

## Methods

**Mice.** All animals used in *Listeria* infection experiments were 6 to 9 week old, housed in a BL2 containment facility. Time to death experiments were performed with the full knowledge and approval of the Standing Committee on Animals at Harvard Medical School (Protocol # 03032). BALB/cByJ, C57BL6/ByJ inbred and CXB recombinant inbred strains of mice were obtained from Jackson Labs. Initial 120 BALB/cByJxC57BL6/ByJ female F2 progeny were bred at Jackson Lab. The second cross was bred at the HMS facility.

***Listeria monocytogenes* infections.** The dose of bacteria allowing reliable differentiation between resistant and susceptible animals was determined in a dose response experiment. Intravenous infection of BALB/cByJ animals by tail vein injection of  $2.5 \times 10^4$  cfu of *Listeria monocytogenes* 10403s (ref. 1) in 300  $\mu$ l of PBS led to death of BALB/cByJ mice at around the 72 hour timepoint, while all animals of the C57BL6/ByJ background were able to recover. Our observation schedule defined an 8-hour window in determination of the death timepoint.

**Phenotyping of the intercross animals.** Animals from the first cross were infected in two groups. The first group of 90 CB6F2/ByJ females and 4 of each of the parental controls were infected at 9 weeks of age by intravenous injection of 300  $\mu$ l of approx.  $3 \times 10^4$  *Listeria monocytogenes* in PBS. For technical reasons, such as known under- or over-dosing of the animals, 4 animals were excluded from the analysis. A second group of 30 CB6F2/ByJ 9 weeks old females and 8 parental controls was infected with the same dose of bacteria. The second cross consisting of 84 8-9 weeks old mixed sex CB6F2/ByJ animals was independently analyzed using an identical experimental setup. The susceptibility of mice to *L. monocytogenes* infection was defined as the number of hours elapsed between injection and death of the animal. All moribund animals were recorded as dead and sacrificed by asphyxiation. All death timepoints were recorded as time of observation values, except for the sacrificed moribund animals for which the values represented the exact time of death. Accordingly, all time of death values are equal to time of observation minus 0-8 hours. Animals surviving past the 240 hour timepoint were considered recovered from the disease. Spleens and additional tail samples were collected postmortem from all animals used in the experiment.

**Histology.** Livers and spleens were removed aseptically from sacrificed infected animals and fixed in 10% buffered formalin (Biochemical Sciences). The organs were embedded in paraffin, cut into 5  $\mu$  sections and stained with hematoxylin and eosin. Intracellular bacteria were visualized by Gram staining of representative sections.

**Genotyping.** Genomic DNA was extracted from all of the CB6F2/ByJ animals using commercial QIAamp tissue kit (QIAGEN). MapPairs primer sets (Research Genetics) for genetic mapping were selected based on PCR product allelic difference of at least 8 bp to allow genotyping using 4% agarose gels. The initial genetic map of CB6F2/ByJ animals covered all autosomes with spacing of approximately 20 cM. Additional markers were added to regions around loci showing linkage in preliminary analysis.

**Data analysis.** For mouse  $i$ , let  $g_i$  denote its genotype at a putative QTL, let  $z_i = 1$  or 0 according to whether or not it survived, and let  $y_i$  denote the log of its time-to-death (left undefined in the case  $z_i = 1$ ). We assume that the  $(g_i, y_i, z_i)$  are mutually independent, that  $\Pr(z_i = 1 | g_i = g) = p_g$ , and that  $y_i | (g_i = g) \sim \text{normal}(\mu_g, \sigma^2)$ . In the case of complete genotype data, maximum likelihood estimates (MLEs) of the parameters  $p_g$ ,  $\mu_g$ , and  $\sigma^2$  are easily obtained. In the case of missing genotype data (especially in the consideration of positions between markers), we calculated  $q_{ig} = \Pr(g_i = g | \text{marker data})$ , under the assumption of no crossover interference, using the hidden Markov model (HMM) technology developed by Baum *et al.*<sup>2</sup> and, following the approach of Lander and Botstein<sup>3</sup>, we used a form of the EM algorithm to obtain MLEs<sup>4</sup>. We calculated the MLEs under four hypotheses:  $H_0$ : all  $p_g$  and all  $\mu_g$  are equal;  $H_1$ : the  $p_g$  are equal but the  $\mu_g$  vary;  $H_2$ : the  $\mu_g$  are equal but the  $p_g$  vary; and  $H_3$ : the  $p_g$  and  $\mu_g$  each vary. Let  $L_i$  be the maximum log (base 10) likelihood under the hypothesis  $H_i$ . Then  $\text{lod}(p, \mu) = L_3 - L_0$ ,  $\text{lod}(p) = L_3 - L_1$ , and  $\text{lod}(\mu) = L_3 - L_2$ . In the case of complete genotype data (but not when there is missing data), the likelihoods for the  $p_g$  and  $\mu_g$  may be separated, and thus  $\text{lod}(p, \mu) = \text{lod}(p) + \text{lod}(\mu)$ . We performed permutation tests<sup>5</sup>, with 1000 permutations, to obtain approximate significance thresholds for the lod scores. Complete details of the statistical methods are provided at <http://biosun01.biostat.jhsph.edu/~kbroman/publications/ms0004.pdf>.

We used a bootstrap analysis<sup>6,7</sup> to obtain 95% confidence intervals for QTL locations.

1. Portnoy, D.A., Jacks, P.S. & Hinrichs, D.J. Role of hemolysin for the intracellular growth of *Listeria monocytogenes*. *J. Exp. Med.* **167**, 1459-1471 (1988).
2. Baum, L.E., Petrie, T., Soules, G. & Weiss, N. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Stat.* **41**, 164-171 (1970).
3. Lander, E.S. & Botstein, D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**, 185-199 (1989).
4. Dempster, A.P., Laird, N.M. & Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Stat. Soc. B.* **39**, 1-38 (1977).
5. Churchill, G.A. & Doerge, R.W. Empirical threshold values for quantitative trait mapping. *Genetics* **138**, 963-971 (1994).
6. Visscher, P.M., Thompson, R. & Haley, C.S. Confidence intervals in QTL mapping by bootstrapping. *Genetics* **143**, 1013-1020 (1996).
7. Manly, B.F.J. *Randomization, bootstrap and Monte Carlo methods in biology*. (Chapman & Hall, London, 1997).