MMSTy study: Objectives were to (1) establish a measure of sleep propensity for a more comprehensive characterization of sleepiness in murine genetics and interventional studies and (2) to characterize sample sizes necessary for statistical differences in effect.

Design: Average multiple sleep latency values were compared in mice, varying strain, circadian time, and forced-wakefulness conditions.

Subjects: Adult male mice of inbred strains were studied.

Interventions: Mice were implanted with electroencephalographic and electromyographic recording electrodes. Twenty-four-hour periods of stable baseline sleep activity (> 600 minutes) were confirmed prior to baseline sleep-latency testing. Average sleep latencies were obtained across 10- and 20-minute nap opportunities within 4 consecutive 30-minute periods. Forced wakefulness protocols were performed prior to additional sleep-latency tests.

Measurements and Results: Sleep-latency testing with 20-minute nap opportunities every 30 minutes revealed a shorter sleep latency in the lights-on period (12.4 minutes ± 0.9 vs 16.5 ± 1, P < .001), a substantial reduction in sleep latencies in mice subjected to 6-hour forced wakefulness (eg, C57BL/6J baseline: 12.4 ± 0.9 minutes, and forced wakefulness, 8.5 ± 0.9 minutes, P < .01), and strain differences in latencies following short-term forced wakefulness (P < .01). Sample sizes for 85% power to detect a 25% reduction in the 20-minute daytime Murine Multiple Sleep Latency Test require fewer than 20 mice per group for commonly used transgenic background strains.

Conclusions: The Murine Multiple Sleep Latency Test is a robust measure of sleep propensity, and the latency varies with homeostatic and circadian influences. The test requires minimal added time to standard murine sleep recordings, yet yields important additive information.

Key Words: assay, sleep latency, sleepiness, mice, phenotype, wakefulness

Citation: Veasey SC; Yeou-Jey H; Thayer P; Fenik P. Murine multiple sleep latency test: phenotyping sleep propensity in mice. SLEEP 2004;27(3):388-93.

INTRODUCTION

MOUSE MODELS ARE INCREASINGLY USED TO HELP IDENTIFY RELATIVE AND SPECIFIC ROLES OF KNOWN NEUROTRANSMITTERS, NEUROMODULATORS AND NEUROPEPTIDES AND NOVEL GENES IN SLEEP-WAKE REGULATION.1-13 These murine models, however, are presently underutilized in the absence of a validated robust objective measure of sleepiness. Ideally, a sleep-propensity test in mice should parallel one of the established objective measures of sleepiness used in humans, for example, the Maintenance of Wakefulness Test14 or the Multiple Sleep Latency Test.15-17 The protocols for both of these assays account for ultradian- and circadian-rhythm variations in sleep latency by acquiring multiple measurements across time.14,15 Thus, a test characterizing overall sleep propensity in mice should involve the measurement of sleep latencies across multiple time points.

There are also several practical requirements for a murine phenotypic assay. For an assay to be used in studies of transgenic mice, it is imperative that the variance be low enough to allow adequate statistical power with small numbers of mice. The test must also be simple to perform and add little time and disruption to the baseline and recovery sleep recordings used for a full characterization of sleep in mice.

In light of the above, we developed the Murine Multiple Sleep Latency Test (MMSLT), paralleling the Multiple Sleep Latency Test (MSLT) for use in humans. In this paper, we compare variations of the protocols and provide sleep-latency data for 2 circadian time points and sleep-deprivation conditions. We show that the test is robust; and show dependence of the sleep-latency values on circadian and homeostatic influences. This test requires small sample sizes of animals for statistical power and adds minimal time to sleep recordings. This assay is suggested as an objective measure of sleep propensity or sleepiness in transgenic mouse models of sleep-wake abnormalities.

METHODS

Animals

Adult (16- to 26-week-old) male C57BL/6J (B6), A/J (AJ), DBA/1J (DBA), 129 SI/SvImJ (129S), F2 mice from parental strains: C57BL/6J and 129 SI/SvImJ, substrain stock number 101045, (B6129SF2) and C3H/HeJ (C3H) mice (Jackson Laboratory, Bar Harbor, ME) were studied. Methods and study protocols were approved in full by the Institutional Animal Care and Use Committee of the University of Pennsylvania and conform to the revised NIH Office of Laboratory Animal Welfare Policy. Food and water were provided ad libitum. Ambient lighting was constant with 12-hour lights-on duration beginning at 7 AM. Ambient temperature was 24°C ± 2°C.

