Incorporating Mouse Genome
Mammalian Genome
Genes and Phenotypes

A major locus conferring susceptibility to infection by *Streptococcus pneumoniae* in mice

Paul Denny,[1] Elaine Hopes,[1⋆] Neill Gingles,[2#⋆] Karl W. Broman,[3] William McPheat,[5] John Morten,[4] Janet Alexander,[2¶] Peter W. Andrew,[2] Steve D.M. Brown[1]

[1]MRC UK Mouse Genome Centre & Mammalian Genetics Unit, Harwell, Oxon, OX11 0RD, UK
[2]Department of Microbiology and Immunology, University of Leicester, Leicester, UK
[3]Department of Biostatistics, Johns Hopkins University, Baltimore, Maryland, USA
[4]Research and Development Genetics, AstraZeneca, Alderley Park, Macclesfield, Cheshire, UK
[5]Cell Biology and Biochemistry, AstraZeneca R&D Mölndal, Pepparedsleden 1, 431 83 Mölndal, Sweden

Received: 18 December 2002 / Accepted: 14 March 2003

Abstract

We have studied the genetics of susceptibility to infection by *Streptococcus pneumoniae* in mice. Linkage analysis of the F_2 generation from a cross between resistant BALB/cO1aHsd and susceptible CBA/CaO1aHsd strains allowed us to map a major locus controlling the development of bacteremia and survival after intranasal infection.

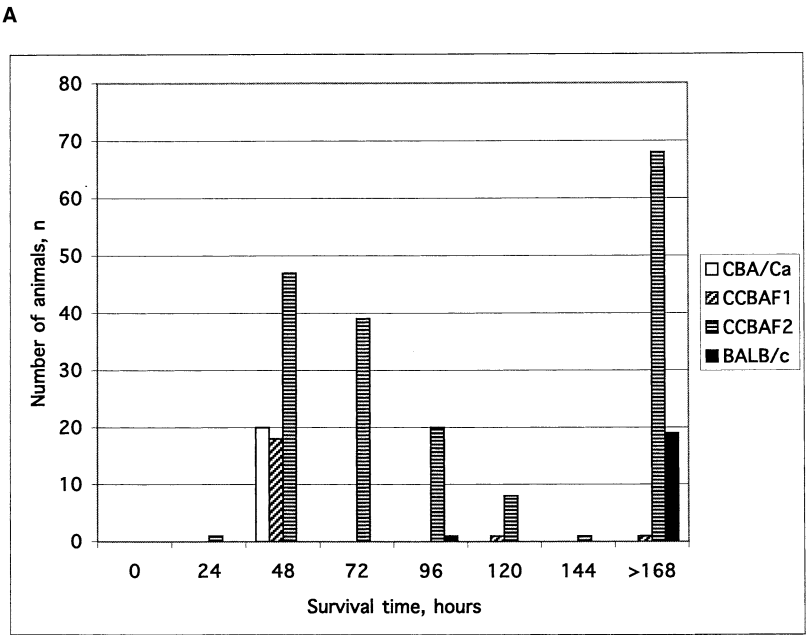
Introduction

Infections caused by the pneumococcus, *Streptococcus pneumoniae*, have accounted historically for more morbidity and mortality than any other bacterium. The pneumococcus is responsible for the majority of cases of community-acquired pneumonia and is also a significant cause of meningitis, bacteremia, and otitis media in children (Mufson 1981). Antibiotics are the standard therapy for pneumococcal disease, but antibiotic resistance is increasingly common in pneumococcal strains and has epidemic potential (Nuorti et al. 1998). Vaccination can offer protection against a spectrum of pneumococcal serotypes, but there are various problems with current vaccination strategies, and horizontal gene transfer may further reduce the effectiveness of both vaccines and antibiotics (Tuomanen 1999). These factors make it clear that a better understanding of the host response to this bacterium is essential for improved preventative and therapeutic treatments.

