Commentary

Mammalian Genome

Mammalian Genome 8, 711–713 (1997).
© Springer-Verlag New York Inc. 1997

Behavioral and functional analysis of mouse phenotype: SHIRPA, a proposed protocol for comprehensive phenotype assessment

Derek C. Rogers,[1] E.M.C. Fisher,[2] S.D.M. Brown,[3] J. Peters,[3] A.J. Hunter,[1] J.E. Martin[4]

[1]Neuroscience Research, SmithKline Beecham Pharmaceuticals, New Frontiers Science Park, Harlow, UK
[2]Neurogenetics Unit, Imperial College School of Medicine at St. Mary's, London, UK
[3]MRC Mouse Genome Centre and Mammalian Genetics Unit, Harwell, UK
[4]Department of Morbid Anatomy, St. Bartholemew's and the Royal London School of Medicine, London, UK

Received: 3 March 1997 / Accepted: 30 May 1997

Abstract. For an understanding of the aberrant biology seen in mouse mutations and identification of more subtle phenotype variation, there is a need for a full clinical and pathological characterization of the animals. Although there has been some use of sophisticated techniques, the majority of behavioral and functional analyses in mice have been qualitative rather than quantitative in nature. There is, however, no comprehensive routine screening and testing protocol designed to identify and characterize phenotype variation or disorders associated with the mouse genome. We have developed the SHIRPA procedure to characterize the phenotype of mice in three stages. The primary screen utilizes standard methods to provide a behavioral and functional profile by observational assessment. The secondary screen involves a comprehensive behavioral assessment battery and pathological analysis. These protocols provide the framework for a general phenotype assessment that is suitable for a wide range of applications, including the characterization of spontaneous and induced mutants, the analysis of transgenic and gene-targeted phenotypes, and the definition of variation between strains. The tertiary screening stage described is tailored to the assessment of existing or potential models of neurological disease, as well as the assessment of phenotypic variability that may be the result of unknown genetic influences. SHIRPA utilizes standardized protocols for behavioral and functional assessment that provide a sensitive measure for quantifying phenotype expression in the mouse. These paradigms can be refined to test the function of specific neural pathways, which will, in turn, contribute to a greater understanding of neurological disorders.

Background

Transgenic and gene-targeting techniques with mice are increasingly common tools in the study of human gene function. Spontaneous and induced mouse mutations are also a critical resource in the characterization of previously uncloned genes of biological importance, and the analysis of phenotypic differences between strains has become important, for example, in dissecting the effect of genetic background. A mouse model of human disease or gene function is, however, of limited value unless properly characterized, and accurate phenotype assessment is the jewel in the crown of genetic manipulation. The vast majority of behavioral data in experimental animals have been obtained from rats, but increasingly structural and functional analysis is focused on the mouse. Phenotype analysis of mice has tended to be qualitative rather than quantitative in nature, with, for example, reports of grooming, eating, and handling response (Chang et al. 1993). Some groups have applied more sophisticated tests of locomotor and cognitive function to mice. For example, animals in which a portion of the mouse genome corresponding to human Chromosome (Chr) 21 was trisomic have demonstrated learning deficits (Reeves et al. 1995), and gene-targeted animals in which the β-amyloid precursor protein gene or the Huntington gene have been disrupted have also displayed clear behavioral and functional deficits (Muller et al. 1994; Nasir et al. 1995). Such studies have usually been carried out in isolation, and these protocols make the assumption, often tacit, that other phenotypic variables are normal.

One key to accurate phenotype analysis in all species is to ensure that the initial stages of assessment are always performed impartially and that features are not missed, or findings anticipated or biased by making premature assumptions about end results. Methods for behavioral, functional, and morphological characterization should ideally be directly comparable between different groups to build a useful body of data by which one can reproducibly assess the effect of a genetic manipulation or strain variation, the usefulness of a model, and the effectiveness of a therapeutic intervention.

