Behavioral Neuroscience
2008, Vol. 122, No. 4, 777–787

0735-7044/08/$12.00 DOI: 10.1037/0735-7044.122.4.777

Data Mining in a Behavioral Test Detects Early Symptoms in a Model of Amyotrophic Lateral Sclerosis

Neri Kafkafi
Maryland Psychiatric Research Center

Daniel Yekutieli
Tel Aviv University

Paul Yarowsky
VA Baltimore Medical Center and University of Maryland School of Medicine

Gregory I. Elmer
Maryland Psychiatric Research Center

"What's wrong with my genetically engineered animal?" is a common yet often difficult to answer question in behavioral phenotyping. We present here a method termed Pattern Array for mining movement patterns and isolating those that best capture an effect of a genetic manipulation. We demonstrate the method by searching for early motor symptoms in the open-field behavior of SOD1 mutant rats, an animal model of amyotrophic lateral sclerosis. Pattern Array was able to identify a unique motor pattern that differentiated the SOD1 mutants from the wild-type controls 2 months before disease onset. This pattern included heavy braking while moving near the arena wall but turning away from it. SOD1 mutants performed this pattern significantly less than wild-type controls in 2 independent data sets. At such early age the SOD1 mutants could not be differentiated from the controls by standard behavioral measures or by subjective observation. The early discovered symptom may enable investigators to test therapies aimed for intervention rather than remediation. Our results demonstrate the feasibility and potential of detecting subtle behavioral effects using data mining strategies.

Keywords: locomotor behavior, rat, open field, SOD1, SEE

Supplemental materials: http://dx.doi.org/10.1037/0735-7044.122.4.777.supp

Exploratory data analysis and data mining are among the most useful paradigms in bioinformatics, yet they are rarely employed for the analysis of behavioral data in laboratory animals. Although many of the standard behavioral tests (especially spatial tests employing computerized tracking, see Crawley, 2000) automatically record large amounts of information-rich data, these data are rarely explored or mined, and are usually used merely for calculating a small set of hardwired cumulative measures, such as the total mileage traversed by the animal or the total time it spent in a certain area. This might limit the ability to detect subtle behavioral effects in knockouts and transgenics. Genetically engineered animal models typically require considerable effort and time to produce, which might prove frustrating if the standard tests then fail in discovering any behavioral effect (e.g., Grammer, Kuchay, Ghishti, & Baudry, 2005; Perez & Palmiter, 2005). In this study we demonstrate an algorithm termed Pattern Array, specifically designed for mining behavioral data using a large number of measures to isolate subtle yet consistent effects of genetic manipulation.

As a test case of the above problem we will discuss here the SOD1G93A (SOD1) rat model of amyotrophic lateral sclerosis (ALS). ALS is the most common form of motor neuron disease, a progressive and ultimately fatal degeneration of nerve cells (Rosen et al., 1993). About 10% of human ALS cases are inherited, of them 2% are caused by mutations in the superoxide dismutase 1 gene (SOD1, see Rosen et al., 1993). Transgenic rats and mice expressing any of several mutant human SOD1 alleles show many attributes of human ALS, including adult-onset muscle weakness as well as severe motor neuron loss (Bruijn et al., 1997; Bruijn, Miller, & Cleveland, 2004; Gurney et al., 1994) culminating in death by 5 months of age. These genetic models are widely used for developing and testing treatments (Howland et al., 2002; Rothstein et al., 2005).

In SOD1 rats the well-described adult onset of the disease typically occurs around postnatal day (PND) 110 with a standard deviation of less than 10 days (Matsumoto et al., 2006). Discovery of putative earlier motor symptoms that could be measured automatically and reliably in younger animals would enable investigators to develop and test treatments for delaying or even preventing the disease. Moreover, such symptoms may prove useful for contrasting symptomologies with nongenetic animal models of ALS (Shaw & Wilson, 2006). However, such

Neri Kafkafi and Gregory I. Elmer, Department of Psychiatry, Maryland Psychiatric Research Center, University of Maryland; Daniel Yekutieli, Department of Statistics and Operations Research, Tel Aviv University; Paul Yarowsky, Department of Veteran Affairs; VA Baltimore Medical Center and Departments of Pharmacology and Experimental Therapeutics, University of Maryland School of Medicine.

Neri Kafkafi and Gregory I. Elmer were given grants from NIH, CEBRA Grant DA022407; Paul Yarowsky was given the VA Merit Award, Department of Veteran Affairs, Baltimore Medical Center.

Correspondence concerning this article should be addressed to Neri Kafkafi, MD Psychiatric Research Center, University of Maryland, School of Medicine, Department of Psychiatry, Maple and Locust Streets, Room #A7, P.O. Box 21247, Baltimore, MD 21228. E-mail: nkafkafi@mprc.umaryland.edu

early symptoms have not been found. Matsumoto et al. recently characterized the behavior of SOD1 mutant rats using several tests, including righting reflex, inclined plane (for testing grip strength), home-cage and open-field activity, but did not detect reliable symptoms before PND 100. Moreover, as we also confirm in this study, Matsumoto et al. could not detect any abnormality in these animals before PND 90 even by subjective observations of their behavior.

This is an example of a common situation in the behavioral phenotyping of animal models: The most immediate hypotheses regarding a behavioral effect of the mutation were already exhausted, and more elaborate hypotheses will have to be tested one-by-one using dedicated (and likely costly and time-consuming) setups with an unknown chance of success. A simple open-field test of these animals, however, would yield a wealth of dynamical motor patterns that to date have mostly been ignored. In this study we describe a method designed to mine these data for reliable premorbid differences between the SOD1 and normal rats.

The data we use here are path coordinates from the SEE (Strategy for the Exploration of Exploration) open-field test (Drai & Golani, 2001), although in principle Pattern Array can be applied to any spatial test, or in fact any kind of behavioral test that records large amounts of data. SEE is a software-based strategy, embedded in the programming environment of Mathematica, for the visualization and analysis of free spatial behavior. It was recently shown to be useful for the behavioral phenotyping of mice (Benjamini, Drai, Elmer, Kafkafi, & Golani, 2001; Drai, Kafkafi, Benjamini, Elmer, & Golani, 2001; Horev, Benjamini, Sakov, & Golani, 2007; Kafkafi, Benjamini, Sakov, Elmer, & Golani, 2005; Kafkafi & Elmer, 2005; Kafkafi, Lipkind, et al., 2003; Kafkafi, Pagis, et al., 2003; Lipkind et al., 2004). These studies show that, in contrast with a common view of open-field behavior as an essentially stochastic phenomenon, it is structured and consists of intrinsic behavioral building blocks. The most basic of these building blocks are "progression segments," consisting of bouts of locomotor movement, and "lingering episodes" ("stops" in their generalized sense, consisting of both arrests and small "nonlocomotor" movements). SEE employs simple properties of these building blocks and their syntax as behavioral measures ("endpoints") for assessing open-field behavior.

In the above SEE studies, behavioral patterns and measures were defined using strategies of exploratory data analysis, employing several types of graphic visualization of the behavior in SEE (Drai & Golani, 2001) and/or watching the actual behavior. Once a behavioral pattern is algorithmically defined in SEE it can be tested over a large database of raw path data (Kafkafi, 2003) to assess its ability to discriminate reliably between different genotypes or treatments. The Pattern Array method develops this exploratory approach further into a data mining strategy, by defining a whole class of behavioral patterns as multiple simultaneous combinations of several ethologically relevant SEE endpoints, such as the distance from the arena wall, the speed and acceleration of movement, direction of movement, and turning. These combinations are then explicitly mined for those that maximize the difference between the experimental groups, in our case the difference between the SOD1 mutants and the wild-type controls.