There is a highly significant genetic contribution to susceptibility to infectious disease in humans (Sorensen et al. 1988; Cooke & Hill 2001). Allelic associations of candidate genes with pneumococcal disease have been detected in case-control studies (Roy et al. 2002a, 2002b), but the limited number of case-confirmed pedigrees or sibling pairs preclude genome-wide linkage analysis. In order to circumvent this problem and allow us to identify host factors controlling susceptibility to pneumococcal infection, we chose to use the laboratory mouse as a model genetic organism. The mouse is the most frequently used animal model for pneumococcal disease. In susceptible strains, a very reproducible pattern of bronchopneumonia with bacteremia follows intranasal infection, and the resulting pathology is very similar to that observed in the disease in humans. One precedent for the power of mouse genetics to aid in studies of human infectious disease is the identification of the NRAMP1 gene (now known as SLC11A1). Natural allelic variants of this gene are associated with differences in susceptibility to infection by leishmania, mycobacteria and salmonella in mice (Blackwell 2001) and tuberculosis in humans (Cervino et al. 2000). We have previously shown marked differences in susceptibility to pneumococcal infection between inbred mouse strains (Gingles et al. 2001) and have exploited this to allow mapping of a major locus controlling survival after infection.

⋆An equal contribution to the work was made by these authors.
#*Present address:* Department of Cell Biology, Division of Vascular Biology, The Scripps Research Institute, La Jolla, California, USA.
¶*Present address:* AstraZeneca R&D Charnwood, Molecular Biology, Bakewell Road, Loughborough, Leicestershire, UK.
Correspondence to: P. Denny; E-mail: paul@har.mrc.ac.uk

 DOI: 10.1007/s00335-002-2261-9 • Volume 14, 448–453 (2003)