Electrode Implantation and Recordings

Surgical implantation of electrodes and electrophysiologic recordings were performed using previously described anesthesia and electrode-implantation methods.18 Immediately following surgery, mice were returned to their home cages (in groups of 3-4 mice) for a 48- to 72-hour recovery period. Afterwards, mice were placed in individual recording cages; recording cables were connected to the mice under brief (30-second) 2% isofluorane anesthesia. Mice were housed inside a sound-attenuated, shielded, and well-ventilated designated sleep-recording room. Recordings were initiated 5 to 7 days after cable attachment. Electroencephalogram (EEG) signals were filtered at 0.3 and 35 Hz (1/2 max, 6 dB/octave), and electromyogram signals were filtered at 1 and 100 Hz and amplified (12A5, AC amps, Grass Telefactor, West Warwick, RI). Optimal combinations of the 2 frontal and 2 occipital EEG electrodes were acquired with an electrode selector board (12 PB_36 Electrode Selector, Grass) and sent to an A/D board (Converter 4801A, ADAC, Woburn, MA) in standard computers. The behavioral-state acquisition

Disclosure Statement

This work was supported by National Institutes of Health grant HL65225.

Submitted for publication September 2003
Accepted for publication November 2003

Address correspondence to: Sigrid Carlen Veasey, MD, 987 Maloney Bldg., Hospital of the University of Pennsylvania, 3600 Spruce St., Philadelphia, PA 19104; Fax: (215) 662 7749; E-mail: veasey@mail.med.upenn.edu

Measuring Sleep Propensity in Mice—Veasey et al
Establishing Stability of Sleep Patterns Prior to Sleep-Latency Testing

Confirmation of a stable baseline sleep pattern is essential to the interpretation of a baseline or forced-wakefulness MSLT. Thus, we used the sleep-analysis data as described above to confirm total sleep times of >600 minutes per day for each of the two 24-hour periods prior to the baseline sleep-latency testing. In addition, we visualized day-night graphs, developed by our group, to illustrate hourly percentiles throughout an entire recording; an example of the sleep analysis graphic is displayed in Figure 1, where hourly percentiles of NREM sleep (top panel) and REM sleep (lower panel) show sleep patterns across days prior to multiple sleep latency testing. An infected mouse may have sleep times >600 minutes per 24 hours, yet show an upward or downward trend in hourly percentiles (instability in sleep-wake patterns) over the 5- to 7-day recording prior to baseline latency testing.21,22 Thus, a second exclusion criterion was a visually detectable change in NREM sleep or REM sleep hourly percentile distributions relative to zeitgeber (lights on) time each day.

MMSLT Protocols

Protocol 1: Four 10-minute nap opportunities beginning at 7 pm

Paralleling the human MSLT protocol, the study began following the major rest period—at the onset of the lights-out period. On the day of testing, ambient darkness was maintained in the recording room for the mice throughout the normal dark cycle. Low-level red light was used to monitor the activity of mice. Prior to each nap opportunity, animals were maintained awake for 20 minutes by gently stroking the mice with a feather whenever they were still or EEG waveforms showed any slow-wave activity. The 10-minute nap opportunities began at 7 pm. Each day, ambient darkness was maintained for the duration of the nap opportunity, as is done in the human MSLT under similar circumstances.15 One-way ANOVA was used to compare groups (strain/forced wakefulness/time differences). For all paired samples (baseline vs sleep deprivation) or within a condition, the average sleep latency across the 4 nap opportunities. In naps during which no sleep occurred, the value used was the duration of the nap opportunity, as is done in the human MSLT under similar circumstances.15 One-way ANOVA was used to compare groups (strain/forced wakefulness/time differences). For all paired samples (baseline vs sleep deprivation), paired t testing was performed with Bonferroni corrections. Unless otherwise stated, values are reported as mean ± SEM. Statistical significance was determined using data analysis software (Graph Pad Prism, San Diego, CA) and defined when probabilities of the null hypothesis were <.05. Statistical power for future studies was performed

Mice that were eating, drinking, grooming, or building a nest were left undisturbed. Four nap opportunities were allowed over 4 consecutive 30-minute periods. For each nap trial, mice were allowed a 10-minute nap opportunity, followed by forced wakefulness, when needed, for the last 20 minutes. Electrographic recordings were observed throughout the study for sleep onsets, defined as 3 consecutive 10-second epochs scored as NREM or REM sleep.

