last two trials), presaline baseline, and preethanol baseline ranged from $r = 0.69$ – 0.80 ($P < 0.01$). Intralaboratory correlations were even higher, ranging from $r = 0.80$ – 0.95 , suggesting the test is highly reliable for detecting the effects of genotype on undrugged performance. Sensitivity to 2.0 g/kg ethanol was also significantly correlated between laboratories ($r = 0.49, P < 0.03$; Fig. 3). When the wild-derived MOLF/Ei strain was removed from this analysis, the correlation increased to $r = 0.69$ ($P < 0.01$).

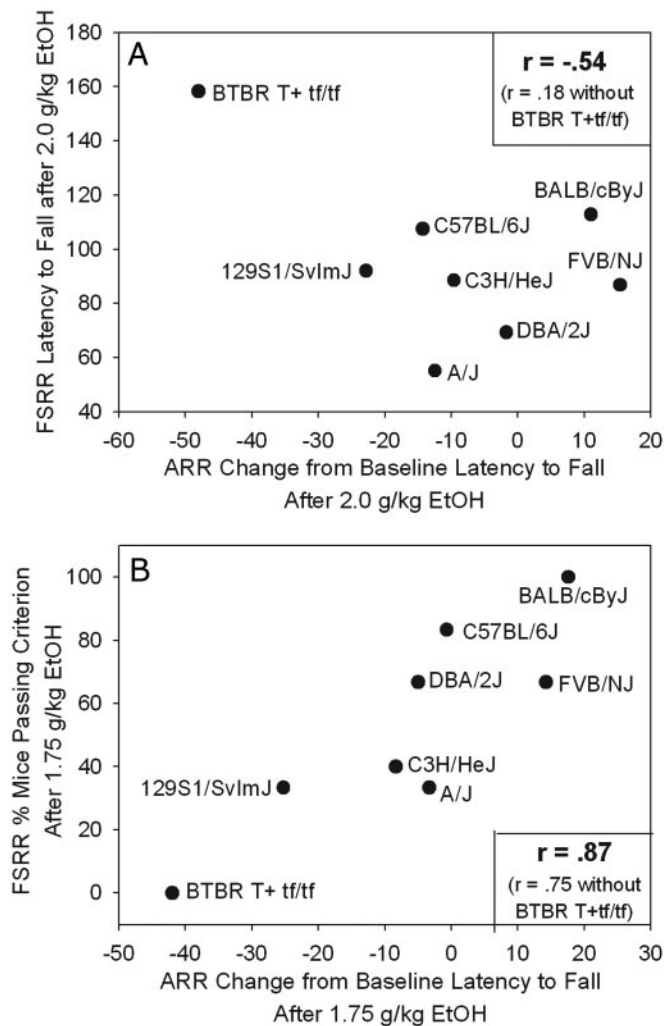


Fig. 5. Correlation of ARR and FSRR performance. Symbols represent strain means. Genetic codetermination was greater when tests were administered at the same postethanol time point ($n =$ six to eight per strain per test). (A) Strain mean latency to fall (sec) immediately after 2 g/kg ethanol on the FSRR plotted against mean change from baseline latency to fall (sec) on the ARR 30 min after 2 g/kg. (B) Percent of mice passing a 3-min criterion test 30 min after 1.75 g/kg ethanol on the FSRR plotted against the strain mean change from baseline latency to fall (sec) on the ARR 30 min after 1.75 g/kg ethanol.

The fixed-speed version of the rotarod task was tested in eight inbred strains. Results are shown in Fig. 4. At each rpm (3, 6.5, and 10), there was a significant effect of ethanol dose ($F > 5.9$, $P < 0.01$) with higher doses leading to shorter latencies to fall. At 3 and 6.5 rpm, there were also significant effects of strain ($F > 3.12$, $P < 0.01$), but at 10 rpm, the effect of strain was suppressed ($F = 1.51$, not significant). We analyzed data only from those mice that were able to pass a 3-min criterion test before drug administration each day. This resulted in elimination of the A/J strain from the 6.5- and 10-rpm analyses because so few mice were able to pass the preinjection criterion test (2 of 23 mice passed at 10 rpm).

We were interested in the genetic relationship between performance on the two versions of the rotarod task. To assess this, we correlated strain means for performance after 2 g/kg on the ARR (20 rpm/min) and the FSRR (3 rpm). Higher strain sensitivity to ethanol on the ARR was correlated with lower sensitivity on the FSRR ($r = -0.54$, not significant; Fig. 5A). Without the BTBR T+^{tf/tf} strain, the correlation of performance

became $r = 0.18$ (not significant), suggesting virtually no genetic relationship between these traits. Because these two assays were performed at different times after ethanol administration (FSRR immediately and ARR at 30 min), we tested another sample of the same eight strains on the FSRR 30 min after injection, when stable brain ethanol levels make the FSRR a pass-fail test. Mice were tested after 1.75 g/kg, because pilot testing showed that 2 g/kg was too high for most strains to be able to perform the task 30 min later. The percentage of mice per strain that passed a 3-min test after ethanol on the FSRR correlated significantly with ARR performance after 1.75 g/kg ($r = 0.87$, $P < 0.01$; Fig. 5B). Even when excluding the BTBR T+^{tf/tf} strain, the correlation was still very high ($r = 0.75$, $P = 0.05$). These analyses showed that there is a substantial shared genetic contribution in the two tasks, but only when they are used at the same postethanol time point.