A large number of mouse strains are in common use. There are considerable behavioral differences between strains, for example, in the performance of learning and memory tasks (Paylor et al. 1994) and response to acoustic startle (Paylor and Crawley in press). Obtaining appropriate control data is, therefore, critically important. In behavioral testing, a number of factors including nutritional state, circadian rhythm, oestrus cycle time, sex, and age-related differences may affect the outcome and must be taken into account in the design of experiments (Gower and Laberty 1993). An appreciation of the range of occasional findings in normal mice is also necessary at both behavioral and structural levels. To date, there is no agreed screening and testing protocol designed to identify and characterize disorders arising from abnormalities in the mouse genome. We propose that SHIRPA, a systematic, objective protocol for phenotype analysis, can provide the framework to meet this need.

SHIRPA—a proposed protocol for phenotype assessment

SHIRPA (Fig. 1) is a three-stage protocol designed as a series of individual tests that in themselves provide quantitative data about an individual performance. Such test-specific performance is directly comparable between animals, over time, and between groups. In addition, the tests also provide the opportunity to define abnormalities or variation in the mouse, with each test providing information about the pattern of function of a particular system, for example the brain and neuromuscular system. Collation and analy-

Correspondence to: D.C. Rogers

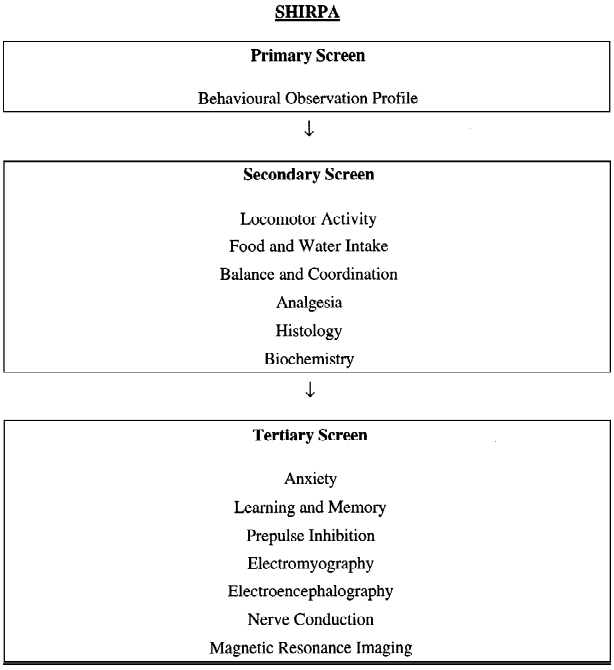

Fig. 1.

Table 1. Tests contributing to the assessment of specific functions.

Muscle and lower motor neurone function		
Body position	Spontaneous activity	Locomotor activity
Gait	Tail elevation	Limb position
Positional passivity	Visual placing	Body tone
Wire maneuver	Limb tone	Abdominal tone
Righting reflex	Passivity	Urination and defecation
Motor performance	Balance and coordination	
Spinocerebellar function		
Body position	Locomotor activity	Gait
Tail elevation	Limb position	Visual placing
Grip strength	Body tone	Limb tone
Abdominal tone	Righting reflex	Motor performance
Balance and coordination		
Sensory function		
Transfer arousal	Gait	Limb position
Touch escape	Visual placing	Pinna reflex
Corneal reflex	Toe pinch	Righting reflex
Analgesia		
Neuropsychiatric function		
Body position	Spontaneous activity	Bizarre behavior
Transfer arousal	Locomotor activity	Food and water intake
Startle response	Touch-escape	Positional passivity
Body tone	Righting reflex	Catalepsy
Fear	Irritability	Aggression
Anxiety	Vocalization	Prepulse inhibition
Learning and memory		
Autonomic function		
Palpebral closure	Piloerection	Startle response
Tail elevation	Skin color	Salivation
Temperature	Food and water intake	Respiration rate
Heart rate	Urination	Defecation