Method

Animals and Testing

We obtained 12 SOD1 mutant (G93A) and 12 Sprague-Dawley wild-type control rats, both males at 5 weeks of age, from Taconic Labs, New York. They were housed 2 to 3 per cage with food and water ad libitum for 2 weeks on a standard dark–light cycle before the beginning of the experiment. The animals were tested at two ages: PND 50 to 55 and PND 75 to 80. Each of these tests included three 30 min open-field sessions—one session per day for 3 consecutive days—and a grip-strength test on the fourth day. All animals were weighed before each testing time point. Open-field tests were conducted using the standard SEE procedure (Drai & Golani, 2001, Kafkafi, et al., 2005). Briefly, the animal was allowed to freely explore a 2.50 m diameter circular arena while its location was tracked using Noldus EthoVision (Wageningen, The Netherlands) video tracking system at a rate of 30 Hz, and the {*Time, X, Y*} coordinates of the path (e.g., Figure 1) were exported to SEE. The grip force test was conducted separately for the fore and hind legs using a metal grid connected to an isometric force transducer (Columbus Instruments, Ohio) in a procedure similar to that described by Derave et al. (2003): The animal was lifted by its tail and made to hold the grid with its fore or hind limbs, and then pulled backward gently until it could no longer hold the grid. The maximal force in grams was recorded in six consecutive trials and the animal's final result was set to their median. The experimental protocols followed the "Principles of Laboratory Animal Care" (National Academy of Sciences, 1996). The animals used in this study were maintained in facilities fully accredited by the American Association for the Accreditation of Laboratory Animal Care (AAALAC).

SEE Behavioral Procedures

SEE (Drai & Golani, 2001) is a software-based strategy in the programming environment of Mathematica for the visualization and analysis of free spatial behavior. The EthoVision path coordinates were imported into SEE (Drai & Golani, 2001), and the SEE Path Smoother procedure (Hen, Sakov, Kafkafi, Golani, & Benjamini, 2004) was used to filter out tracking noise. Because the animals were not very active we pooled data from the three successive sessions for each animal in each age. It should be stressed that data was never pooled over different animals or over different ages. At a tracking rate of 30 Hz the data file of each animal at each age thus included 30 coordinates per second $\times$ 60 s $\times$ 30 min $\times$ 3 sessions = 162,000 data points. Using the standard SEE procedure the path was further divided into segments of progression and lingering (small local movements during stopping, see Drai, Benjamini, & Golani, 2000; Kafkafi, Mayo, Drai, Golani, & Elmer, 2001). The lingering component of behavior has too small spatial resolution to be reliably measured by current tracking technology, therefore our analysis concentrated on progression segments. Depending on the activity of the animal, the number of data points within progression segments usually consisted of 10,000 to 50,000 (i.e., about 5 to 25 cumulative minutes) per pooled three sessions. Mutants and controls did not consistently differ in their general activity (see Results section and Figure 3, second row, left).

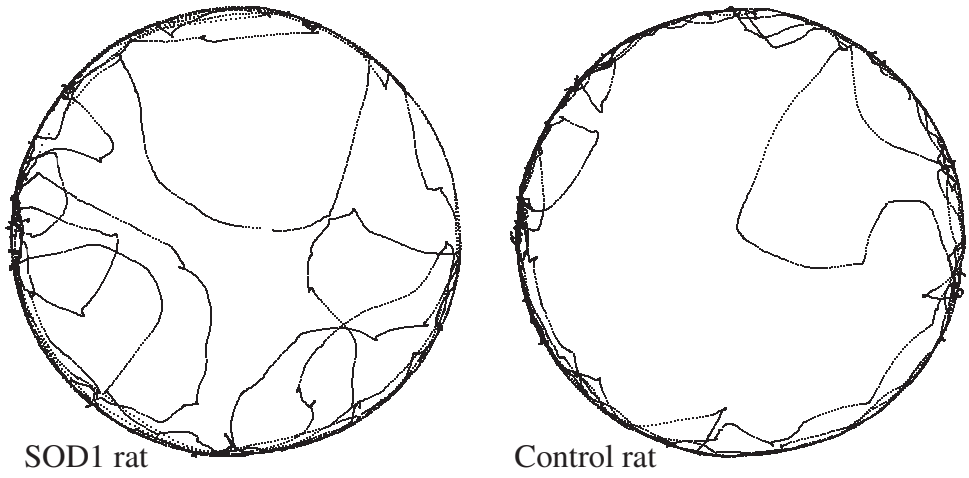

Figure 1. Example path plots. Representative path plots from one SOD1 rat (left) and one control rat (right) in the open-field arena. Only progression (movement) segments are shown. Each data point represents 1/30 seconds. The coordinates of these points are the input for the Pattern Array method.

The Pattern Array algorithm itself, described in the Algorithm section, was programmed in Mathematica using the SEE package (Drai & Golani, 2001) and the SEE Experiment Explorer package (Kafkafi, 2003).

Behavioral Attributes

The Pattern Array method was realized by transforming the path data into nine variables or "attributes", and partitioning the range in each attribute into several bins using cutoff values (see Table 1 and Algorithm section for full explanation). These attributes and cutoff values are based on previous studies of open-field behavior and large amount of preliminary data in both mice and rats. In all attributes the distribution of data was consistently unimodal with very similar modes between different animals. Following is a description of these nine attributes and the rationale for using them.

d: Momentary distance from arena wall. This is a well-established attribute of rodent open-field behavior (e.g., Ramos, Berton, Mormede, & Chaouloff, 1997) measuring thigmotaxis (wall-hugging). Both rats (Broadhurst, 1975) and mice (Leppanen, Ewalds-Kvist, & Selander, 2005) have been genetically selected for increased and decreased thigmotaxis. In mice, the distance from the wall was shown to be a factor in the intrinsic organization of the behavior (Lipkind et al., 2005), and the wall has a strong effect on the direction of movement even when the animal is at distance from it (Horev et al., 2007). Because the animals tend to stay much more near the wall we used four bins of increasing distance: 0 to 8 cm from the wall (approximately the range of maintaining physical contact with the wall), 8 to 20 cm from the wall (close proximity but not physical contact), 20 to 40 cm (slightly away from the wall), 40 to 125 cm (very far from the wall).

v: Momentary speed of movement. Speed was shown to be a key variable in the intrinsic categorization of behavior to progression and "lingering" in both mice (Drai et al., 2000) and rats (Kafkafi et al., 2001). We chose four bins of speed based on this experience: 0 to 20 cm/s (slow), 20 to 40 cm/s (medium), 40 to 60 cm/s (fast) and above 60 cm/s (very fast). The speed was computed

Table 1
Behavioral Attributes Used in Pattern Array

Symbol	Attribute definition	Units	Number of bins	Bin edges
d	Momentary distance from arena wall	cm	4	0, 8, 20, 40, 125
v	Momentary speed of movement	cm/s	4	0, 20, 40, 60, ∞
a	Momentary acceleration of movement	cm/s^2	5	−∞, −30, −5, 5, 30, ∞
j	Momentary jerk (change in acceleration) of movement	cm/s^3	5	−∞, −300, −50, 50, 300, ∞
h	Momentary movement direction relative to wall (heading)	degrees	5	−90, −30, −5, 5, 30, 90
c_4	Momentary path curvature in a 4 cm scale	degrees/cm	5	−∞, −10, −2, 2, 10, ∞
c_{16}	Momentary path curvature in a 16 cm scale	degrees/cm	5	−∞, −5, −1, 1, 5, ∞
t_s	Time from start of progression segment	s	3	0, 0.2, 1, ∞
t_e	Time to end of progression segment	s	3	0, 0.2, 1, ∞

Note. The path data are transformed into the nine attributes shown, and the range in each attribute is partitioned into several bins. Thus each of the path data points in Figure 1 is classified into cells or "patterns" as shown in Figure 2. See the Method section for more of the rationale and calculation of each attribute.

and noise filtered using the LOWESS algorithm as described in Hen et al. (2004) with a moving window width of 0.4 s.