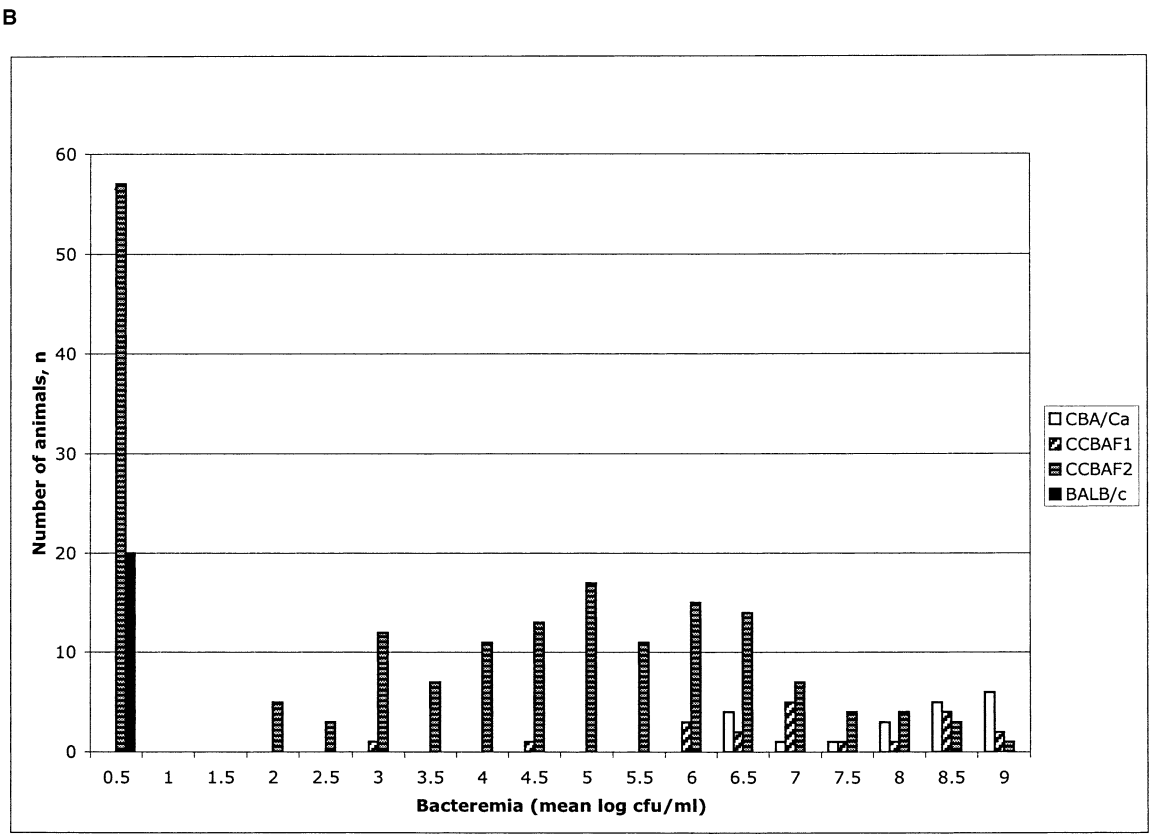


Fig. 1. Plots of survival time (in hours) after infection (**A**) and bacteremia at 24 h after infection (**B**) for animals from parental strains, F_1 and F_2 generations. Those animals in the ">168" h category lived beyond the 168-h time point and were considered to have recovered.

Materials and methods

Mice. All experiments described in this study had local ethics committee approval and were done under an appropriate Home Office license. Inbred and F_1 hybrid mice used in this study, BALB/cO1aHsd (abbreviated to BALB/c), CBA/CaO1aHsd (=CBA/Ca), and (BALB/c × CBA/Ca)F_1, were obtained from Harlan-Olac, Bicester, UK. The CCBA F_2 generation mice were bred in the Division of Biomedical Ser-

vices, University of Leicester, by using a panel of 8 male and 16 female CCBA F_1 mice. Progeny were weaned 20 days after birth and then caged separately. Eight groups generated 254 CCBA F_2 mice.

Phenotyping. Mice were obtained at 7 weeks old and infected when more than 9 weeks old. After arrival, all animals were kept for a minimum of 2 weeks to acclimatize. Following infection, mice were housed in an isolator. Mice were screened for pre-existing antibody to pneumococcal capsule by using ELISA (Gingles et al. 2001). No antibodies were detected. The type 2 *S. pneumoniae* strain used was D39 (NCTC 7466) from the National Collection of Type Cultures, Central Public Health Laboratory, London, UK. Before use in infection experiments, pneumococci were passaged through mice, and a frozen standard inoculum was prepared as described before (Gingles et al. 2001). Bacterial culture and the intranasal administration of the bacterial challenge to mice were as described (Gingles et al. 2001). Briefly, for challenge, mice were lightly anesthetized with fluothane over oxygen and challenged with 10^6 colony-forming units (CFU) of type 2 *S. pneumoniae* administered into the nostrils. Following infection, mice were closely monitored for the visual development of symptoms as described before (Morton and Griffiths 1985). Mice were monitored for 7 days, and mice alive at that time were considered to have survived the infection. The end point of the assay was moribund. The time of becoming moribund was recorded, and the animal was killed by cervical dislocation. In experiments to determine pneumococcal growth in the blood, 100 μl of blood was taken from the tail vein at 24 h post-infection, and CFU was determined by serial dilution and plating as described before (Gingles et al. 2001).

Genotyping. Genomic DNA was prepared from tail clips following challenge, using a salting out method (Miller et al. 1988). All microsatellite markers used in the initial genome scan were taken from the Whitehead/MIT database (Dietrich et al. 1996) except for tumor necrosis factor *(Tnf)* (MGD-MRK-28090, Mouse Genome Database: http://www.informatics.jax.org/). The genome scan was performed in two stages: on an initial set of 92 F_2 generation animals, and then on a further 92 animals, with 137 markers distributed across the autosomes. Markers had an average spacing of ~12 cM, with the largest interval being about 32 cM. Polymerase chain reaction (PCR) amplification was carried out by using incorporation of fluorescently labeled-dCTP from Applied Biosystems (AB), based on a high-throughput method (Rhodes et al. 1998). Genotyping marker panels were organized taking into account PCR product size, allelic variation, and the three different fluorophores. Maximum number of markers per set was 12, with most sets containing between 6 and 9 markers. A 1:200 dilution of genomic DNA was used for PCR amplification in HotStarTaq PCR buffer with 1.5 mM magnesium chloride (Qiagen), 200 μM of each unlabeled dNTP (Amersham Pharmacia), 1.3 μM of each primer (Research Genetics; Mouse MapPairs, or custom synthesized by MWG or Genset), 0.2U HotStarTaq (Qiagen), and one of either 1 μM RG6, 0.5 μM R110, or 8 μM Tamra-labeled dCTP (AB). Between 1 and 15 μl of PCR products from the marker panel for each individual, predetermined, were pooled and coprecipitated by a standard ethanol/ammonium acetate method; pellets were resuspended in 8 μl sterile deionized water. Excess labeled dCTPs were removed with mini Sephadex G50 fine columns, 1 μl of prepared products was run on a 6% wt/vol acrylamide gel, by using 12-cm well-to-read plates at 750 Volts for 2.5 h. Raw gel data were analyzed with the AB Genescan software version 3.1.2; allelic differences were analyzed with the AB software package Genotyper version 1.1, and data were stored in Microsoft Excel spreadsheets for further analysis.