sis of data from all the tests can be summed to give a specific profile of function (Table 1). This approach mimics the diagnostic process of general, neurological, and psychiatric examination in humans. Measurement of locomotor activity, for example, gives data regarding the integrated function of cortical arousal, cerebral locomotor control, and neuromuscular function. This test can be taken alone as a coarse functional indicator, but when it forms part of a screening panel including tests of motor strength and neural control, such as grip strength and limb tone, it helps to define specific deficits. Thus, it may become apparent that a poor performance in locomotor activity tests reflects a muscular weakness in the absence of central nervous system dysfunction. The functional tests are performed in a sequence that is easy to follow while disturbing the animal as little as possible. The primary and secondary screens are of high throughput such that an experienced operator can assess up to 50 mice per day and use equipment that is relatively simple and inexpensive. The information derived from each test in the primary and secondary screens can also be grouped as tests contributing to an assessment of specific functions.

The primary screen is based on that developed by Irwin (1968), which has been widely used for screening drug candidates by pharmaceutical laboratories. This standard method provides a behavioral and functional profile by observational assessment of mice. This test will indicate defects in gait or posture, motor control and coordination, changes in excitability and aggression, salivation, lacrimation, piloerection, defecation, muscle tone, and temperature. It also provides a gross measure of analgesia. All parameters are scored to provide a quantitative assessment that enables comparison of results both over time and between different laboratories.

The secondary screen involves a comprehensive behavioral and functional screening battery and pathological analysis. This includes measurement of spontaneous locomotor activity in the horizontal and vertical planes, and during these tests, food and water intake is also monitored. Tests of motor performance are carried out to confirm and quantify effects observed in the behavioral profile. Balance and coordination is quantified with an accelerating rota-rod, which measures the ability of the mice to remain on a rotating drum (Jones and Roberts 1968). The effects on the perception of pain are measured with the mouse hot plate test (O'Callaghan and Holtzman 1975), a well established test of nociception. In addition, histopathological analysis of animals is carried out, and biochemical studies to measure serum urea, electrolyte, and blood glucose levels are performed to identify major metabolic deficits. These tests are carried out in conjunction with mass spectroscopic analysis of dried blood spot samples to screen for amino acid and intermediate metabolism defects.

These primary and secondary protocols provide a basic phenotype assessment that can be used in a wide range of applications, including the characterization of spontaneous and induced mutants, and the analysis of transgenic and knockout phenotypes. Additional specialist investigations can be added to any stage; for example, in a large-scale mutagenesis program we have an interest in picking up possible behavioral abnormalities that may be related to human schizophrenia. For these studies, we have introduced prepulse inhibition (PPI) testing into the primary screen, using a paradigm developed for rats and recently applied to the study of mice (Willott and Carlson 1995), in which the response to a loud acoustic stimulus is attenuated if it is shortly preceded by a low-level prepulse stimulus. The phenomenon is present in unaffected humans and animals, but is reduced in schizophrenic patients (Braff et al. 1978).

The tertiary screening stage in our protocol is tailored to the analysis of neurological mutants and is suitable for the assessment of existing or potential models of neurological disease, as well as the assessment of phenotypic variability that may be the result of known or uncharacterized genetic influences. There are a number of limitations to this approach, and the identification of suitable models for psychiatric disorders such as depression is difficult at present. A number other mouse models of neurological disorders are, however, relatively well characterized. In our protocol, anxiety is measured by open-field activity (Gerlai et al. 1996), in which

patterns of exploration in a brightly lit arena are recorded, and the elevated plus maze (Lister 1987). In addition, other models with proven construct validity such as the light-dark box (Crawley 1981) can also be used. Learning and memory models have been used previously in the assessment of mutant and gene-targeted mice (Bach et al. 1995; Reeves et al. 1995; Tonegawa et al. 1995), and our protocol uses the Morris water maze (Morris 1981), in which mice have to locate a submerged platform using distal cues. This comprehensive assessment is complemented by electromyography, electroencephalography, nerve conduction, and magnetic resonance imaging techniques, which employ well-established methods in the analysis of structural and functional abnormalities of the nervous system.