a: Momentary acceleration of movement. Acceleration was shown to be a key variable discriminating the behavior of different genotypes of mice reliably and with high broad-sense heritability (Kafkafi, Pagis, et al., 2003). We chose five unequal bins of acceleration that produce approximately similar frequencies: less than −30 cm/s (strong deceleration, meaning heavy braking), −30 to −5 cm/s (mild deceleration), −5 to 5 cm/s (approximately uniform speed) 5 cm/s to 30 cm/s (mild acceleration), more than 30 cm/s (high acceleration). The acceleration was computed and noise filtered using the LOWESS algorithm as described in Hen et al. (2004) with a moving window width of 0.4 s.

j: Momentary jerk of movement. Jerk (the derivative of acceleration according to time, or the second derivative of the speed) was chosen because speed peaks were shown to be a meaningful component of rodent behavior (Drai et al., 2000; Kafkafi et al., 2001), and jerk is required to distinguish between speed peaks (near-zero acceleration and negative jerk) and local minima of speed (near-zero acceleration and positive jerk). We chose five unequal bins of jerk that produce approximately similar frequencies: less than −300 cm/s (very negative jerk, meaning a strong decrease in acceleration), −300 to −50 cm/s (mild decrease in acceleration), −50 to 50 cm/s (approximately uniform acceleration), 50 cm/s to 300 cm/s (mild increase in acceleration), more than 300 cm/s (strong increase in acceleration). The jerk was computed and noise filtered using the LOWESS algorithm as described in Hen et al. (2004) with a moving window width of 0.4 s.

h: Momentary movement direction ("heading") relative to wall. Horev at al., (2007) showed the effect of the wall on heading even from a distance. We computed the heading in degrees relative to the arena wall, with negative values representing movement toward the wall and positive values away from it. We chose five unequal bins of heading that produce approximately similar frequencies: −90° to −30° (moving toward the wall), −30° to −5° (moving slightly toward the wall), −5° to 5° (moving approximately parallel to the wall), −90° to −30° (moving slightly away from the wall), 30° to 90° (moving away from the wall). The direction was computed and noise filtered using the LOWESS algorithm as described in Hen et al. (2004) with a moving window width of 0.4 s.

c_4: Path curvature in a 4 cm scale. This attribute measures the momentary turning (change of direction) in a unit of path length. Kafkafi and Elmer (2005) showed that the curvature of the path has high heritability in the mouse and can be used to differentiate inbred strains with high replicability across laboratories. They also showed that the curvature measured in a 4 cm scale (smaller than the animal body) is not necessarily correlated with the curvature measured in a 16 cm scale (approximately body length in rats). For this reason we use the curvature in both scales as attributes in this study. The curvature in 64 cm scale, also used in the above study, was not used here because only a small portion of the segments were longer than 64 cm. Curvature was computed as detailed in Kafkafi and Elmer except for one difference: rather than using the sign to differentiate between left and right turning we used it here to differentiate between the direction toward the arena wall or away from it. As in *h*, negative curvature values indicate turning toward the wall and positive curvature values indicate turning in the direction away from the wall. We chose five unequal bins of curvature that produced approximately similar frequencies: less than −10 degree/cm (turning sharply toward the wall), −10 to −2 degree/cm (turning slightly toward the wall), −2 to 2 degree/cm (moving approximately straight ahead), 2 to 10 degree/cm (turning slightly away of the wall), more than 10 degree/cm (turning sharply away of the wall). As is Kafkafi and Elmer the curvature was computed relative to a distance rather than a time unit (because calculating it over very small distances is very sensitive to measurement error), meaning it represents different time windows depending on the speed, for example, in a typical speed of 16 cm/s using 4 cm scale implies a time window of 4/16 or 0.25 s.

c_{16}: Path curvature in a 16 cm scale. See the previous attribute for properties of path curvature and computing it in different distance scales. We chose five unequal bins of curvature that produced approximately similar frequencies: less than −5 degree/cm (turning sharply toward the wall), −5 to −1 degree/cm (turning slightly toward the wall), −1 to 1 degree/cm (moving approximately straight ahead), 1 to 5 degree/cm (turning slightly away from the wall), more than 5 degree/cm (turning sharply away from the wall).

t_s: Time from start of progression segment. Progression segments were shown to be a primary natural primitive of rodent spatial behavior (Drai et al., 2000; Kafkafi et al., 2001). By definition, a progression movement starts and ends with complete immobility. Our experience suggests that certain movement patterns may be affected if they take place immediately after the beginning of the segment, immediately before it ends, or anywhere in the middle. We therefore chose three unequal bins of time from the start of the segment: less than 0.2 s (6 data points in our 30 Hz measurement rate), 0.2 to 1.0 s and more than 1.0 s.

t_e: Time to end of progression segment. See previous attribute for the rationale. We chose three unequal intervals: less than 0.2 s, 0.2 to 1.0 s (30 data points) and more than 1.0 s.

Algorithm

Pattern Array is designed to test a very large number of possible movement patterns in parallel, and isolate only those patterns in which a significant difference between the experiment groups is detected (in our case a difference between the SOD1 mutants and the wild-type controls, see Figure 1). The main question is thus how to dissect behavior into many possible patterns in a meaningful way. We achieve this by transforming each coordinate of the path into a vector of several "attributes" (variables of movement) and dissecting the range of each attribute into several bins by introducing cutoff values. This strategy may be thought of as a generalization of many standard measures used in current behavioral tests, such as "center time". Center time is widely used in open-field tests to measure the animal's tendency to venture into the center of the arena, by defining a rather arbitrary cutoff value on the distance from the wall and counting the frequency of staying in a distance larger than this cutoff. Pattern Array generalizes this concept in three ways: by testing several different cutoff values in the "attribute" of distance from the wall, by testing several cutoff values in each of several additional attributes, and finally by testing intersections of several attributes. Only those cutoff values that manage to show an effect are kept.

The chosen attributes are variables that were shown in previous studies to be relevant to open-field behavior, such as momentary

distance from the wall, momentary speed and momentary change of direction. Using the above cutoff values the range of each attribute is divided into several bins, thus dividing the attribute space into many "cells" (e.g., Figure 2), each cell corresponding to a specific combination of attribute values, or a motor pattern (e.g., Figure 4 showing a particular event belonging to the cell highlighted in Figure 2). For each of these patterns the difference between the experiment and control group is then tested in the relative frequency of performing it.

As with data mining strategies in many other fields, the quandary with such an approach is the multiplicity problem. That is, when simultaneously screening a large number of possible movement patterns we need to prohibit the occurrence of false positives and provide valid statistical inference for the selected patterns (Benjamini & Yekutieli, 2005). We achieve this by using the most conservative multiple-comparisons criterion—the Bonferroni criterion: using a corrected significance level of α/n, where n is the number of potential movement patterns, thus ensuring that the probability of discovering even a single false movement pattern is less than α. Using the more adaptable false discovery rate (FDR, see Benjamini et al., 2001) is also an option in Pattern Array, but for the purpose of the present study it would not represent any advantage, because it functions almost identically to the Bonferroni in cases where very few zero hypotheses are rejected (i.e., very few discoveries). To further ensure the validity of the inference for the selected movement patterns we divide the animals into independent "training set" and "test set." The training set is used for isolating the best patterns as described above, while the statistical inference for the selected movement patterns only is based on the independently distributed test set. In the following we detail the algorithm step-by-step.

Input. The inputs for the algorithm are the (*T, X, Y*) coordinates of the animal's path in the arena belonging only to progression segments (see Method section and Figure 1, more details in Drai et al., 2000; Kafkafi et al., 2001) measured at a rate of 30 Hz. Progression segments are typically 6 to 300 data points in length (i.e., 0.2 to 10 s in duration) and a session typically includes several hundred of them to a total of 10,000 to 50,000 data points per animal.

