The Old School, Brewhouse Hill, Wheathampstead,
Hertfordshire AL4 8AN, UK

Animal Welfare 2006, 15: xxx-xxx
ISSN 0962-7286

Reducing the number of animals used in behavioural genetic experiments using chromosome substitution strains

MC Laarakker*, F Ohl and HA van Lith

Department of Animals, Science & Society, Division of Laboratory Animal Science, Faculty of Veterinary Medicine, Utrecht University, PO Box 80.166, 3508 TD Utrecht, The Netherlands
* Contact for correspondence and requests for reprints: M.C.Laarakker@vet.uu.nl

Abstract

Chromosome substitution strains (also called consomic lines or strains) are strains in which a single, full-length chromosome from one inbred strain — the donor strain — has been transferred onto the genetic background of a second inbred strain — the host strain. Based on the results obtained from behavioural tests with the two parental strains, the minimum number of animals from each of the host and consomic strains that are required for a successful behavioural genetic analysis can be estimated. Correct application of statistical knowledge can lead to a further reduction in the number of animals used in behavioural genetic experiments using chromosome substitution strains.

Keywords: *animal welfare, behavioural genetics, chromosome substitution strain, mice, QTL, reduction*

Introduction

A quantitative trait locus (QTL) is a position on the genome that is associated with genetic differences for a quantitative trait. Over the past decade, methods for genome analysis of animal models have been developed to identify and locate QTLs. Chromosome substitution strains, also called consomic strains, can accelerate the identification and mapping of QTLs. Chromosome substitution strains are produced by transferring a single, full-length chromosome from one inbred strain — the donor strain — onto the genetic background of a second strain — the host strain — by repeated backcrossing (Singer *et al* 2004). Because the host and donor strain are genetically very diverse, the consomic panels (a set of chromosome substitution strains) can be used as a general genetic discovery tool. Therefore, panels of chromosome substitution strains are an advantage to researchers studying the genes affecting developmental, physiological and behavioural processes.

The Division of Laboratory Animal Science, Utrecht University, is specifically interested in behavioural genetic research and plans to use a commercially available set of mouse chromosome substitution strains (van Lith 2005). Determination of the number of animals required per strain of both the host strain and the consomic strain for the genetic analysis is one of the most important and difficult decisions one has to make. Based on the results obtained from behavioural tests with the two parental strains, the minimum number of animals from the host and consomic strains that are required for a successful behavioural genetic analysis can be estimated. This paper demonstrates that, by using statistical knowledge in a correct way, it is possible to reduce the number of laboratory animals used in behavioural genetic experiments using chromosome substitution strains.

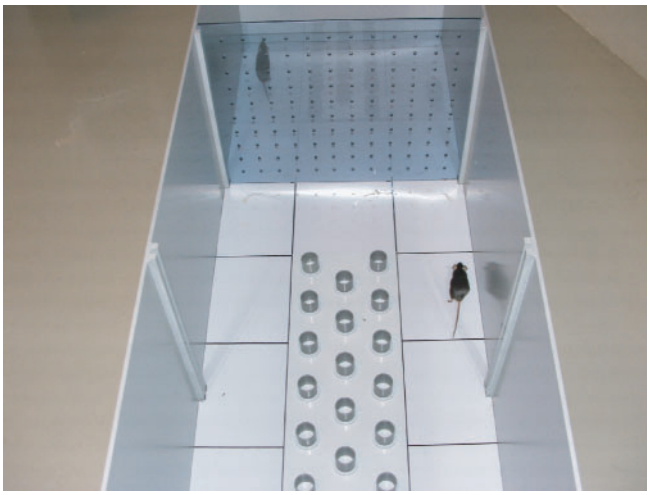
Materials and methods

The protocol of the experiment was approved by the Animal Experiments Committee of the Academic Biomedical Centre, Utrecht, and peer-reviewed by the scientific and ethical committee of the Department of Animals, Science & Society, Utrecht University.