Discussion

Our results illustrate the importance of using multiple task parameters when comparing different genotypes on the rotarod, as testing with a single set of parameters may lead to an inaccurate assessment of the true genotypic differences. One example of this is the effect of ethanol dose on ARR performance. Had we compared C57BL/6J with BALB/cByJ mice after a 1 g/kg dose of ethanol at 20 rpm/min, we would have concluded that they did not differ in ethanol sensitivity. However, after 2 g/kg, C57BL/6J mice were much more sensitive than BALB/cByJ mice. In addition, had we only tested strains at 2 g/kg ethanol, we would have missed the enhanced performance seen in many strains given low (1–1.25 g/kg) doses of ethanol (Fig. 1).

It was also clear that the rotarod may not be the ideal test of ataxia or intoxication for certain genotypes because of their propensity to jump from the rod instead of running on top. This was observed often in BTBR T+^{tf/tf} and the wild-derived SPRET/Ei mice. In addition, some strains may be more prone to holding onto the rod and passively rotating around instead of actively performing the task. Previous reports have noted this behavior (8, 16, 17), which we consider to be a different response that should be treated separately from balance and walking performance. We avoided this potential confound by using a rod diameter (6.3 cm) large enough to prevent the majority of mice from being able to hold on and rotate passively (13).

Motor coordination (18), as well as motor learning (19, 20), are thought to require functional integration of frontoparietal and motor cortex, cerebellum, and striatal circuitry. It is not surprising to see differences in acquisition, retention, and peak rotarod performance among a set of inbred strains (Table 1, Fig. 2). These differences are likely to reflect differences in structure and/or function of these essential brain regions, and different neuronal sensitivity to ethanol in these brain regions may also explain the strain differences in ethanol-induced ataxia.

We encourage the use of the rotarod for future studies of motor incoordination in new genotypes such as null-mutant mice, but recommend certain test parameters. We suggest using rotation rates no greater than 10 rpm for the FSRR and acceleration rates ≤ 30 rpm/min for the ARR. Intermediate acceleration rates produced graded results (data not shown), and the highest rates (≥ 40 rpm/min) reduced the sensitivity of the apparatus to detect strain and drug-dose effects. Future testing should use rod diameters of at least 6 cm to prevent mice from passively rotating on the rod instead of running on top. For intoxication studies, we recommend that mice be trained to a stable level of performance before drug administration. This insures that changes seen after drug administration are the result of drug action and not effects on learning on the apparatus. It also establishes that genetic differences in drug sensitivity are not confounded with genetic differences in undrugged perfor-

mance. In our hands, 10 massed trials was sufficient to produce peak performance across 21 strains, but some genotypes may differ. We also recommend testing multiple drug doses. Although drugs were known to improve rotarod performance (3, 4), we were surprised to learn that low doses of ethanol could improve performance. Testing multiple doses can allow the detection of both increased and decreased performance on the ARR.

Much has been made of the occasional inability of laboratories to reproduce findings of other laboratories (15). Perhaps one reason for this is that laboratories are often comparing behavior by using somewhat different apparatus and/or testing protocols. For example, the elevated plus maze, one of the more common tests of anxiety in rodents, differs frequently across laboratories in lighting conditions and in the wall height of the open, more anxiogenic portion of the maze, which affects alley selection of the subjects (21). Many different tests of learning and memory are used, often as if they were interchangeable, but genetic influence is likely to depend on specific parameters for these tests, too. Our studies show that procedural variables that may seem small to the experimenter can have pronounced effects on the pattern of genetic differences. The current set of experiments may represent a good strategy to improve interlaboratory reliability. By first identifying the most sensitive procedure for detecting genetic effects, we were able to detect highly reliable strain differences first within and then across our two laboratories (Fig. 3). By establishing the important procedural parameters of other behavioral tests of ataxia and tasks in other domains (e.g., anxiety and learning), we may be better positioned to design studies with maximal sensitivity to detect genetic differences. This should lead to more repeatable results both within and across laboratories. Even when such measures were taken, though, some laboratory-specific results were seen. The BTBR

T^{+/tf/tf}, AKR/J, and 129S1/SvImJ strains acquired the ARR test differently in Portland and Edmonton (Fig. 2). The poor performance of the BTBR strain in Portland was a surprise. In our first experiment in Portland (Table 1), this strain performed more like those tested in Edmonton. One explanation for these differences is that for these two experiments in Portland, there were different experimenters. Experimenter effects have been shown to affect behavioral results in other tests (22). Further, BTBR mice were not present in each shipment of mice for the second experiment, so the observed site differences may have originated from differential shipping effects. We believe that there will always be some genetic results that are laboratory specific. Systematic studies can reveal some of the bases for such effects (22).