Protocol 2: Four 10-minute Nap Opportunities Beginning at 2 pm (Baseline and 6-Hour Forced Wakefulness)

Protocol 2 differed from Protocol 1 only by time of day for the MMSLT and that the protocol was conducted under 2 conditions: baseline sleep and 6-hour total sleep deprivation (8 AM-2 PM). MMSLTs were performed from 2 PM to 4 PM following baseline and 6-hour forced wakefulness conditions. Thus, after sufficient baseline sleep was confirmed for each mouse, a baseline MMSLT was performed. The following day, mice were deprived of all sleep for the first 6 hours of the lights-on period using gentle handling, as above, with wakefulness confirmed electrographically. Immediately following sleep loss, a second MMSLT was performed beginning at 2 PM (forced wakefulness MMSLT).

Protocol 3: Four 20-minute Nap Opportunities Beginning At 7 pm

Protocol 3 differed from Protocol 1 by changing the duration of the nap opportunities to 20-minutes and changing the duration of wakefulness between nap opportunities to 10 minutes, so that the overall intervals and circadian time for testing were unchanged.

Protocol 4: Four 20-Minute Nap Opportunities Beginning at 2 pm (Baseline and 6 Hours Forced Wakefulness)

Protocol 4 differed from Protocol 2 by changing the duration of the nap opportunities to 20 minutes and changing the duration of wakefulness between nap opportunities to 10 minutes, so that the overall intervals and circadian time for testing were unchanged. After normal baseline sleep was confirmed for each mouse, a baseline MMSLT was performed. The following day, mice were kept awake with gentle handling as above from 8 AM to 2 PM, with wakefulness confirmed electrographically. Immediately following sleep loss, a second MMSLT was performed beginning at 2 PM (forced-wakefulness MMSLT).

Protocol 5: Four 20-Minute Nap Opportunities Beginning at 2 pm, Effects of 0, 2-Hour and 6-Hour Forced Wakefulness on Latency

This protocol was developed to assess dose-responsiveness for the MMSLT values and the duration of sleep loss. Three conditions of sleep-deprivation duration were tested, with separate groups of mice for each condition. The 3 conditions tested were sleep deprivation 0 hours, 2 hours, and 6 hours of enforced wakefulness, all completed at 2 PM for the onset of the MMSLT, using 20-minute nap opportunities.

Sleep and MMSLT Analysis

Behavioral state parameters were analyzed as previously reported, using factor analysis of variance (ANOVA) with Bonferroni corrections for the number of comparisons. The primary variable for analysis was the average sleep latency across the 4 nap opportunities. In naps during which no sleep occurred, the value used was the duration of the nap opportunity, as is done in the human MSLT under similar circumstances.15 One-way ANOVA was used to compare groups (strain/forced wakefulness/time differences). For all paired samples (baseline vs sleep deprivation), paired t testing was performed with Bonferroni corrections. Unless otherwise stated, values are reported as mean ± SEM. Statistical significance was determined using data analysis software (Graph Pad Prism, San Diego, CA) and defined when probabilities of the null hypothesis were <.05. Statistical power for future studies was performed.
RESULTS

Protocols

MMSLT Protocol 1 (10-minute MMSLT, 7 pm)

Average total sleep time per 24 hours for each of the 2 baseline 24-
hour periods prior to latency testing was 720 ± 73 minutes (n = 9). The average sleep latency for the four 10-minute nap opportunities (MMSLT-
10 minute) beginning at 7 PM in B6 mice was 10 ± 0 minutes (n = 9). That is, all mice remained awake for all 10-minute nap opportunities in the 7 PM to 9 PM study period.