Step 1. Each data point is quantified using m "attributes" of movement. In this study we used $m = 9$ attributes, defined in Table 1, that were found in previous studies to be relevant to open-field behavior. The distance from the wall d was shown to measure heritable thigmotactic behavior (e.g., Broadhurst, 1975; Leppanen et al., 2005; Lipkind et al., 2005; Ramos et al., 1997). The momentary speed v was shown to be a key variable in the intrinsic categorization of behavior into progression and "lingering" in both mice (Drai et al., 2000) and rats (Kafkafi et al., 2001). The acceleration a was shown to have high heritability and reliably in mouse inbred strains (Kafkafi, Pagis, et al., 2003). The jerk j (the derivative of acceleration according to time, or the second derivative of the speed) was chosen because speed peaks were shown to be a meaningful component of rodent behavior (Drai et al., 2000; Kafkafi et al., 2001) and the jerk is required to distinguish between speed peaks (near-zero acceleration and negative jerk) and local minima of speed (near-zero acceleration and positive jerk). The momentary heading h (direction of movement relative to the arena wall) was proposed by Horev at al. (2007) as an important aspect of open-field behavior in the mouse. The path curvature in a scale of 4 cm and 16 cm (c_4 and c_{16}, respectively) were shown to discriminate several mouse inbred strains with high heritability and reliability (Kafkafi & Elmer, 2005). Finally, attributes t_s and t_e quantify the temporal location of the data point within progression segments (Drai et al., 2000; Kafkafi et al., 2001), thus making it possible to mine patterns that always take place in, for example, the beginning or end of progression segments. Further details regarding the computation and rationale of the attributes can be found in the Method section.

The identity of attributes and their number m can be adapted depending on the objective of the study. For example, in an ongoing study using Pattern Array for classification of drugs into psychopharmacological classes in mice (Kafkafi, Yekutieli, &

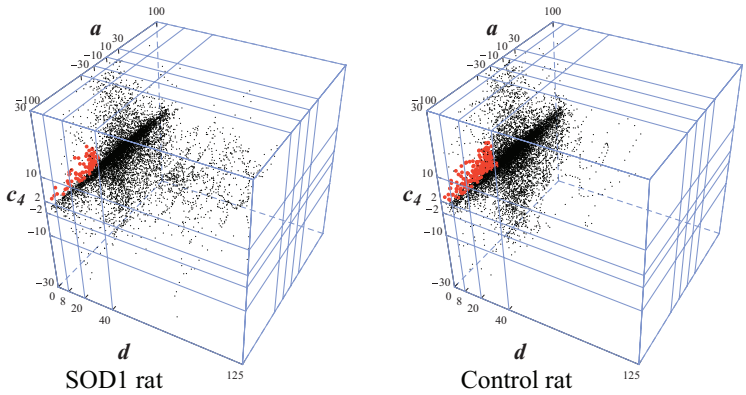

Figure 2. Three-dimensional attribute spaces. Attribute spaces for one SOD1 rat (left) and one control rat (right), corresponding to the same sessions shown in Figure 1. Each point in the attribute space corresponds to a point in the path plot. The three attributes chosen here are the distance from the arena wall (d), the acceleration (a) and the curvature of the path (c_4). Grid lines show the division of the attribute space into "cells." Points falling into one of the cells are highlighted (orange). Dividing their number by the total number of points gives the relative frequency of performing this pattern. The highlighted cell here is $P\{1,^{*},1,^{*},^{*},4,^{*},^{*},^{*}\}$, which is the pattern found to best differentiate the SOD1 mutants from the controls.

Elmer, in press) the same attributes were used with the addition of a tenth attribute: the time from injection. This attribute was added because in drug-induced behavior the development in time during the session is expected to have an important effect. In the present study we do not except a motor deficiency to change much during the session, or even during consecutive sessions within the same age, and therefore it is preferable to discard this attribute and decrease m. Further considerations for the choice of m are detailed in the Discussion section.

The path of the animal in the arena (see Figure 1) can thus be thought of as a set of trajectories (progression segments), each typically including several tens of data points, each data point consisting of a nine-dimensional vector of the form (d, v, a, j, h, c_4, c_{16}, t_s, t_e) in the attribute space. However, in practice using all nine attributes at once is excessive (see Step 3) so in this study we consider only subspaces of up to four attribute subsets. Figure 2 illustrates data in one three-attributes subspace.

Step 2. The range of each attribute is partitioned into several disjointed intervals (bins), thus dividing each attribute subspace into many "cells" (grid lines in Figure 2). Table 1 shows the number of bins for each attribute and the cutoff values chosen for the bin edges. Note that the bins were not necessarily chosen to be of equal length, but rather to contain data points in approximately equal frequencies (see below). For example, rats and mice typically move near the wall much more frequently than in the center of the arena, therefore the distance from the wall d was divided using cutoffs of 0, 8, 20, 40, and 125 cm (horizontal axes in Figure 2). The loss of information due to partitioning continuous variables into discrete levels was probably small because the distributions of all attributes were unimodal with very constant modes.

Step 3. Dividing the attribute space using all $m = 9$ dimensions at once would result in a huge number of cells, each corresponding to a very overly specified movement pattern (i.e., specifying nine different requirements in its definition) mostly including too few data points for a significant sample size and highly vulnerable to random variation. In this study we thus limit ourselves to all attribute subspaces of up to four dimensions out of the total nine. For example, Figure 2 shows the three-dimensional subspace including the three attributes d, a, and c_4. Each cell is denoted by an ID of the form $P\{i_1,i_2,i_3, \ldots ,i_m\}$ corresponding to the m-dimensional attribute vector, were i is the index of the bin in the corresponding attribute according to Table 1, and an asterisk denotes a attribute that is not relevant to the definition of the cell. For example, $P\{1,^*,1,^*,^*,4,^*,^*,^*\}$ denotes a cell in which values of the first attribute belong to the first bin in that attribute (i.e., distance from the wall $0 < d < 10$ cm), values of the third attribute belong to the first bin in that attribute (i.e., acceleration $a < -30$ cm/s^2), values of the sixth attribute belong to the fourth bin in that attribute (i.e., path curvature $2 < c_4 < 10$ degree/cm), and the other attributes (asterisks) are irrelevant to the definition of the pattern and can accept any value. Note that for all practical purposes $P\{1,^*,1,^*,^*,4,^*,^*,^*\}$ is a three-dimensional vector, not a nine-dimensional vector. Limiting the algorithm to four relevant attributes at most means that a total of 50,674 possible patterns having at least five asterisks are considered.

Step 4. In each cell we consider the relative frequency of data points falling into this cell, using the Logit transformation:

$$LogitFrequency(P\{i_1,i_2,i_3, \ldots ,i_m\}) = \log\left(\frac{k + 1/3}{l - k + 1/3}\right)$$

where k is the number of data points falling in this cell and l is the total number of data points for this animal (see Figure 2). The Logit transform is routinely used in statistics (e.g., logistic regression) to transform proportions bounded between 0 and 1 to real valued variables more amenable to statistical analysis, and adding 1/3 is a standard procedure for correcting the behavior of the transformation when $k = 0$.

Step 5. We discarded cells with very small support—attribute combinations that were hardly exhibited by most or all animals. Many combinations of attribute values are rarely used due to trivial physical constraints on movement (e.g., accelerating during a sharp turn at high speed), and other combinations are simply things that rats in general prefer to avoid (e.g., running toward the center of the arena at high speed and near-zero acceleration). Such physical and behavioral constraints, however, may differ across the experiment groups, and we attempted to avoid discarding a cell that generally has a low frequency if it is significantly more frequent in one of the groups. Therefore we computed the median *LogitFrequency* in each group of training set samples and discarded the cell only if the maximal group median (whatever group it is) is lower than *FrequencyCutoff*. In this study *FrequencyCutoff* was set to -5.5, which in an animal with typical activity of $l =$ 30,000 data points would mean using this pattern for about 120 data points (i.e., 4 cumulative seconds) or 0.4% of total progression time. After this step we are thus left with B_{nonneg} – the set of nonnegliable movement patterns.