Our data show the sensitivity of behavioral genetic results to specifics of the apparatus and protocol, and underscore the importance of adopting test parameters appropriate for a given experimental question. Further, they imply that more than a single test variant is required to characterize a complex behavioral domain, such as ataxia. Single variants of the rotarod task are often used to characterize targeted mutants, but such findings may not generalize widely. Systematic work with several genotypes and multiple tasks, each capturing different parts of the behavioral domain, will be necessary to understand the paths from behavioral assays through their component traits and biological substrates to specific genes.

We thank Chia-Hua Yu, Jason P. Schlumbohm, and Sean Cooper for their expert technical assistance. This work was funded by National Institutes of Health Grants AA12714, AA10760, DA37262, and AA13463; the N. L. Tartar Trust; a grant from the U.S. Department of Veterans Affairs; the Mouse Phenome Project; and Natural Sciences and Engineering Research Council (Canada) Grant 45825.

1. Bogo, V., Hill, T. A. & Young, R. W. (1981) *Neurotoxicology* **2**, 765–787.
2. Dunham, N. W. & Miya, T. S. (1957) *J. Am. Pharm. Assoc.* **46**, 208–209.
3. Jones, B. J. & Roberts, D. J. (1968) *J. Pharm. Pharmacol.* **20**, 302–304.
4. Watzman, N. & Barry, H., III (1968) *Psychopharmacologia* **12**, 414–423.
5. Crabbe, J. C., Gallaher, E. J., Cross, S. J. & Belknap, J. K. (1998) *Behav. Neurosci.* **112**, 668–677.
6. Gallaher, E. J., Jones, G. E., Belknap, J. K. & Crabbe, J. C. (1996) *J. Pharmacol. Exp. Ther.* **277**, 604–612.
7. Crabbe, J. C., Gallaher, E. J., Phillips, T. J. & Belknap, J. K. (1994) *Behav. Neurosci.* **108**, 186–195.
8. Tarantino, L. M., Gould, T. J., Druhan, J. P. & Bucan, M. (2000) *Mamm. Genome* **11**, 555–564.
9. Schafer, G. L. & Crabbe, J. C. (1996) *Alcohol Clin. Exp. Res.* **20**, 1604–1612.
10. Kóks, S., Volke, V., Veraksits, A., Runkorg, K., Sillat, T., Abramov, U., Bourin, M., Huotari, M., Mannisto, P. T., Matsui, T. & Vasar, E. (2001) *Psychopharmacology* **158**, 198–204.
11. Korpi, E. R., Koikkalainen, P., Vekovischeva, O. Y., Makela, R., Kleinz, R., Uusi-Oukari, M. & Wisden, W. (1999) *Eur. J. Neurosci.* **11**, 233–240.
12. Ogura, H., Matsumoto, M. & Mikoshiba, K. (2001) *Behav. Brain Res.* **122**, 215–219.
13. Rustay, M. R., Wahlsten, D. & Crabbe, J. C. (2003) *Behav. Brain Res.*, in press.
14. Boehm, S. L., II, Schafer, G. L., Phillips, T. J., Browman, K. E. & Crabbe, J. C. (2000) *Behav. Neurosci.* **114**, 401–409.
15. Crabbe, J. C., Wahlsten, D. & Dudek, B. C. (1999) *Science* **284**, 1670–1672.
16. Hilber, P. & Caston, J. (2001) *Neuroscience* **102**, 615–623.
17. Watzman, N., Barry, H., III, Kinnard, W. J., Jr., & Buckley, J. P. (1967) *Arch. Int. Pharmacodyn. Ther.* **169**, 362–374.
18. Massaquoi, S. G. & Hallett, M. (1998) in *Parkinson's Disease and Movement Disorders*, eds. Jankovic, J. & Tolosa, E. (Williams & Wilkins, Baltimore), pp. 623–686.
19. Hikosaka, O., Nakamura, K., Sakai, K. & Nakahara, H. (2002) *Curr. Opin. Neurobiol.* **12**, 217–222.
20. Jueptner, M., Frith, C. D., Brooks, D. J., Frackowiak, R. S. & Passingham, R. E. (1997) *J. Neurophysiol.* **77**, 1325–1337.
21. Wahlsten, D., Rustay, N. R., Metten, P. & Crabbe, J. C. (2003) *Trends Neurosci.*, in press.
22. Chesler, E. J., Wilson, S. G., Lariviere, W. R., Rodriguez-Zas, S. L. & Mogil, J. S. (2002) *Nat. Neurosci.* **5**, 1101–1102.