MMSLT Protocol 2 (10-minute MMSLT, 2 pm, baseline-sleep loss)

Baseline MMSLT-10 minute was performed, beginning at 2 PM, in all AJ, C3H, B6, B6129SF2, and 129S mice with adequate EEG and electromyogram signals to visually score sleep onset. Total sleep times per 24 hours for each day prior to baseline testing was > 650 minutes, and sleep hourly patterns appeared stable for all mice included in analysis. Samples sizes and average sleep latencies for each strain are presented in Table 1. One-way ANOVA showed an overall strain difference in sleep latencies (P < .05). Multiple comparisons analysis revealed significantly lower baseline sleep latencies in AJ mice, compared to B6129SF2 mice (6.8 ± 0.7 minutes vs 9.6 ± 0.2 minutes, P = .05). Baseline sleep-latency values did not differ for B6, 129S, or B6129SF2 mice groups. Following sleep deprivation from 8 AM until 2 PM, all tested strains demonstrated significant reductions in average sleep-latency values, by paired t tests. Following enforced wakefulness, the latency for B6 mice was 3 minutes longer than for AJ (P < .01) and C3H (P < .01). DBA mice had the shortest sleep latency following enforced wakefulness (3.0 ± 0.4 minutes). Data for baseline and forced-wakefulness sleep latencies are provided in Table 1.

MMSLT Protocol 3 (20-minute MMSLT, 7 pm)

B6 mice were studied to assess the effect of extending the nap opportunity to 20 minutes at the onset of the lights-out period (MMSLT-20 minutes). Relative to the 7 PM onset 10-minute test (Protocol 1), fewer mice (2 of 8) demonstrated average sleep latencies greater than 20 minutes. The group average was 16.5 ± 1 minutes (n = 8). This latency is significantly greater than the average sleep latency obtained in Protocol 1 using a nap duration of 10 minutes (unpaired t = 6.1, P < .001).

MMSLT Protocol 4 (20-minute MMSLT, 2 pm, baseline-enforced wakefulness)

Three strains of mice were studied in this MMSLT–20-minute nap opportunity at 2 PM: B6, 129S, and B6129SF2. With the extended nap opportunity, none of the mice had an average sleep latency of 20 minutes. All strains demonstrated significant reductions in sleep latencies following short-term enforced wakefulness, as detailed in Table 2. Sleep latencies during the first nap opportunity were compared with latencies in the ensuing 3 naps and were found to show neither significant decreases (repeated-measures ANOVA) nor a trend toward change (nap order and sleep-latency value, as illustrated in Figure 2. Average baseline sleep latencies did not differ between the 3 strains. In contrast, following short-term enforced wakefulness, the sleep latency in B6129SF2 mice was significantly less than in B6 mice (4.5 ± 0.7 minutes versus 8.5 ± 0.9 minutes, P = .05). Sleep latencies were compared for 2 protocols (2 and 4). The 10-minute nap opportunity in Protocol 2 resulted in a significant underestimate of the sleep latency when compared to the latencies with the 20-minute nap (Protocol 4) (Protocol 2 sleep latency, 8.8 ± 0.5 minutes vs Protocol 4 sleep latency, 12.4 ± 0.9 minutes, P < .01).

MMSLT Protocol 5 (10-minute MMSLT, 2 pm, 0, 2, and 6 hours sleep loss)

A dose response in sleep latency to duration of enforced wakefulness was identified with nonparametric, in light of multiple values scored as > 10 minutes (Spearman r = .72; 95% CI, -0.87 to -0.43; P = .0001). The Runs test for linearity demonstrated a curvilinear dependence on forced-wakefulness duration for sleep latencies, with a majority of points on the best-fit straight line, as shown in Figure 3.

Sample Sizes for Adequate Statistical Power

The data collected in the above 20-minute protocols were used to determine sample sizes necessary for adequate statistical power.