Animals

This study was performed using naive male mice from two commercially available mouse inbred strains: A/J, the donor strain, and C57BL/6J, the host strain (n = 9 per strain; The Jackson Laboratory, Bar Harbor, USA). The mice were 4–6 weeks old at arrival, and were housed in a room of the laboratory animal facility at the Department of Animals, Science & Society (Utrecht University) for two weeks for habituation before the behavioural testing started; testing was carried out in the same room. The animals were socially housed in Macrolon–2 cages (three mice per cage) and maintained under a reversed light:dark cycle (white light: 1900h–0700h; red light: 0700h–1900h) with food and water available *ad libitum*. In each cage, animals were provided with a shelter, tissues (Kleenex®: Kimberly–Clark Professional BV, Ede, The Netherlands) and paper shreds (EnviroDri®: Tecnilab–BMI BV, Someren, The Netherlands) as environmental enrichment. Humidity was kept at a constant level of 50% and the ambient temperature was maintained at 21.0 ± 2.0ºC. During the habituation period all mice were handled at

Figure 1



The modified hole–board test. The experimental animal was placed in the testing compartment, with the hole–board in the centre. For socially housed mice, group mates were placed in the group compartment behind the partition to prevent isolation stress.

least four times per week for a few minutes by the person who performed the behavioural experiments (MC Laarakker), this included picking up the animal at the tail base and placing it on the hand or arm, and restraining it by hand for a few seconds at random times of the day.

The modified hole–board test

The behavioural tests were performed using a modified hole–board (mHB) test (Figure 1; Ohl 2003). The mHB test combines the features of an open–field and a hole–board test. In summary, the mHB consists of an opaque grey polyvinylchloride (PVC) box (150 × 50 × 50 cm, length × width × height) that is divided into two compartments by a transparent partition perforated with 120 holes each 3 mm in diameter. The smaller compartment (50 × 50 × 50 cm, length × width × height) is used as a group compartment and is enriched with a tissue (Kleenex®: Kimberly–Clark Professional BV, Ede, The Netherlands); all the cage mates of the tested animal are placed in this group compartment, allowing the tested animal both olfactory and visible contact in order to prevent isolation stress. The experimental compartment (100 × 50 × 50 cm, length × width × height) consists of two areas, one protected area — the box — which is surrounded by the protective walls of the set-up, and an unprotected area — the board. Black lines divide the box into 10 rectangles (20 × 15 cm, length × width) and 2 squares (20 × 20 cm, length × width). The board (60 × 20 × 0.5 cm, length × width × height) is placed in the centre of the experimental compartment, and contains 23 PVC cylinders (3 × 3 cm, diameter × height), positioned across the board in three rows. The board is lit with an additional red light lamp (80 W), such that the board is illuminated with approximately 35 lux, whereas the box is only illuminated with 1–3 lux.

Behavioural testing was performed between 1000h and 1400h under red light conditions; all behavioural tests were videotaped from above the experimental compartment and from the side through the partition wall. At the start of the behavioural test, all three mice from one cage were placed in the group compartment of the mHB for 10 min in order to allow for habituation. The mice were subsequently placed in the experimental compartment one at a time and allowed free exploration for 5 min while a trained observer (MC Laarakker) scored the behaviour by hand, using the program Observer 4.1 (Noldus, Wageningen, The Netherlands). The test set-up was cleaned with water and a damp towel between each mouse.

As in previous studies (Ohl *et al* 2001; references cited in Ohl 2003), several parameters for anxiety–related behaviour (eg latency until the first board entry and number of board entries), locomotion, exploration, risk assessment, memory, arousal, immobility and social affinity were measured. However, only the results for the two parameters 'latency until the first board entry' and 'number of board entries' have been presented here.

Statistical analyses

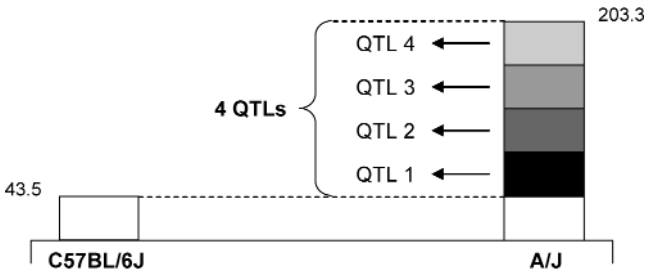
All statistical analyses were carried out using the SPSS computer program. Two-tailed probabilities were estimated throughout. Continuous data (parameter 'latency until the first board entry') were summarised as means ± standard error of the mean (SEM), whereas discrete data on the ordinal scale (parameter 'number of board entries') were presented as medians with the interquartile range (IQR). The Kolmogorov-Smirnov one-sample test was used to check normality of the continuous data; all continuous results within the A/J and C57BL/6J strains were normally distributed. Significant differences in the continuous data between A/J and C57BL/6J mice were calculated using the unpaired Student's *t*-test. The unpaired Student's *t*-tests were performed using pooled (for equal variances) or separate (for unequal variances) variance estimates. The equality of variances was tested using an *F* test. For the unpaired Student's *t*-test with separate variance estimates, SPSS uses the Welch correction. The significance of differences for the ordinal data between A/J and C57BL/6J mice was calculated using the Mann-Whitney *U* test.