Data analysis. Marker order was assessed with MAPMAKER/EXP version 3.0b (Lander et al. 1987). Linkage analysis was performed by using a nonparametric approach, similar to that described (Kruglyak and Lander 1995), but using an extension of the Kruskal-Wallis statistic (Sokal and Rohlf 1995). This statistic deserves some explanation. The Kruskal-Wallis test is a nonparametric statistical test, based on ranks, for comparing the averages of more than two samples. (In the case of an intercross, we are interested in comparing the average phenotypes in the three groups defined by the genotypes of the mice at a particular locus.)

Let n denote the number of mice and R_i denote the rank (or average rank, in the case of ties) of the phenotype of individual i. Consider a particular position in the genome, the location of a putative QTL. In most cases, the precise QTL genotype for a mouse will be unknown, but we may calculate the probability that mouse i has QTL genotype j, given multipoint marker data. Let this probability be denoted p_{ij}. Our extension of the Kruskal-Wallis statistic is the following:

$$H = \frac{12}{n(n+1)} \sum_j \frac{(n - \sum_i p_{ij})(\sum_i p_{ij})^2}{n \sum_i p_{ij}^2 - (\sum_i p_{ij})^2} \left[\frac{\sum_i p_{ij} R_i}{\sum_i p_{ij}} - \frac{n+1}{2} \right]^2$$

In the case of ties among the phenotypes, let t_k denote the number of values in the kth group of ties. In place of the statistic H (defined above), we use H/D, where $D = 1-\Sigma_k\ (t_k^3-t_k)/(n^3-n)$. In the case of complete genotype data at the putative QTL (i.e., at a genetic marker for which all mice were genotyped), the probabilities p_{ij}, are either 0 or 1, and the above statistic reduces to the usual Kruskal-Wallis statistic. We converted the statistic H/D into an approximate LOD score by dividing by 2 ln 10 = 4.61. Statistical significance was determined by permutation tests (Churchill and Doerge 1994); 10,000 permutation replicates were used.

Results

Genetic cross. Our previous work showed that BALB/cO1aHsd and CBA/CaO1aHsd are resistant and susceptible to infection by *S. pneumoniae*, respectively (Gingles et al. 2001). We therefore crossed these strains to produce (BALB/cO1aHsd × CBA/CaO1aHsd)F_1 hybrids (=CCBAF1) and then intercrossed hybrid mice to produce the (BALB/cO1aHsd × CBA/CaO1aHsd)F_2 (CCBAF2) generation, to allow us to follow the inheritance of susceptibility. We chose to follow two traits, survival time and the development of invasive pneumococcal bacteremia, as measured by the numbers of bacteria in the blood. Mice were infected by nasal instillation of pneumococci, and bacteremia was measured 24 h later (Gingles et al. 2001). The course of disease symptoms was subsequently followed for a period of 168 h after infection, and the time that animals became moribund was recorded. Animals alive at 168 h were considered to have survived the infectious challenge (see Materials and methods).

The $CCBAF_1$ generation is susceptible to infection, based on survival time and development of bacteremia, but clearly less so than the CBA/CaO1aHsd parent (Fig. 1). Male and female CCBAF1 hybrids exhibited no differences in bacteremia or survival time (data not shown). A complex phenotypic distribution was seen in the $CCBAF_2$ generation, with many animals surviving the full length of the experiment, and others succumbing to infection early on, in a way similar to the CBA/CaO1aHsd parent (Fig. 1). Many $CCBAF_2$ animals also failed to develop significant bacteremia, and there is good, but not complete, inverse correlation between bacteremia level and survival time (correlation coefficient=−0.7).