---

Table 1—Ten-Minute Murine Multiple Sleep Latency Test, beginning at 2 PM

<table>
<thead>
<tr>
<th>Strain</th>
<th>Baseline Time, min</th>
<th>Baseline vs Forced Wakefulness Statistical Comparison</th>
<th>6-Hour Forced Wakefulness Time, min</th>
<th>6-Hour Forced Wakefulness vs Baseline Statistical Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>AJ</td>
<td>6.8 ± 0.7</td>
<td>vs B6129F2 n = 8</td>
<td>4.1 ± 0.5</td>
<td>3 min less than B6, q = 6.8 P = .05</td>
</tr>
<tr>
<td></td>
<td>n = 18</td>
<td></td>
<td>3 min less than 129S, q = 5.6</td>
<td>q = 4.8</td>
</tr>
<tr>
<td></td>
<td>vs A/J n = 9</td>
<td></td>
<td>7.0 ± 0.5</td>
<td>3 min more than A/J, q = 6.8 NS</td>
</tr>
<tr>
<td>B6</td>
<td>6.8 ± 0.5</td>
<td>none</td>
<td>4.0 ± 0.5</td>
<td>3 min more than 129S, q = 5.6 NS</td>
</tr>
<tr>
<td></td>
<td>n = 17</td>
<td></td>
<td>4 min more than C3H, q = 6.3</td>
<td>q = 4.1</td>
</tr>
<tr>
<td></td>
<td>vs C3H n = 9</td>
<td></td>
<td>7.0 ± 0.5</td>
<td>3 min more than C3H, q = 6.3 NS</td>
</tr>
<tr>
<td></td>
<td>vs DBA n = 7</td>
<td></td>
<td>7.0 ± 0.5</td>
<td>3 min more than DBA, q = 7.1 NS</td>
</tr>
<tr>
<td></td>
<td>vs 129S n = 7</td>
<td></td>
<td>7.0 ± 0.5</td>
<td>3 min more than 129S, q = 7.1</td>
</tr>
<tr>
<td>B6129SF2</td>
<td>9.6 ± 0.2</td>
<td>vs AJ n = 7</td>
<td>4.5 ± 0.7</td>
<td>NS</td>
</tr>
<tr>
<td></td>
<td>n = 8</td>
<td></td>
<td>4.5 ± 0.7</td>
<td></td>
</tr>
<tr>
<td>129</td>
<td>8.9 ± 0.9</td>
<td>none</td>
<td>4.0 ± 0.5</td>
<td>3 min more than A/J, q = 6.8 NS</td>
</tr>
<tr>
<td></td>
<td>n = 10</td>
<td></td>
<td>3 min more than 129S, q = 5.4</td>
<td>P = .001</td>
</tr>
<tr>
<td></td>
<td>vs C3H n = 8</td>
<td></td>
<td>4.0 ± 0.5</td>
<td>3 min more than C3H, q = 5.4 NS</td>
</tr>
<tr>
<td></td>
<td>vs DBA n = 8</td>
<td></td>
<td>4.0 ± 0.5</td>
<td>3 min more than DBA, q = 7.1 NS</td>
</tr>
<tr>
<td></td>
<td>vs 129S n = 8</td>
<td></td>
<td>4.0 ± 0.5</td>
<td>3 min more than 129S, q = 7.1</td>
</tr>
<tr>
<td>C3H</td>
<td>9.2 ± 0.3</td>
<td>none</td>
<td>4.0 ± 0.5</td>
<td>3 min more than A/J, q = 6.8 NS</td>
</tr>
<tr>
<td></td>
<td>n = 8</td>
<td></td>
<td>4.0 ± 0.5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>vs DBA n = 8</td>
<td></td>
<td>4.0 ± 0.5</td>
<td>3 min more than DBA, q = 7.1 NS</td>
</tr>
<tr>
<td></td>
<td>vs 129S n = 8</td>
<td></td>
<td>4.0 ± 0.5</td>
<td>3 min more than 129S, q = 7.1</td>
</tr>
<tr>
<td>DBA</td>
<td>N/A</td>
<td></td>
<td>N/A</td>
<td>P &lt; .001</td>
</tr>
<tr>
<td></td>
<td>n = 12</td>
<td></td>
<td>3 min less than B6, q = 8.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>vs B6 n = 8</td>
<td></td>
<td>3 min less than B6, q = 8.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>vs 129S n = 8</td>
<td></td>
<td>3 min less than 129S, q = 7.1</td>
<td></td>
</tr>
</tbody>
</table>

Data are strain across 4 nap opportunities beginning at 2 PM, occurring every 30 minutes, and are expressed as mean sleep latency in minutes ± SEM. Multiple comparisons across strains for baseline and following 6 hours of forced wakefulness were drawn with Tukey-Kramer, where q > 4.1 = P < .05. Paired t differences were used to determine effect of 6 hours of enforced wakefulness on sleep latency for each strain.