Step 6. Discover the movement patterns in B_{nonneg} differing in relative frequency between experimental groups. In this study we apply a simple t-test to compare the training set mean *LogitFrequency* values between the SOD1 animals and the wild-type controls, and screen the subset of potentially significant movement patterns $B_{pot\text{-}sig} \subseteq B_{nonneg}$ using the Bonferroni criterion. That is, we test each null hypothesis at level α/n where n is the number of comparisons (i.e., the number of cells in B_{nonneg}) at a level of $\alpha = .05$.

Step 6a (optional). Within the remaining potentially significant patterns $B_{pot\text{-}sig}$ a high level of cross-pattern correlation might exist, especially because some of these patterns overlap in their definition. In this case it is possible to use a variety of procedures to screen these patterns further in a way that reduces cross-correlation. However, in our case the objective was to find at least one pattern that discriminates the mutant SOD1 animals from the control animals, and generally their behavior is so similar that very few differences, if at all, are likely to be found. We thus simply picked the most significant pattern out of $B_{pot\text{-}sig}$.

Step 7. Use test set samples to validate the discrimination ability of the movement patterns discovered in the training set. According to Benjamini and Yekutieli (2005) the test set inference must only be corrected for multiplicity for the $B_{pot\text{-}sig}$ screened patterns. If no patterns are found significant in the training set it may be possible to add the data from the test set to the training set to increase the sample size and hopefully detect a significant pattern in Step 6, at the cost of leaving no data for cross-validation in Step 7.

Results

Applying the Pattern Array algorithm in this study we divided the animals before the experiment into two batches A (7 mutants vs. 7 controls) and B (5 mutants vs. 5 controls). Batch A at PND 50 was used as the training set, and the discovered pattern was cross-validated in batch B at PND 50 and in both batches A and B at PND 80. One very inactive control animal in batch A had to be discarded from this Pattern Array analysis, both in PND 50 and 80, but it was still considered for the analysis of body weight, grip force, activity, and center time (see Figure 3).

In Step 5 of the algorithm, out of the total 50,674 behavioral patterns, 11,831 patterns were found common enough in the training set to pass *FrequencyCutoff*. The Bonferroni criterion at $\alpha = .05$ was thus set to $0.05/11831 = 4.226 \times 10^{-6}$, and only two patterns were found to be more significant than this criterion. The more significant of the two ($p = 2.9 \times 10^{-6}$) was $P\{1,^{*},1,^{*},^{*},4,^{*},^{*},^{*}\}$. This ID vector shows that six out of the nine attributes were irrelevant to this pattern and may take any value (asterisks). Of the others, the first attribute refers to the distance from the arena wall d, and the index of 1 in this place (see Table 1) denotes the lowest level of this attribute, which is less than 8 cm from the wall. The third attribute refers to the acceleration a, and 1 denotes the most negative acceleration level, actually a strong deceleration. The sixth attribute c_4 refers to path curvature (change of direction) in a scale of 4 cm (Kafkafi & Elmer, 2005), and the index 4 denotes a slight turn in the direction away from the arena wall. Thus $P\{1,^{*},1,^{*},^{*},4,^{*},^{*},^{*}\}$ is defined as braking strongly while moving very close to the wall but turning slightly away from it. An actual example of a rat performing this pattern can be seen in Figure 4, and additional examples out of the data are shown in the animations (see online supplemental information). At PND 50 the wild-type controls performed this pattern on average for about 1.8%

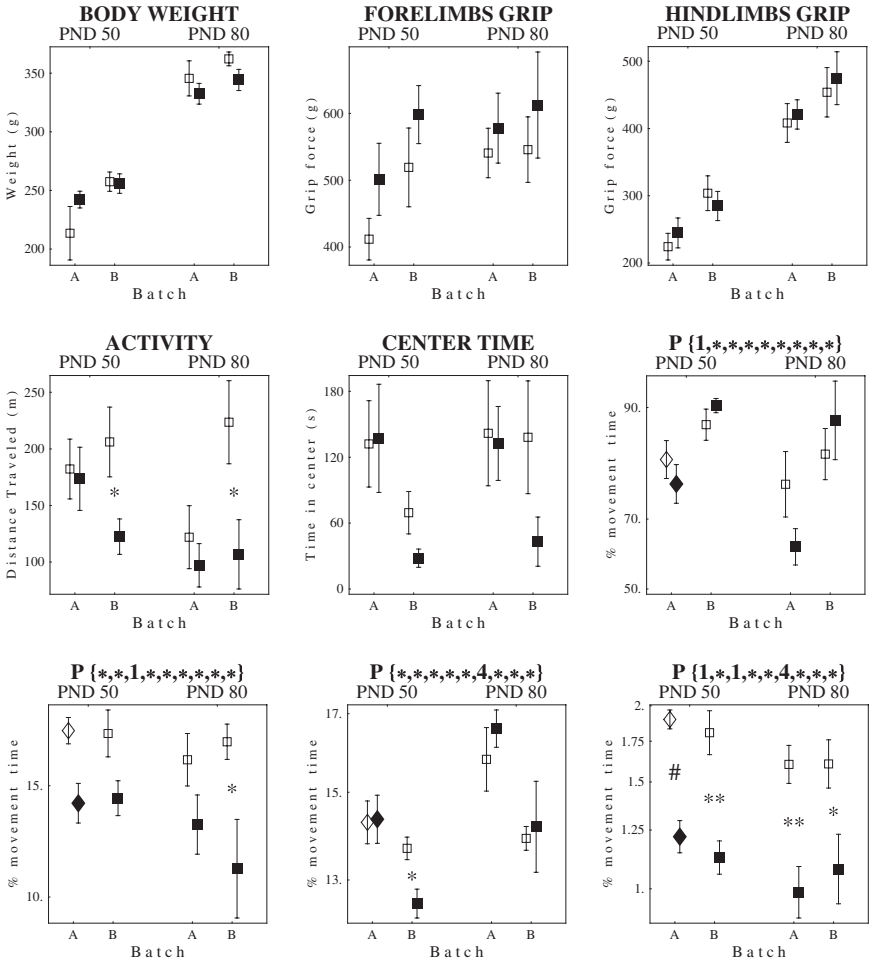

Figure 3. Results using several measures. Results from SOD1 mutants (closed squares) and Sprague–Dawley controls (open squares) in nine different measures including patterns $P\{1,^{*},^{*},^{*},^{*},^{*},^{*},^{*},^{*}\}$, $P\{^{*},^{*},1,^{*},^{*},^{*},^{*},^{*},^{*}\}$, $P\{^{*},^{*},^{*},^{*},^{*},4,^{*},^{*},^{*}\}$ and the discovered pattern $P\{1,^{*},1,^{*},^{*},4,^{*},^{*},^{*}\}$. Animals were divided into two batches, batch A ($n = 7$) and batch B ($n = 5$). Each batch was tested at two ages: PND 50 and PND 80. All results show group means and *SE*s. $^{*}p < .05$. $^{**}p < .01$. $\#p < .0000042$ (Bonferroni criterion at a level of .05 for the training set). Note that in the $P\{1,^{*},1,^{*},^{*},4,^{*},^{*},^{*}\}$ graph (bottom right), batch A in PND 50 (diamonds instead of squares) was used as the training set for discovering the pattern itself.