Linkage analysis. Classical quantitative trait locus (QTL) mapping methods assume that a trait follows a "normal" distribution, clearly not the case for our traits. We therefore used a non-parametric statistical analysis (see Materials and methods). In an initial genome scan performed on a group of 92 $CCBAF_2$ animals using a set of polymorphic markers (Dietrich et al. 1996; Rhodes et al. 1998), a single, suggestive (LOD ≈ 3) linkage with survival was seen on proximal Chr 7 near *D7Mit77* (data not shown). We genotyped a further 92 $CCBAF_2$ animals, and the evidence of linkage in the combined data became highly significant (LOD=6.34; $P < 0.001$) (Fig. 2A). The same region showed significant linkage with bacteremia (LOD=4.62; $P < 0.01$) (Fig. 2B). No other suggestive loci were detected (above LOD=2.3; data not shown).

It is known that there is a predisposition to pneumococcal infection in male humans (Aszkenasy et al. 1995), so we reanalyzed our data after dividing by sex. The evidence for linkage, for each phenotype, comes largely from the male mice (see Fig. 2). However, the sexes do not show a significant difference in the effect of the QTL. For example, the proportions of surviving mice by genotype at *D7Mit77* (Table 1) do not show a significant sex difference. For each sex, approximately half of the BALB/c homozygotes survive, while very few of the

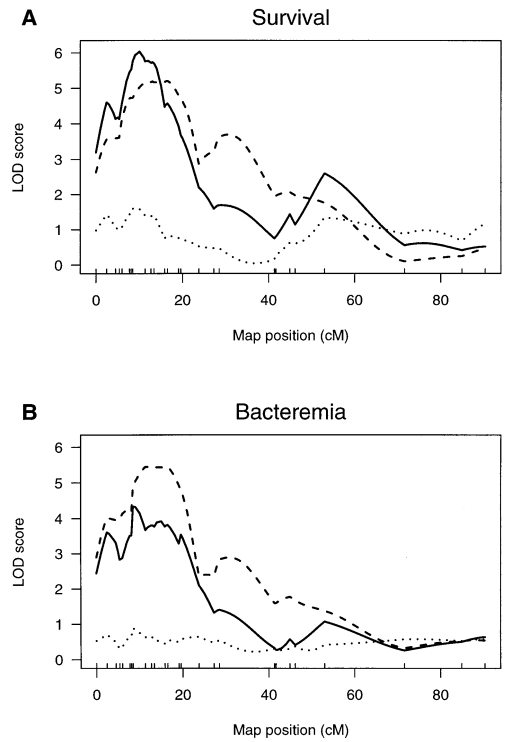


Fig. 2. LOD curves for Chr 7 for the survival (**A**) and bacteremia (**B**) phenotypes. The dashed, dotted, and solid curves correspond to the results for males, females, and the two sexes combined, respectively. Genome- wide significance of this linkage was $P = 0.0001$.

Table 1. Proportion of $CCBAF_2$ mice surviving experiment by genotype at *D7Mit77*

	Genotype			
	C/C	*C/CBA*	*CBA/CBA*	*Any genotype*
Males	13/23 (57)	20/39 (51)	0/24 (0)	33/86 (38)
Females	10/19 (53)	21/54 (39)	4/25 (16)	35/98 (36)
Overall	23/42 (55)	41/93 (44)	4/49 (8)	68/184 (37)

Numbers in parentheses are percentages.

$P = 0.09$ that there is a difference between sexes in $CCBAF_2$.

C = BALB/cO1aHsd allele; CBA=CBA/CaO1aHsd allele.

CBA/Ca homozygotes survive. Regardless of sex, the BALB/c allele at the QTL behaves as a dominant, with reduced penetrance, conferring infection resistance or survival.

Discussion

We have detected a major quantitative trait locus conferring susceptibility to infection *by S. pneumoniae* in mice. This locus influences the development of invasive disease and survival after infection. No other loci reached even a suggestive LOD threshold, but this is not sufficient to conclude that the Chr 7 locus is the only locus segregating in the cross.