Results

Parental strains

Nine mice from the strains A/J and C57BL/6J were used as donor and host strains respectively. Based on former mouse strain comparisons (Ohl *et al* 2001) two discriminating parameters were selected: (1) latency until the first board entry and (2) number of board entries. Significant differences between the two parental strains were found for both parameters (see below). The obtained strain differences prompted the investigation into the chromosomal location of the QTL involved by testing a set of chromosome substitution strains between the A/J and C57BL/6J strains.

Figure 2



Calculation of the mean value for a consomic strain containing one QTL using the behavioural data obtained for the parameter 'latency until the first board entry'.

Latency until first board entry

The parameter 'latency until the first board entry' was considered to be a continuous variable. The mean ± standard error of the mean [SEM] for this parameter were: C57BL/6J = 43.5 ± 13.7 s (n = 9, all male); A/J = 203.3 ± 38.7 s (n = 9, all male). Using a two-tailed unpaired Student's *t*-test with Welch correction, there was a significant difference between the two strains and the parameter 'latency until the first board entry', $t = 3.89$, df = 9.97, $P = 0.003$.

Using two standard behavioural tests, the open–field test and the dark–light transmission test, Singer *et al* (2004) found that there are probably 3–4 non-linked (ie located on different chromosomes) QTLs present per parameter. If the effects of these QTLs are assumed to be additive and that every QTL has the same magnitude of effect, then this means that a chromosome substitution strain with n = 9 animals (all male) can have a mean of 83.5 ± 20.0 s, assuming that the four QTLs are non-linked. The mean value for a consomic strain containing one QTL was calculated using the following equation (and see Figure 2):

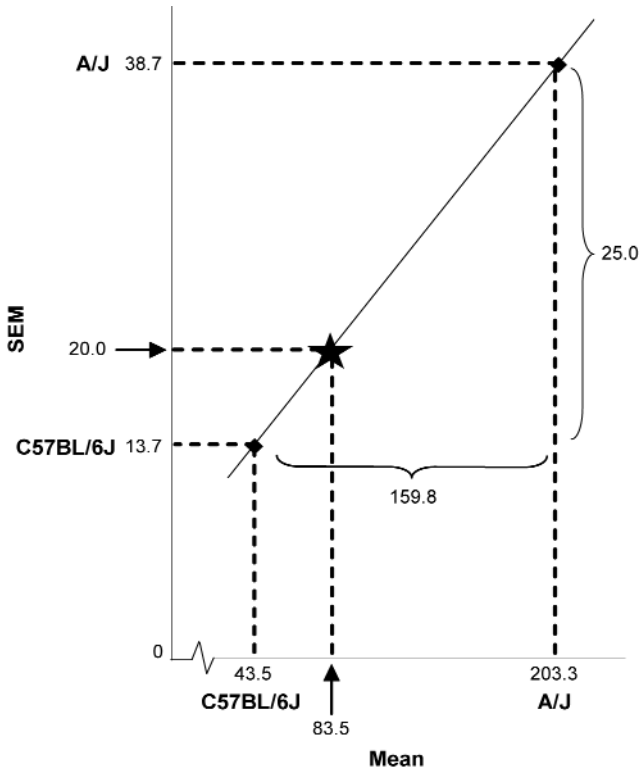
$$43.5 + \left(\frac{203.3 - 43.5}{4} \right) \approx 83.5$$

The standard error of the mean (based on n = 9) for a consomic strain with one QTL (linear interpolation; Figure 3) was calculated as follows:

$$13.7 + \left(\left[\frac{38.7 - 13.7}{203.3 - 43.5} \right] \times (83.5 - 43.5) \right) \approx 20.0$$

In order to compare the chromosome substitution strain with its host strain, α has to be adjusted because of a greater probability of a Type 1 error attributable to multiple comparisons. Belknap (2003) proposes an α value of between 0.003 and 0.004. A chromosome substitution strain cannot be expected to differ significantly from the host strain (C57BL/6J) with respect to the parameter 'latency until the first board entry' when n = 9 animals per strain are used (two-tailed Student's *t*-test, $t = 1.65$, df = 16, $P = 0.119$). Therefore, in order to obtain a significant result, the number of animals (n) was