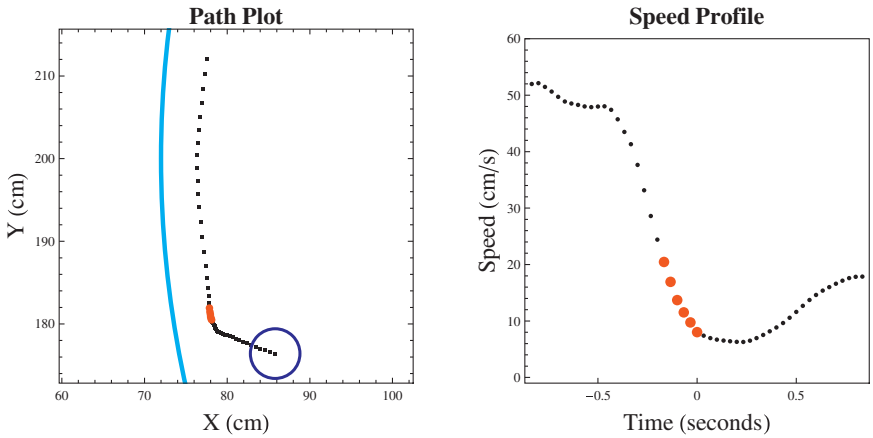

Figure 4. A rat performing the discovered pattern. Path plot (left) and speed profile (right) of a particular example out of the data of a Sprague–Dawley rat performing the discovered pattern $P\{1,^{*},1,^{*},^{*},4,^{*},^{*},^{*}\}$. Each data point represents 1/30 seconds and the six points belonging to the pattern are highlighted in red. The arc in the path plot (cyan) denotes the arena wall and the circle (blue) represents the rat, going from top to bottom of the graph. Notice the turn out of the wall in the path plot and the strong deceleration (negative slope) in the speed profile.

of their progression time (Figure 3, bottom right), whereas the SOD1 animals performed it on average for only 1.2% of their progression time. This pattern was then tested for significance in the test set.

Our results in the standard behavioral measures agree with the inability of previous studies to discover any consistent SOD1 effect before PND 100 (Matsumoto et al., 2006). Figure 3 shows the results in body weight, grip strength of forelimbs and hind limbs, and the open-field behavior using two standard measures: general activity (distance traveled) and center time. Body weight and grip strength of either forelimbs or hind limbs have increased as expected from PND 50 to PND 80, but there was no significant difference between mutants and controls. In batch B the mutants were significantly less active in both ages, PND 50: $t = 2.42$; $p < .05$; PND 80: $t = 2.44$; $p < .05$; but this difference was not replicated in batch A. No significant differences were found in the center time.

Furthermore, no single attribute in Pattern Array could differentiate the SOD1 mutants by itself. Figure 3 also shows results in three single-attribute patterns of Pattern Array: $P\{1,^{*},^{*},^{*},^{*},^{*},^{*},^{*},^{*},^{*}\}$ (i.e., moving near the wall), $P\{^{*},^{*},1,^{*},^{*},^{*},^{*},^{*},^{*}\}$ (i.e., braking strongly), and $P\{^{*},^{*},^{*},^{*},^{*},4,^{*},^{*},^{*}\}$ (i.e., turning slightly away from the wall). These patterns are shown here because their intersection is the discovered three-attributes pattern $P\{1,^{*},1,^{*},^{*},4,^{*},^{*},^{*}\}$. As Figure 3 shows, $P\{1,^{*},^{*},^{*},^{*},^{*},^{*},^{*},^{*}\}$ failed in significantly differentiating the SOD1 rats from the controls. $P\{^{*},^{*},^{*},^{*},^{*},4,^{*},^{*},^{*}\}$ just barely passed significance in one comparison out of four, batch B, PND 50: $t = 3.01$; $p < .05$; and so did $P\{^{*},^{*},1,^{*},^{*},^{*},^{*},^{*},^{*}\}$ Batch B, PND 80: $t = 2.42$; $p < .05$. Thus each of these three more general patterns could not, by itself, consistently differentiate the SOD1 animals.

In contrast, the intersection of these three patterns, the screened pattern $P\{1,^{*},1,^{*},^{*},4,^{*},^{*},^{*}\}$, consistently differentiated the two groups (Figure 3, bottom right) with the SOD1 mutants always performing it significantly less than the wild-type controls. Note that the small variability in batch A at PND 50 (diamonds instead of squares) might be misleading because these data were the "training set" used for the very discovery of this pattern, and by definition it was the most significant out of the 50,674 tested patterns. However, this pattern was also significant in the test sets: batch B at PND 50: $t = 4.5$; $p < .01$; $n = 5, 5$; batch A at PND 80: $t = 4.0$; $p < .01$; $n = 7, 7$; and batch B at PND 80: $t = 2.5$; $p < .05$; $n = 5, 5$, all using t-test. Because here we only consider a single pattern there is no need to correct the test set results for multiplicity. Batch A in PND 50 and PND 80 are the same animals in different ages, and therefore the second result is not independent of the first. However, batch B is an independent validation of batch A in both PND 50 and 80. Note also that batch B not only replicated the differences discovered in batch A, but also the absolute frequencies of performing the pattern were similar. There was actually no overlap at all between the results of the SOD1 animals and the controls in batch A (both PND 50 and 80) and batch B (in PND 50), and only very slight overlap in batch B in PND 80. This indicates that the discovered pattern may be powerful enough to diagnose single SOD1 animals with high confidence.

Discussion

Implications for the SOD1 Rat Model of ALS

SOD1 mutant animals are generally considered presymptomatic before PND 80 in mice (Chiu et al., 1995; Derave et al., 2003; Fischer et al., 2004; Weydt, Hong, Kliot, & Moller, 2003) and PND 90 in rats (Matsumoto et al., 2006). In full agreement with these studies we could not detect a consistent and significant effect of the mutation in our rats at either PND 50 or PND 80 using the grip strength (the common measure of disease onset in the SOD1 animals) or other standard measures and tests. Furthermore, in agreement with Matsumoto et al. and our own past experience, no difference between the SOD1 rats and the wild-type controls could even be detected by subjective observation of their behavior before PND 90. In contrast, the Pattern Array method was able to discover a movement pattern that significantly and consistently dif-

ferentiated the SOD1 rats at PND 50 and 80 in comparatively small group sizes (5 to 7 animals). This difference may be related to the denervation found in hind leg muscles—*gastrocnemius*, *soleus*, and *tibialis anterior*—found in SOD1 mice, which included 40% of end-plates by PND 47 and continued to progress up to the time of death (Fisher et al., 2004). The discovered premorbid symptom may enable investigators to test treatments for delaying or even preventing the disease.

Performance of the isolated pattern slightly decreased in both mutants and controls as a function of age (Figure 3, bottom right). This is reasonable because they were heavier by PND 80, which makes strong braking more difficult. The pattern did not detect any increase of the difference between mutants and controls from PND 50 to PND 80, which would be expected if early symptoms were getting worse as the mutants approach the age of disease onset. However, note that we only used here the simplest application of Pattern Array, in which patterns are mined based on *t*-test comparison between the experiment and control group in one case (the PND 50). A more advanced study may test a training set along several ages using, for example, two-way analysis of variance (ANOVA) of Genotype $\times$ Age with age as a repeated measure. By mining patterns with both large genotype effect and large Genotype $\times$ Age interaction effect, Pattern Array can be configured to discover a pattern that more specifically tracks disease progression. In principle many kinds of statistical tests can be used in Pattern Array, depending on the design of the data and the objective of testing. In any case, additional studies will be needed to establish the normal range of the effect in additional animals as well as its replicability in additional laboratories.

Although it is not clear yet why this specific pattern is performed less by the mutants, Pattern Array can be used to explore the results of similar patterns to gain some insight regarding the important characteristics, as Figure 3 illustrates with the three single-attribute patterns. The consistency of results in $P\{^*,^*,1,^*,^*,^*,^*,^*\}$ suggests that the mutants are generally deficient in strong deceleration, which for physical reasons (Newton's second law) is proportional to the force acting on the feet. This is consistent with the denervation in 40% of the end-plates of the hind leg muscles, reported in SOD1 mice by PND 47 (Fisher et al., 2004). However, the grip force test was not able to discover this deficiency, perhaps because of the inevitable stress associated with it or because of some other confounding conditions. It may be that the other two components of the discovered pattern—moving near the wall and the slight turn—are important not in themselves (as suggested by their inconsistency in Figure 3) but because they happen to provide specific conditions in which the difference in braking is more pronounced and consistent. It is of course possible to continue exploring the results in any number of additional patterns, although multiplicity considerations should to be addressed in such case. It should be stressed, however, that no single behavioral component in this study was able to detect the early symptom in a reliable manner. Such reliable detection was only achieved by a unique interaction of three components, an interaction that was not likely to have been foreseen from the outset. Additional studies focused on the discovered pattern will be needed to explore whether and how it is correlate with a neurophysiological endpoint, for example, by any early denervation in the hind leg muscles (Fisher et al., 2004).