The BALB/c and CBA/Ca strains are almost completely resistant and susceptible, respectively, to infection by *S. pneumoniae* (Gingles et al. 2001). Hybrid $CCBAF_1$ mice are intermediate in susceptibility between the parental strains, indicating partial dominance of the susceptible phenotype. If we consider only the genotype at *D7Mit77* in the $CCBAF_2$ generation, few of the CBA/Ca homozygotes survive to the end of the experiment (Table 1), suggesting that the CBA/Ca allele confers susceptibility (and that the BALB/c allele confers resistance).

A complex series of processes lead from inhalation of pneumococci to invasive infection and ultimately death (McCullers and Tuomanen 2001). Cellular and molecular studies of the BALB/c and CBA/Ca strains suggest that defects in the innate immune system are responsible for the differences in susceptibility to infection. Lungs of infected, susceptible mice (CBA/Ca) are deficient in neutrophils (Gingles et al. 2001) and mast cells (Kerr et al. 2001), by comparison with the resistant strain (BALB/c). Several genes encoding molecules involved in inflammation or innate immunity—for example, *Tyrobp*, *Hcst*, and *Zfp36*—lie in the genetic interval between *D7Mit341* and *D7Mit247* that is most likely to encompass the QTL (Fig. 3). However, this interval is 11 Mb and contains at least 250 genes (http://www.ensembl.org/Mus_musculus). We are, therefore, following a number of approaches in order to reduce the interval containing the QTL to a size amenable to complete sequencing in both parental strains. High-resolution mapping of interval-specific congenic strains (Darvasi 1997; Lyons et al. 2000) should allow us to eliminate many candidate genes, and microarray expression profiling (Aitman et al. 1999) may provide support for some of the remaining genes. It will, however, still be necessary to demonstrate a function for any putative mutation by using transgenic complementation or mutagenesis

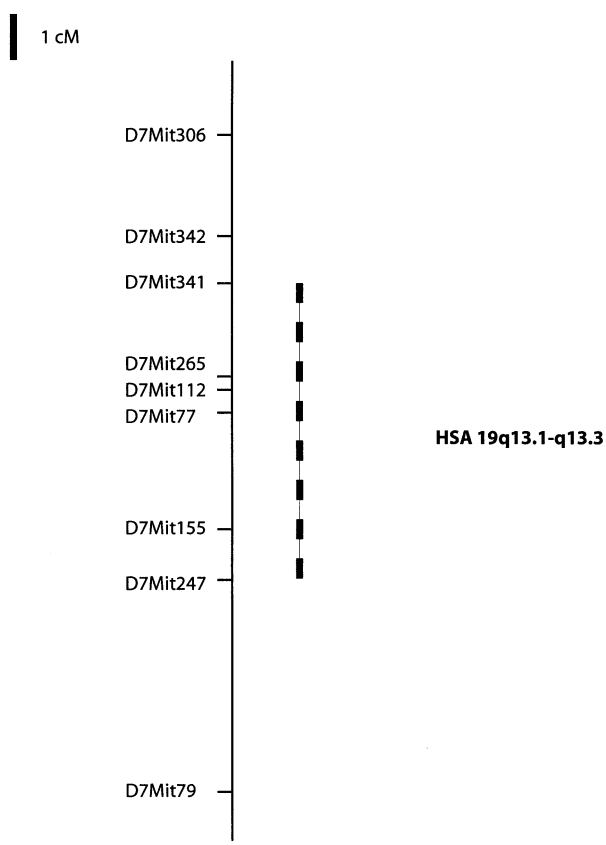


Fig. 3. Genetic map of proximal Chr 7 showing the most likely location of the survival QTL. The −1.5 LOD support interval (dashed line) spans about 7 cM, between *D7Mit341* and *D7Mit247*, with the peak LOD score close to *D7Mit77*. This region shows complete conservation of synteny with human Chr 19q13.1–q13.3.

tests (Nadeau and Frankel 2000). Identification of the etiological mutation will allow better understanding of the molecular basis of pneumococcal infection susceptibility and perhaps lead to improved approaches to prevention or therapy in humans.

Acknowledgments

We thank Martin Farrall and Jonathan Flint for discussions about linkage analysis, and Jackie King for technical support. This work was supported by the Medical Research Council and AstraZeneca UK Limited.