Figure 3



Calculation of the standard error of the mean for a consomic strain with one QTL using the behavioural data obtained for the parameter 'latency until the first board entry'.

increased. If n is doubled, the results for the host strain and the consomic strain (eg for chromosome 1) become: C57BL/6J = 43.5 ± 9.4 s (n = 18, all male); C57BL/6J–Chr1^{A}/NaJ (the consomic strain = 83.5 ± 13.7 s (n = 18, all male). Using an unpaired two-tailed Student's *t*-test with Welch correction, $P = 0.022$ ($t = 2.40$, df = 34), which, according to Belknap (2003), was not significant. The results were statistically significant when n = 27 animals per strain were used, when $P = 0.004$ ($t = 2.97$, df = 52). The predicted results were: C57BL/6J = 43.5 ± 7.6 s (n = 27, all male); C57BL/6J–Chr1^{A}/NaJ = 83.5 ± 11.1 s (n = 27, all male).

This suggested that behavioural testing should start with 27 mice of the C57BL/6J host strain and 6 mice per chromosome substitution strain, that is, to start with a ratio of 4.5:1 (27:6), as suggested by Belknap (2003). This experiment predicted the following results: C57BL/6J = 43.5 ± 7.6 s (n = 27, all male); C57BL/6J–Chr1^{A}/NaJ = 83.5 ± 25.3 s (n = 6, all male); $P = 0.053$ ($t = 2.02$, df = 31).

With a value of $P < 0.053$, then for the chromosome substitution strain for which this is the case (assuming 4 consomic strains are found, and using the parameter 'latency until the first board entry'), 21 additional male animals would be tested. After behavioural testing, the statistical analysis would be repeated, but with 27 animals for both the chromosome substitution strain and the host strain (C57BL/6J), it is most likely that a value close to $P = 0.004$ will be obtained.

Table 1 The number of board entries made by A/J and C57BL/6J male mice.

	A/J (n = 9)	C57BL/6J (n = 9)
	0	1
	0	6
	0	10
	0	12
	0	14
	4	15
	5	16
	6	17
	13	20
median	0	14
IQR	5.5	8.5

Table 2 The number of board entries made by consomic mice (C57BL/6J–Chr1^{A}/NaJ) and C57BL/6J male mice.

	C57BL/6J–Chr1^{A}/NaJ (n = 9)	C57BL/6J (n = 9)
	0	1
	2	6
	3	10
	3	12
	4	14
	4	15
	6	16
	12	17
	15	20
median	4	14
IQR	6.5	8.5

Board entries

The parameter 'board entries' was considered to be a discrete variable on the ordinal scale. See Table 1 for the the number of board entries made by the A/J and C57BL/6J mice, and the median with the IQR. For this parameter the two inbred strains were significantly different (two-tailed Mann-Whitney *U* test, $U = 7.5$, $P = 0.0032$). Assuming there were also four QTLs with equal affects involved for this parameter then the results for the chromosome substitution strain and the host strain could be calculated (see Table 2).

Estimating the median for a consomic strain with one QTL was calculated as follows:

$$\left(\frac{14-0}{4}\right) = 3.5$$

Because an odd number of animals were used, the median was either 3 or 4. Considering the most unfavourable situation (which means the difference between host strain and consomic strain is smaller) a median of 4 was chosen. The IQR was estimated as follows:

$$5.5 + \left(4\left[\frac{8.5-5.5}{14}\right]\right) \approx 6.36$$

Because the numbers of board entries were counts and nine animals were used, the IQR became 6.25 or 6.5; therefore, similar to 6.36. Again, the most unfavourable situation was considered and the largest IQR was chosen, ie 6.5.

Although it is not correct for discrete data, the mean and SEM for the parameter 'number of 'board entries' were calculated to guarantee comparability of the shape of the distribution between the A/J, consomic and C57BL/6J strain: for the A/J strain, mean ± SEM = 3.1 ± 1.5 and for the C57BL/6J 12.3 ± 2.0; a chromosome substitution strain could be 5.4 ± 1.6.