Potential of Pattern Array for Behavioral Phenotyping

A considerable portion of current research in the field of behavioral phenotyping may be described as attempts to answer the question "what's wrong with my genetically engineered animal?" using a battery of behavioral tests (Crawley, 2000). The SOD1 mutant rat is discussed here as one typical case of an animal model in which the standard tests fail to detect some desirable effect (for other examples, see Grammer et al., 2005; Perez & Palmiter, 2005). Even when significant behavioral effects are detected they might prove difficult to reproduce in other laboratories (Crabbe, Wahlsten, & Dudek, 1999; Kafkafi, et al., 2005) or in slightly different conditions (Chesler, Wilson, Lariviere, Rodriguez-Zas, & Mogil, 2002; Valdar et al., 2006; Wahlsten, Rustay, Metten, & Crabbe, 2003). The discovery of reliable behavioral endpoints with predictive validity, even before a good understanding of their etiology is achieved, can significantly improve intervention research (Willner, 1991). However, despite the obvious need and the large amounts of already existing raw data, simple strategies of data mining and exploratory data analysis are rarely employed in the field. The current phenotyping databases (e.g., the Mouse Phenome Database, see Paigen & Eppig, 2000) only store endpoint results in standard behavioral measures, not the raw data themselves. Data mining strategies like Pattern Array can test tens of thousands of hypotheses in parallel, thus improving the chances of discovering an effect by making better use of data from currently available tests. Although this study utilized data from the open-field SEE test, the Pattern Array algorithm can in principle be adapted in a relatively straightforward manner, by choosing appropriate sets of attributes, to other spatial tests employing automated tracking, and even to other tests that record large amounts of raw data. In many cases such data may have already been measured and stored, but used only for computing a small number of traditional behavioral endpoints.

There are several additional methods that can be adapted for mining subtle behavioral differences, each likely to have its advantages and disadvantages. For example, linear discriminant analysis (LDA) or partial least squares (PLS) can be used without arbitrarily partitioning the range of continuous measures of behavior into discrete patterns, although it would still be necessary to specify the series of potential attributes. An important consideration in choosing our approach in this study was the interpretability and utility of any finding. In LDA and similar standard discrimination methods a finding would consist of an abstract formula for computing the posterior probability that an animal has the SOD1 mutation out of its path data. At best this formula would tell us which are the important attributes and whether they increase or decrease this probability. In contrast, the approach we take in this study explicitly specifies the differentiating behavioral pattern, and can trace it back to particular events in the data (e.g., Figure, 4 and Animations 1-4 in the online supplemental material) or a videotaped session. It is thus possible to focus on this pattern, possibly reproduce it in specialized experimental setups, and explore how it is affected by the SOD1 mutation. Note that once focusing on a single pattern as the endpoint of interest the measurement process becomes technically much simpler because the data-mining algorithm is not required anymore. Another important advantage of our method is that it tests a broad range of patterns from the very general (single-attribute "patterns," each typically occurring for

about 20% or more of the total movement time) to the very specific (four-attribute patterns, each typically occurring for less than 1%).

It is important to stress that any data-mining approach would have to depend on some a priori scheme for dissecting behavior, for example, the introduction of attributes and cutoff values in Pattern Array, and in this sense would still require, like any other standard behavioral test, certain insight into the nature of the expected behavioral effect. However, by testing a large number of combinations in parallel, data mining methods may considerably enhance the power of human insight to detect behaviors that are affected by the genetic manipulation of interest, especially complex combinations that are not easy to guess from the outset. The use of a strict multiple-comparisons criterion in the training set and cross-validation in a test set can ensure that the parameters were not selected to discover circumstantial differences in the particular data set.

Like most data mining approaches, Pattern Array is constrained by statistical considerations of multiple comparisons, in this case the total number of patterns to be tested, which is determined by the number of attributes and the number of bins in each attribute. Too few patterns will decrease the chance to isolate one that is affected by the genetic manipulation, although too many patterns will result in an overly restrictive multiple-comparisons criterion, which might result in failing to identify an affected pattern as significant even if it was tested. Increasing the number of animals and the number of data points per animal will generally increase the level of significance of true positives, thus increasing the number of patterns that can be tested and the chance to pinpoint on the most appropriate pattern. Using density functions instead of arbitrary cutoffs may eliminate artifacts and enable more precise framing of the looked-for difference, but this again will increase the number of hypotheses to be tested. Thus density functions may prove effective in later experiments, after the nature of the difference was already established and irrelevant attributes can be dropped from the analysis.

Pattern array does not treat path data as a time series because this again would necessitate a huge number of comparisons. Instead it captures patterns of movement by employing dynamical attributes such as momentary speed, momentary acceleration, and momentary change of direction. Note, however, that the dynamical attributes in this study all have a short time scale (mostly estimated with a window size of half a second or less) and are therefore appropriate for detecting brief behavior patterns of the kind that is usually associated with motor symptoms. Such attributes are unlikely to detect more prolonged behaviors that last several seconds or even minutes, and are usually associated with more cognitive functions. In principal it is possible to use Pattern Array with attributes designed for longer time scales, but for a given session duration this will decrease the number of data points and consequently the power to detect an effect. In this study we focused on attributes of short time scale because the symptoms of the SOD1 mutation were expected to be motor in nature.

Data mining strategies can be used to search for behaviors related to any experimental manipulation of interest, and therefore have potential in fields such as psychopharmacology and toxicology. For example, in an ongoing study the same Pattern Array algorithm and attributes are employed in drug-injected mice, screening patterns that best classify psychomotors, opioids, and NMDA antagonists (Kafkafi et al., in press). The Pattern Array method fits well into the approach proposed by Kafkafi et al. (2005) of keeping databases of raw behavioral data from many experiments, treatments, and laboratories. Once a new pattern is detected in one experiment, using Pattern Array or any other method, it can be immediately tested over the whole database, thus gaining insight into its meaning, consistency, and generality. In such a strategy the data from each experiment may be useful beyond merely confirming or rejecting the original hypothesis.

References

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., & Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. *Behavioural Brain Research, 125,* 279–284.

Benjamini, Y., & Yekutieli, D. (2005). False discovery rate-adjusted multiple confidence intervals for selected parameters. *Journal of the American Statistical Association, 100,* 469, March 2005, 71–81(11).

Broadhurst, P. L. (1975). The Maudsley reactive and nonreactive strains of rats: A survey. *Behavior Genetics, 5,* 299–319.

Bruijn, L. I., Becher, M. W., Lee, M. K., Anderson, K. L., Jenkins, N. A., Copeland, N. G., et al. (1997). ALS-linked SOD1 mutant G85R mediates damage to astrocytes and promotes rapidly progressive disease with SOD1-containing inclusions. *Neuron, 18,* 327–338.

Bruijn, L. I., Miller, T. M., & Cleveland, D. W. (2004). Unraveling the mechanisms involved in motor neuron degeneration in ALS. *Annual Review of Neuroscience, 27,* 723–749.

Chesler, E. J., Wilson, S. G., Lariviere, W. R., Rodriguez-Zas, S. L., & Mogil, J. S. (2002). Influences of laboratory environment on behavior. *Nature Neuroscience, 5,* 1101–1102.

Chiu, A. Y., Zhai, P., Dal Canto, M. C., Peters, T. M., Kwon, Y. W., Prattis, S. M., et al. (1995). Age-dependent penetrance of disease in a transgenic mouse model of familial amyotrophic lateral sclerosis. *Molecular Cell Neuroscience, 6,* 349–362.