SHIRPA utilizes standardized protocols for behavioral and functional assessment that provide a sensitive measure for quantifying phenotype expression in the mouse. These paradigms can be refined to test the function of specific neural pathways, which will, in turn, contribute to a greater understanding of neurological and affective disorders.

Methods

SHIRPA (**S**mithKline Beecham Pharmaceuticals; **H**arwell, MRC Mouse Genome Centre and Mammalian Genetics Unit; **I**mperial College School of Medicine at St Mary's; **R**oyal London Hospital, St Bartholomew's and the Royal London School of Medicine; **P**henotype **A**ssessment, Fig. 1) involves three stages; the first two give a detailed general phenotype assessment, and the tertiary stage provides a specialized screen. The protocol detailed below has a tertiary screening stage primarily tailored to neurological deficits, but alternative specialist screening methods can be used; for example, blood pressure, heart rate, and angiographic studies may be carried out to identify cardiovascular defects. Full details of the SHIRPA protocol are available on the Harwell Web site: http://www.mgu.har.mrc.ac.uk/MGU-welcome.html

Primary screen. The primary screen provides a behavioral observation profile (Fig. 1), and assessment of each animal begins by observing undisturbed behavior in a viewing jar. In addition to the scored behaviors, the observer looks for manifestation of bizarre or stereotyped behavior, and convulsions, compulsive licking, self-destructive biting, retropulsion, and indications of spatial disorientation are recorded. Thereafter, the mouse is transferred to the arena for testing of transfer arousal and observation of normal behavior. There follows a sequence of manipulations with tail suspension and use of the grid across the width of the arena. To complete the assessment, the animal is restrained in a supine position to record autonomic behaviors prior to measurement of the righting reflex. Throughout this procedure, vocalization, urination, and general fear, irritability, or aggression are recorded.

Secondary screen. The secondary screen involves a comprehensive behavioral screening battery and pathological analysis (Fig. 1): *Spontaneous Locomotor Activity (LMA)*—carried out in a bank of activity monitors containing infrared beams; LMA and rearing are recorded over a period of 24 h; *Food and Water Intake*—monitored during the LMA test by weighing of the food and water dispensers; *Motor Performance*—the ability to grip a horizontal bar with forelimbs or maintain position on an inclined plane are recorded; *Balance and Coordination*—mice are placed on the moving drum of a rota-rod, which is then accelerated until the animals fall from the drum onto a plate to stop the timer; *Analgesia*—mice are placed on a hot-plate set at 52.5°C, and latency to forepaw licking or hindpaw fanning is recorded; *Histology*—a full study is made of all external and internal structures, including X-ray study of the whole animal prior to dissection for skeletal structures, macroscopic inspection, weighing of all organs, and histological sampling. The mice are perfusion fixed with formol saline for studies of the brain, spinal cord, peripheral nerve and muscle studies. In addition, frozen tissue is taken for studies of metabolic or mitochondrial disorders, and ultrastructural studies may require paraformaldehyde or glutaraldehyde fixation. Depending on the nature of disorder under study, histological techniques may be used in addition to the standard hemotoxylin and eosin for organs and Luxol fast blue-cresyl violet for brain and spinal cord. These techniques include a wide range of histochemical methods to identify inclusion material and immunohistochemical studies of glial, myelin and cytoskeletal proteins, and ubiquitin; *Biochemistry*—a standard multichannel analyzer is used for serum analysis of urea and electrolytes, blood glucose levels are analyzed, and mass spectroscopic analysis of dried blood spots is performed.