The mean for a consomic strain containing one QTL (Figure 4) was calculated as follows:

$$3.1 + \left(\frac{12.3-3.1}{4}\right) = 5.4$$

The SEM for a consomic strain containing one QTL (linear interpolation; Figure 5) was calculated as follows:

$$1.5 + \left(\left[\frac{2.0-1.5}{12.3-3.1}\right](5.4-3.1)\right) \approx 1.6$$

It appears that with the values presented in Table 2 for the C57BL/6J–Chr1^{A}/NaJ (n = 9) consomic strain, the preferred results were obtained; Mann-Whitney *U* test, $U = 15.5$, $P = 0.0269$. Using 9 consomic animals and 27 host strain animals (median = 14; IQR = 6.0), $P = 0.0060$. If 27 consomic animals (median = 14, IQR = 3.0) and 27 host strain animals (median = 14, IQR = 6.0) were tested (see section *Latency until the first board entry*), $P = 0.0001$ ($U = 139.5$). For this parameter (data were analysed using the Mann-Whitney *U* test), α was also adjusted because of the increased probability of a Type 1 error. Here as well, the criteria of Belknap (2003) can be used ($P < 0.004$). Applying these rules, and using 27 consomic and 27 host strain animals for this parameter would satisfy the criteria. The criteria would not be satisfied using 9 consomic and 27 host strain animals ($U = 46.5$, $P = 0.006$,); therefore, our suggestion is to use 12 consomic mice for this parameter.

Discussion

When more than one statistical test is performed when analysing data from a study, statisticians demand that a more stringent criterion be used for statistical significance than the conventional $P < 0.05$. As previously explained in the section *Latency until the first board entry*, multiple comparisons (comparing consomic strains with the host strain) were carried out, therefore there was an increased probability of a Type 1 error. A method to adjust for multiple tests is known as the Bonferroni adjustment: α is divided by the number of comparisons that are carried out, and, for example, Singer *et al* 2004 and Singer *et al* 2005 use this method. A complete set of chromosome substitution strains of the mouse contains of 21 different strains

Table 3 Overview of the number of laboratory animals for the parameter 'latency until the first board entry' using different statistical methods.

$\alpha = 0.0024$	$\alpha = 0.004$	$\alpha = 0.004$ + multiple phase approach
Singer *et al* 2005	Belknap 2003	This article
726	594	237
33 × 22	27 × 22	27 + 6 × 21 + 21 × 4

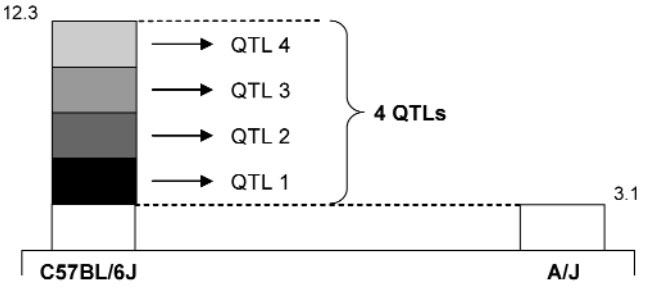
(19 autosomes, the X and Y chromosome). Because we plan to compare all 21 available consomic strains with the host strain, the adapted α of Singer *et al*'s (2004; 2005) method would give a value of ≈ 0.0024.

Belknap's (2003) method is actually based on the method developed by Dunnett (1955) and applies to the situation where different test groups (in this case the consomic strains) are compared to only one control group (in this case the host strain). For example, the continuous parameter 'latency until the first board entry', an α of 0.004 (suggested by Belknap 2003) when compared with $\alpha = 0.0024$ (suggested by Singer *et al* 2004; Singer *et al* 2005) implies a reduction in the total number of animals that have to be tested (726 versus 594 animals). A further reduction in the number of laboratory animals used in such an experiment would be possible if the behavioural tests were started with 27 C57BL/6J host strain animals and 6 animals per consomic strain (according to Belknap [2003] a 4.5:1, or 27:6, ratio is the most efficient) and extra animals (21 in case of the parameter 'latency until the first board entry' and 6 in case of the parameter 'board entry') of the appropriate consomic strains were only tested if $P < 0.05$. Table 3 shows an overview of the number of animals used for testing the parameter 'latency until the first board entry'. The reduction in the number of laboratory animals can be derived from this.