Crabbe, J. C., Wahlsten, D., & Dudek, B. C. (1999). Genetics of mouse behavior: Interactions with laboratory environment. *Science, 284*(5420), 1670–1672.

Crawley, J. N. (2000). *What's wrong with my mouse? Behavioral phenotyping of transgenic and knockout mice*. New York: Willey-Liss.

Derave, W., Van Den Bosch, L., Lemmens, G., Eijnde, B. O., Robberecht, W., & Hespel, P. (2003). Skeletal muscle properties in a transgenic mouse model for amyotrophic lateral sclerosis: Effects of creatine treatment. *Neurobiology of Disease, 13,* 264–72.

Drai, D., Benjamini, Y., & Golani, I. (2000). Statistical discrimination of natural modes of motion in rat exploratory behavior. *Journal of Neuroscience Methods, 96,* 119–131.

Drai, D., & Golani, I. (2001). SEE: A tool for the visualization and analysis of rodent exploratory behavior. *Neuroscience and Biobehavioral Reviews, 25,* 409–426.

Drai, D., Kafkafi, N., Benjamini, Y., Elmer, G. I., & Golani, I. (2001). Rats and mice share common ethologically relevant parameters of exploratory behavior. *Behavioural Brain Research, 125,* 133–140.

Fischer, L. R., Culver, D. G., Tennant, P., Davis, A. A., Wang, M., Castellano-Sanchez, A., et al. (2004). Amyotrophic lateral sclerosis is a distal axonopathy: Evidence in mice and man. *Experimental Neurology, 185,* 232–240.

Grammer, M., Kuchay, S., Chishti, A., & Baudry, M. (2005). Lack of phenotype for LTP and fear conditioning learning in calpain 1 knock-out mice. *Neurobiology of Learning and Memory, 84,* 222–227.

Gurney, M. E., Pu, H., Chiu, A. Y., Dal Canto, M. C., Polchow, C. Y., Alexander, D. D., et al. (1994). Motor neuron degeneration in mice that express a human Cu, Zn superoxide dismutase mutation. *Science, 264*(5166), 1772–1775.

Hen, I., Sakov, A., Kafkafi, N., Golani, I., & Benjamini, Y. (2004). The

dynamics of spatial behavior: How can robust smoothing techniques help? *Journal of Neuroscience Methods 133,* 161–172.

Horev, G., Benjamini, Y., Sakov, A., & Golani, I. (2007). Estimating wall guidance and attraction in mouse free locomotor behavior. *Genes Brain Behavior, 6*(1), 30–41.

Howland, D. S., Liu, J., She, Y., Goad, B., Maragakis, N. J., Kim, B., et al. (2002). Focal loss of the glutamate transporter EAAT2 in a transgenic rat model of SOD1 mutant-mediated amyotrophic lateral sclerosis (ALS). *Proceedings of the National Academy of Sciences, USA, 99,* 1604–1609.

Kafkafi, N. (2003). Extending SEE for large-scale phenotyping of mouse open-field behavior. *Behavior Research Methods, Instruments, & Computers, 35,* 294–301.

Kafkafi, N., Benjamini, Y., Sakov, A., Elmer, G. I., & Golani, I. (2005). Genotype-environment interactions in mouse behavior: A way out of the problem. *Proceedings of the National Academy of Sciences of the United States of America, 102,* 4619–4624.

Kafkafi, N., & Elmer, G. I. (2005). Texture of locomotor path: A replicable characterization of a complex behavioral phenotype. *Genes Brain Behavior, 4,* 431–443.

Kafkafi, N., Lipkind, D., Benjamini, Y., Mayo, C. L., Elmer, G. I., & Golani, I. (2003). SEE locomotor behavior test discriminates C57BL/6J and DBA/2J mouse inbred strains across laboratories and protocol conditions. *Behavioral Neuroscience, 117,* 464–477.

Kafkafi, N., Mayo, C., Drai, D., Golani, I., Elmer, G. (2001). Natural segmentation of the locomotor behavior of drug-induced rats in a photobeam cage. *Journal of Neuroscience Methods, 109,* 111–21.

Kafkafi, N., Pagis, M., Lipkind, D., Mayo, C. L., Benjamini, Y., Elmer, G. I., et al. (2003). Darting behavior: A quantitative movement pattern for discrimination and replicability in mouse locomotor behavior. *Behavioural Brain Research, 142,* 193–205.

Kafkafi, N., Yekutieli, D., & Elmer, G. I. (in press). A data mining approach to in vivo classification of psychopharmacological drugs. *Neuropsychopharmacology.*

Leppanen, P. K., Ewalds-Kvist, S. B., & Selander, R. K. (2005). Mice selectively bred for open-field thigmotaxis: Life span and stability of the selection trait. *Journal of General Psychology 132,* 187–204.

Lipkind, D., Sakov, A., Kafkafi, N., Elmer, G. I., Benjamini, Y., & Golani, I. (2005). New replicable anxiety-related measures of wall vs. center behavior of mice in the open field. *Journal of Applied Physiology, 97,* 347–359.

Matsumoto, A., Okada, Y., Nakamichi, M., Nakamura, M., Toyama, Y., Sobue, G., et al. (2006). Disease progression of human SOD1 (G93A) transgenic ALS model rats. *Journal of Neuroscience Research, 83,* 119–33.

National Academy of Sciences. (1996). *Guide for the care and use of laboratory animals.* (DHEW Publication No. 86–23). Washington, DC: Author. Available at http://www.nap.edu/readingroom/books/labrats/

Paigen, K., & Eppig, J. T. (2000). A mouse phenome project. *Mammalian Genome, 11,* 715–717.

Perez, F. A., & Palmiter, R. D. (2005). Parkin-deficient mice are not a robust model of parkinsonism. *Proceedings of the National Academy of Sciences, USA, 102,* 2174–2179.

Ramos, A., Berton, O., Mormede, P., & Chaouloff, F. (1997). A multiple-test study of anxiety-related behaviours in six inbred rat strains. *Behavioural Brain Research, 85*(1), 57–69.

Rosen, D. R., Siddique, T., Patterson, D., Figlewicz, D. A., Sapp, P., Hentati, A., et al. (1993). Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature, 362*(6415), 59–62.

Rothstein, J. D., Patel, S., Regan, M. R., Haenggeli, C., Huang, Y. H., Bergles, D. E., et al. (2005). Beta-lactam antibiotics offer neuroprotection by increasing glutamate transporter expression. *Nature, 433,* 73–77.

Shaw, C. A., & Wilson, J. M. B. (2006). Environmental toxicity and ALS: Novel insights from an animal model of ALS-PDC. In. H. Mitsumoto, S. Przedborski & Gordon, P. H., (Eds.), *Amyotrophic lateral sclerosis* (pp. 435–448). New York: Taylor & Francis.

Valdar, W., Solberg, L. C., Gauguier, D., Cookson, W. O., Rawlins, J. N., Mott, R., et al. (2006). Genetic and environmental effects on complex traits in mice. *Genetics, 174,* 959–84.

Wahlsten, D., Rustay, N. R., Metten, P., & Crabbe, J. C. (2003). In search of a better mouse test. *Trends in Neurosciences, 26,* 132–136.

Weydt, P., Hong, S. Y., Kliot, M., & Moller, T. (2003). Assessing disease onset and progression in the SOD1 mouse model of ALS. *Neuroreport, 14,* 1051–1054.

Willner, P. (1991). Methods for assessing the validity of animal models of human psychopathology. In A. A. Boulton, G. B. Baker, & M. T. Martin-Iversen (Eds.), *Animal models in psychiatry, 1* (18th ed., pp. 1–24). Totowa, NJ: Humana.

Received November 20, 2007
Revision received February 27, 2008
Accepted March 3, 2008 ■