Tertiary screen. The tertiary screening stage is designed to utilize existing or potential models of neurological disease, as well as the assessment of phenotypic variability that may be the result of known or uncharacterized genetic influences. Anxiety is measured by open-field activity and the elevated plus maze, and learning and memory are assessed with the Morris water maze. Where appropriate, electromyography, electroencephalography (EEG), nerve conduction, and magnetic resonance imaging (MRI) techniques are also employed.

References

Bach ME, Hawkins RD, Osman M, Kandel ER, Mayford M (1995) Impairment of spatial but not contextual memory in CaMKII mutant mice with a selective loss of hippocampal LTP in the range of the θ frequency. Cell 81, 905–915

Braff D, Stone D, Callaway E, Geyer MA, Glick I, Bali L (1978) Prestimulus effects on human startle reflex in normals and schizophrenics. Psychophysiology 15, 339–343

Chang PL, Lambert DT, Pisa MA (1993) Behavioural abnormalities in a murine model of human lysosomal storage disease. Neuroreport 4, 507–510

Crawley JN (1981) Neuropharmacologic specificity of a simple animal model of the behavioral actions of benzodiazepines. Pharmacol Biochem Behav 15, 695–699

Gerlai R, Millen KJ, Herrup K, Fabien K, Roder J (1996) Impaired motor learning performance in cerebellar En-2 mutant mice. Behav Neurosci 110, 126–133

Gower AJ, Laberty Y (1993) The aged mouse as a model of cognitive decline with special emphasis on studies in NMRI mice. Behav Brain Res 57, 163–173

Irwin S (1968) Comprehensive observational assessment: 1a. A systematic, quantitative procedure for assessing the behavioural and physiologic state of the mouse. Psychopharmacologia 13, 222–257

Jones BJ, Roberts DJ (1968) The quantitative measurement of motor incoordination in naive mice using an accelerating Rota-Rod. J Pharm Pharmacol 20, 302–304

Lister RG (1987) The use of a plus maze to measure anxiety in the mouse. Psychopharmacology 92, 180–185

Morris RGM (1981) Spatial localisation does not require the presence of local cues. Learn Motiv 12, 239–260

Muller U, Cristina N, Li ZW, Wolfer DP, Lipp HP, Rulicke T, Brandner S, Aguzzi A, Weissmann C (1994) Behavioral and anatomical deficits in mice homozygous for a modified beta-amyloid precursor protein gene. Cell 79, 755–765

Nasir J, Floresco SB, O'Kusky JR, Diewert VM, Richman JM, Zeisler J, Borowski A, Marth JD, Phillips AG, Hayden MR (1995) Targeted disruption of the Huntington's disease gene results in embryonic lethality and behavioral and morphological changes in heterozygotes. Cell 81, 811–823

O'Callaghan JP, Holtzman SG (1975) Quantification of analgesic activity of narcotic antagonists by a modified hot-plate procedure. J Pharmacol Exp Ther 192, 497–505

Paylor R, Crawley JN (1997) Inbred strain differences in prepulse inhibition of the mouse startle response. Psychopharmacology, in press

Paylor R, Tracey R, Wehner J, Rudy JW (1994) DBA/2 and C57BL/6 mice differ in contextual fear but not auditory fear conditioning. Behav Neurosci 108, 810–817

Reeves RJ, Irving NG, Moran TH, Wohn A, Kitt C, Sisodia SS, Schmidt C, Bronson RT, Davisson MT (1995) A mouse model for Down syndrome exhibits learning and behaviour deficits. Nature Genet 11, 177–183

Tonegawa S, Li Y, Erzurumlu RS, Jhaveri S, Chen C, Goda Y, Raylor R, Silva AJ, Kim JJ, Wehner JM, Stevens CF, Abeliovich A (1995) The gene knockout technology for the analysis of learning and memory, and neuronal development. Progr Brain Res 105, 3–14

Willott JF, Carlson S (1995) Modification of the acoustic startle response in hearing-impaired C57BL/6J mice: prepulse augmentation and prolongation of prepulse inhibition. Behav Neurosci 109, 396–403