Table 2—Twenty-Minute Murine Multiple Sleep Latency Test, beginning at 2 PM

<table>
<thead>
<tr>
<th>Strain</th>
<th>Baseline Time, min</th>
<th>Baseline vs Forced Wakefulness Statistical Comparison</th>
<th>6-Hour Forced Wakefulness Time, min</th>
<th>6-Hour Forced Wakefulness vs Baseline Statistical Comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td>B6</td>
<td>12.4 ± 0.9</td>
<td>none</td>
<td>8.5 ± 0.9</td>
<td>4 min more than B6129F2, q = 4.6 P = .05</td>
</tr>
<tr>
<td></td>
<td>n = 13</td>
<td></td>
<td>8.5 ± 0.9</td>
<td>q = 4.6</td>
</tr>
<tr>
<td>B6129SF2</td>
<td>12.5 ± 1.0</td>
<td>none</td>
<td>4.5 ± 0.7</td>
<td>3 min more than B6, q = 4.6 P = .001</td>
</tr>
<tr>
<td></td>
<td>n = 8</td>
<td></td>
<td>4.5 ± 0.7</td>
<td>q = 4.7</td>
</tr>
<tr>
<td>129</td>
<td>12.7 ± 1.1</td>
<td>none</td>
<td>7.1 ± 0.7</td>
<td>none</td>
</tr>
<tr>
<td></td>
<td>n = 11</td>
<td></td>
<td>7.1 ± 0.7</td>
<td></td>
</tr>
</tbody>
</table>

*Data are expressed as mean ± SEM.
†Multiple comparisons across strains were drawn with Tukey-Kramer, where q > 4.1 = P < .05.
in future studies for a desired statistical power and effect size, and data are summarized in Table 3. Sample sizes necessary for 80% power to detect an effect size of 25% difference in sleep latency from the latencies observed in the present study were determined for B6, B6129SF2, and 129S strains using a 2-group t test with a .05 two-sided significance level. For future studies on mice with a B6 genetic background, the baseline MMSLT is expected to be approximately 12.4 minutes with a SD of 3 (as in our results above). Assuming a proportional group SD for a comparison condition or strain, the sample sizes necessary to detect a 25% effect size with 80% power would be 15 mice in the B6 control condition and 10 in the comparison group. If the variance for the comparison condition were higher, a larger number of mice would be needed to achieve statistical power. For this reason, the comparison-group expected sample sizes reported below are written as requiring at least that number in the group. Mice on a B6129SF2 background should have, for control groups, a baseline latency of approximately 12.5 minutes with a SD of 2.6 and, thus, require fewer mice for the same power and effect: 10 mice in the control B6129SF2 group and at least 9 mice in the group to be compared. For mice on a 129S background, the control group should have a MMSLT of 12.7 minutes with a SD of 3.2. For the above desired power, 13 mice would be needed in the control group and at least 10 in the comparison group. For similar statistical power but under forced-wakefulness conditions using mice on a B6 background, assuming a sleep latency following forced wakefulness of 8.5 minutes with a SD of 3.4, the sample size necessary for the control B6 group should be 17 and the comparison group, at least 11 mice. For B6129F2 background mice, the control group would have a sleep latency after forced wakefulness close to 4.5 minutes with a SD of 1.85; thus, to detect a 25% effect size with 80% power, would require sample sizes of 37 for the control group and at least 25 for the comparison group. Finally, for mice on a pure

### Table 3—Suggested Sample Sizes for Murine Multiple Sleep Latency Testing in Selected Strains of Mice to Achieve 80% Power to Detect a 25% Difference in Latency

<table>
<thead>
<tr>
<th>Background Strain</th>
<th>Sample size necessary to detect lower sleep latency under baseline conditions</th>
<th>Sample size necessary to detect lower sleep latency following short-term forced wakefulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>B6</td>
<td>≥ 15 in control group</td>
<td>≥ 17 in control group</td>
</tr>
<tr>
<td>B6129SF2</td>
<td>≥ 10 in control group</td>
<td>≥ 11 in mutant group</td>
</tr>
<tr>
<td>129</td>
<td>≥ 13 in control group</td>
<td>≥ 10 in mutant group</td>
</tr>
</tbody>
</table>

This calculation is based on the assumption that SD in the mutant group is similar to that in the control group for that background strain.