The two parameters 'latency until the first board entry' and 'number of board entries' are related to each other (Spearman's coefficient of rank correlation, $R_s = -0.9215$, n = 18, $P = 0.000$); therefore, a multivariate method, such as Hotelling's T^2 test, could have been used. Furthermore, Turri *et al* (2004) demonstrated that multivariate analysis, when compared with univariate analysis, has an increased power to detect QTLs when the genetic effects are correlated. However, several assumptions are necessary for proper application of the Hotelling's T^2 test. One of the assumptions is that dependent variables should have a multivariate normal distribution. Because 'latency until the first board entry' is a continuous variable and 'number of board entries' is a discrete variable, the joint distribution can never be bivariate normal; therefore we have not attempted to analyse these two parameters jointly.

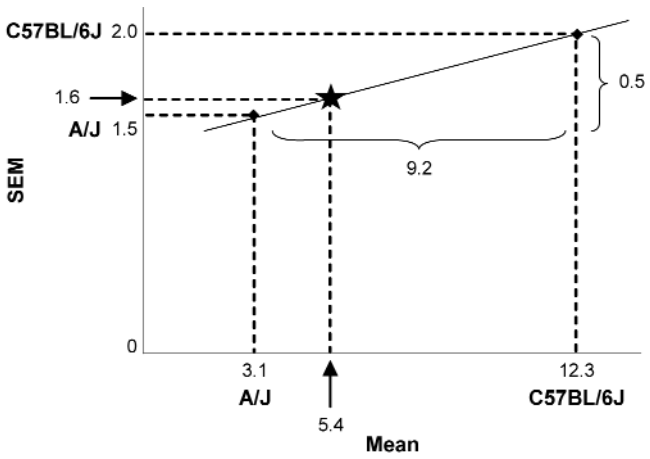
Based on the results obtained from the behavioural tests using the two parental strains, the minimum number of animals from the host and consomic strains that are needed for a successful behavioural genetic analysis can be estimated. By adjusting α to 0.004 (Belknap 2003), and performing a multiple phase approach (our suggestion), a reduction in the number of animals used in these experiments can be obtained. The multiple phase approach that we suggest is in fact a form of sequential analysis, which Russell and Burch (1959, reprinted 1992) had already suggested in 1959 as one method of reducing the number of animals used in experiments.

Figure 4



Calculation of the mean value for a consomic strain containing one QTL using the behavioural data from the parameter 'number of board entries'.

Figure 5



Calculation of the standard error of the mean for a consomic strain with one QTL using the behavioural data obtained for the parameter 'number of board entries'.

Acknowledgements

The authors would like to thank two anonymous referees for their useful comments on the manuscript.

References

Belknap JK 2003 Chromosome substitution strains: some quantitative considerations for genome scans and fine mapping. *Mammalian Genome 14*: 723-732

Dunnett CW 1955 A multiple comparison procedure for comparing several treatments with a control. *Journal of the American Statistical Association 50*: 1096-1121

Ohl F 2003 Testing for anxiety. *Clinical Neuroscience Research 3*: 233-238

Ohl F, Sillaber I, Binder E, Keck ME and Holsboer F 2001 Differential analysis of behavior and diazepam-induced alterations in C57BL/6N and BALB/c mice using the modified hole board test. *Journal of Psychiatric Research 35*: 147-154

Russell WMS and Burch RL 1959 (reprinted 1992) *The Principles of Humane Experimental Technique*. Universities Federation for Animal Welfare: Wheathampstead, UK

Singer JB, Hill AE, Burrage LC, Olszens KR, Song J, Justice M, O'Brien WE, Conti DV, Witte JS, Lander ES and Nadeau JH 2004 Genetic dissection of complex traits with chromosome substitution strains of mice. *Science 304*: 445-448

Singer JB, Hill AE, Nadeau JH and Lander ES 2005 Mapping quantitative trait loci for anxiety in chromosome substitution strains of mice. *Genetics 169*: 855-862

Turri MG, DeFries JC, Henderson ND and Flint J 2004 Multivariate analysis of quantitative trait loci influencing variation in anxiety-related behaviour in laboratory mice. *Mammalian Genome 15*: 69-76

van Lith HA 2005 Mapping of quantitative trait loci underlying strain differences in treatment susceptibility in rodents. http://www.vet.uu.nl/site/viavet_english/viavet_skeleton/Framework/main_frame_dtml?path=viavet_english/faculty/departments&skeleton_url=http://www.vet.uu.nl/site/viavet_english/viavet_skeleton (accessed 2 November 2005)