References

1. Aitman TJ, Glazier AM, Wallace CA, Cooper LD, Norsworthy PJ et al. (1999) Identification of *Cd36* (*Fat*) as an insulin-resistance gene causing defective fatty acid and glucose metabolism in hypertensive rats. Nat Genet 21, 76–83
2. Aszkenasy O, George R, Begg N (1995) Pneumococcal bacteremia and meningitis in England and Wales 1982 to 1992. CDR Rev 5, 45–58
3. Blackwell, JM (2001) Genetics and genomics in infectious disease susceptibility. Trends Mol Med 7, 521–526
4. Cervino AC, Lakiss S, Sow O, Hill AV (2000) Allelic association between the NRAMP1 gene and susceptibility to tuberculosis in Guinea-Conakry. Ann Hum Genet 64, 507–512
5. Churchill GA, Doerge RW (1994) Empirical threshold values for quantitative trait mapping. Genetics 138, 963–971
6. Cooke GS, Hill AVS (2001) Nat Rev Genet 2, 967–977
7. Darvasi A (1997) Interval-specific congenic strains (ISCS): an experimental design for mapping a QTL into a 1-centimorgan interval. Mamm Genome 8, 163–167
8. Dietrich WF, Miller J, Steen R, Merchant MA, Damron–Boles D et al. (1996) A comprehensive genetic map of the mouse genome. Nature 380, 149–152
9. Gingles NA, Alexander JE, Kadioglu A, Andrew PW, Kerr A et al. (2001) Role of genetic resistance in invasive pneumococcal infection: identification and study of susceptibility and resistance in inbred mouse strains. Infect Immun 69, 426–434
10. Kerr AR, Irvine JJ, Search JJ, Gingles NA, Kadioglu A et al. (2001) Role of inflammatory mediators in resistance and susceptibility to pneumococcal infection. Infect Immun 70, 1547–1557
11. Kruglyak L, Lander ES (1995) A nonparametric approach for mapping quantitative trait loci. Genetics 139, 1421–1428
12. Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ et al. (1987) MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1, 174–181
13. Lyons PA, Armitage N, Argentina F, Denny P, Hill NJ et al. (2000) Congenic mapping of the type 1 diabetes locus, Idd3, to a 780-kb region of mouse chromosome 3: identification of a candidate segment of ancestral DNA by haplotype mapping. Genome Res 10, 446–453
14. McCullers JA, Tuomanen EI (2001) Molecular pathogenesis of pneumococcal pneumonia. Front Biosci 6, 877–889
15. Miller SA, Dykes DD, Polesky HF (1988) A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16, 1215
16. Morton DB, Griffiths PH (1985) Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment. Vet Rec 116, 431–436
17. Mufson MA (1981) Pneumococcal infections. JAMA (J Am Med Assoc) 246, 1942–1948
18. Nadeau JH, Frankel WN (2000). The roads from phenotypic variation to gene discovery: mutagenesis versus QTLs. Nat Genet 25, 381–384
19. Nuorti JP, Butler JG, Cratcher JM, Guevara R, Welch D et al. (1998) An outbreak of multidrug-resistant pneumococcai pneumonia and bacteremia among unvaccinated nursing home residents. N Engl J Med 338, 1861–1868
20. Rhodes M, Straw R, Fernando S, Evans A, Lacey T et al. (1998) A high-resolution microsatellite map of the mouse genome. Genome Res 8, 531–542
21. Roy S, Hill AV, Knox K, Griffiths D, Crook D (2002a) Association of common genetic variant with susceptibility to invasive pneumococcal disease. BMJ 324, 1369
22. Roy S, Knox K, Segal S, Griffiths D, Moore CE et al. (2002b) MBL genotype and risk of invasive pneumococcal disease: a case-control study. Lancet 359, 1569–1573
23. Sokal R, Rohlf R (1995) In: Biometry, New York: WH Freeman, pp 423–427
24. Sorensen TI, Nielsen GG, Andersen PK, Teasdale TW (1988) Genetic and environmental influences on premature death in adult adoptees. N Engl J Med 318, 727–732
25. Tuomanen E (1999) Molecular and cellular biology of pneumococcal infection. 2, 35–39