### Table 4—Suggested complete sleep recording with Murine Multiple Sleep Latency Test protocol

<table>
<thead>
<tr>
<th>Day after electrode implantation</th>
<th>Time of day</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>Connect recording cable.</td>
<td></td>
</tr>
<tr>
<td>6-9</td>
<td>Record baseline sleep.</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>1:40 PM (or 6 h, 40 min relative to lights on)</td>
<td>Baseline MMSLT: gentle handling for 20 min, followed by nap opportunity for 20 min every 30 min, with remainder of time as gentle handling for total of 4 naps.</td>
</tr>
<tr>
<td>11</td>
<td>Continue sleep recordings.</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>8 AM-2 PM (or 1 to 7 h relative to lights on)</td>
<td>Forced wakefulness, confirmed with electroencephalogram recordings monitored throughout forced wakefulness.</td>
</tr>
<tr>
<td>12</td>
<td>2 PM (or 7 h relative to lights on)</td>
<td>Begin first nap of MMSLT, and continue as for baseline MMSLT.</td>
</tr>
<tr>
<td>12</td>
<td>3:30 PM (or 8 h, 30 min relative to lights on)</td>
<td>Recovery-sleep recording.</td>
</tr>
<tr>
<td>14</td>
<td>Complete recording and download data for analysis.</td>
<td></td>
</tr>
</tbody>
</table>

MMSLT refers to Murine Multiple Sleep Latency Test.
129S background, the expected control-group forced-wakefulness sleep latency should be 7.1 minutes with a SD of 1.7. Thus, to detect a 25% effect with similar power and assuming a SD for the comparison group that is proportional to the SD of the control group, would require 12 control mice and at least 9 comparison mice.

DISCUSSION

We have designed a novel phenotypic assay for sleep propensity appropriate for implementation in most genetic or interventional behavioral-state studies in mice. The primary variable, average sleep latency, varies with circadian phase and with prior sleep amount in a dose-dependent manner. Thus, the assay is sensitive to circadian and homeostatic influences on sleep.20 The test requires minimal additional technical or recording time for standard mouse-sleep recordings; data from 8 or more mice may be obtained in 2 hours and are immediately ready for statistical analysis. Variance for the test is sufficiently low to require small numbers of mice for sufficient statistical power for effect sizes in sleep latency of 25%, for mice on either a C57BL/6J or 129SvImJ background.

Phenotyping sleep in mice is rapidly evolving. Numerous variables have been described, including sleep-state amount, sleep architecture, and spectral analysis of EEG signals for behavioral states.18,21-28 From the clinical perspective, however, important objective outcome measures are either measures of alertness, including the Maintenance of Wakefulness Test,14,29-31 measures of vigilance,32,33 and cognitive performance34-36 or measures of sleepiness, including the MSLT.15-17 These clinical measures of alertness and sleepiness are sensitive assays for characterizing the impact of inadequate sleep and intrinsic sleep disorders such as narcolepsy and obstructive sleep apnea on persons’ function. In this study, we have focused on developing a measure of sleepiness in mice. We believe, however, that tests of alertness and tests of sleepiness may provide complementary, rather than redundant, information. Tests of vigilance and higher cognitive throughput should also be incorporated into the phenotypic characterization of sleep-wake behaviors in mice.

The proposed MMSLT was designed to parallel the human MSLT, as a measure of propensity toward sleep,15 but as a test for mice, this assay has several important differences to acknowledge. In humans, the MSLT occurs with 4 or 5 nap opportunities distributed over an 8- to 10-hour time period and, thus, has the potential to measure sleep latencies at very different circadian time points and across several ultradian rhythms of sleepiness. Mice, in contrast to humans, do not have such prolonged periods of wakefulness, and an 8- to 10-hour protocol would result in significant forced wakefulness that, in turn, could impact upon later naps. For this reason, the total recording period has been shortened, and we have shown, as in Figure 2, that the latencies do not shorten across the four naps that would support a test-induced sleep-loss condition. This assay spans at least one ultradian cycle, yet characterizes sleepiness over a relatively short fragment of the circadian period. Another important difference factored into test design is the marked behavioral arousal over a relatively short fragment of the circadian period. This circadian arousal at lights off can obscure, or delay, the homeostatic pressure toward sleep.18 This heightened arousal during the test period is problematic for data analysis. Naps during which sleep does not occur are assigned a value of 20 minutes. This value, however, represents latencies > 20 minutes and is, by definition, a nonparametric measure. Further, by using a value of 20 minutes to describe what might be much higher, differences between groups or differences between forced-wakefulness conditions may be obscured in this “ceiling” value. Performing the test at 2 PM and extending the nap opportunity to 20 minutes markedly increases the number of sleep onsets observed. For these reasons, we have selected a 2-hour period with 4 nap opportunities and recommend completion of the test several hours before the active period begins.

This protocol was designed for implementation in transgenic studies, where background strains are predominantly either B6 or a 129 substrain. Murine transgenic studies, ideally, should have minimal variance between background strains for the phenotype being tested so that differences observed in the phenotype between a mutant mouse and the control mice may be attributed to the gene under study. The desire to have minimal strain differences also holds for mutagenesis studies in mice, where a second strain is necessary to identify the gene, and it is preferred to have minimal strain differences to better detect the mutant phenotypic difference. In contrast, where differences in inbred strains are desired, as in quantitative trait loci studies, a more appropriate time to perform the MMSLT may be in early dark period using longer sleep opportunities.2 The utility of the MMSLT for quantitative trait loci studies will require further research. In the interim, for transgenic and interventional studies in mice with B6 or 129 backgrounds, we recommend the following protocol for the MMSLT (as detailed in Table 4); this may be incorporated into a standard sleep study with a 6-hour forced-wakefulness protocol. The present study suggests that studies using mice bred to a pure background strain will require far fewer mice for the forced-wakefulness sleep-latency assay. When using control mice on mixed backgrounds, it may be necessary to perform a small pilot study of the control mice to determine SDs and actual sample sizes before ordering and implementing expensive transgenic mice.

By selecting a test time that minimizes variance, it is conceivable that variances between mutants and control background mice will also be reduced. However, using the 20-minute nap-opportunity protocol, all 3 strains tested (B6, 129S, and B6129SF2) showed significant reductions in sleep latency following short-term forced wakefulness, suggesting that increased sleepiness can be measured in all tested strains. Thus, the proposed protocol (Table 4) should be able to detect baseline sleepiness in mutant mice on any of these genetic backgrounds, provided adequate sample sizes are used, as outlined in this report. At the same time, the average baseline latencies were 12 minutes for the 3 strains, and this suggests that there is room to also detect reduced sleepiness in mutant mice on any of the 3 backgrounds, when the study is adequately powered. In contrast, there is a ceiling effect for the 10-minute nap protocol, and this does not leave room to detect mutants with less sleepiness. Thus, the 20-minute nap-opportunity protocol is recommended for most test conditions, as the initial test. In future studies, should mice of a particular background strain have increased variance or increased sleepiness, a second test at a different circadian time may be employed.

One intriguing finding in the sleep-latency data following forced wakefulness using the 20-minute nap-opportunity protocol was that sleep latency in B6129SF2 mice was less than the latency for each parent strain. Typically, behavioral tests in F2 mice groups show intermediate traits between parental strains, and increased variance,41 but this is not always true.42,43 In contrast, several tests, particularly learning paradigms, have shown F1 responses to differ from both parental strains.44-46 The F2 mice in our study were 8 weeks younger than either parent strain (16 weeks vs 22-26 weeks) and were studied separately. Although this is not a large age difference, it may contribute to a larger homeostatic response.47 The mice were extremely active in the first part of the forced-wakefulness protocol. It will be of interest to determine if the F1 and F2 sleep-latency responses to short-term sleep loss show intermediate traits or greater sleep propensity than parent strains in aged-matched groups for young and old groups.

The human MSLT was designed to detect both sleep propensity and REM-sleep-onset periods.15 However, we have designed the MMSLT to characterize sleepiness, rather than REM-sleep latencies. The test could be used to measure REM-sleep-onset periods but should be validated in the murine models of narcolepsy, and it is possible that nap durations and intervals might need modification (likely longer nap opportunities) to enhance sensitivity and specificity of the test.1,37

In summary, we have developed and validated an assay to characterize sleep propensity in mice. This test has been validated for circadian and homeostatic responses in sleep propensity. This assay is suggested for the characterization of sleepiness in transgenic mice undergoing standard sleep studies. Sensitivity of this assay for detection of sleepiness in models of intrinsic sleep disorders should substantiate the utility of the

SLEEP, Vol. 27, No. 3, 2004
assay. This test requires little added time and effort to a sleep-recording study in mice and should enable mice to sample for background strains of B6 or 129S to detect 25% differences in baseline and forced-wakefulness sleep latencies.

REFERENCES