Genet. Sel. Evol. 37 (2005) 539–561

DOI: 10.1051/gse:2005015

Original article

A genome scan for quantitative trait loci affecting the *Salmonella* carrier-state in the chicken

Pierre TILQUIN[a], Paul A. BARROW[b], José MARLY[c], Frédérique PITEL[d], Florence PLISSON-PETIT[d], Philippe VELGE[c], Alain VIGNAL[d], Philippe V. BARET[a], Nat BUMSTEAD[†,b], Catherine BEAUMONT[e*]

[a] Unité de Génétique, Faculté d'ingénierie biologique, agronomique et environnementale, Université catholique de Louvain, Croix du Sud 2 bte 14, B-1348 Louvain-la-Neuve, Belgium
[b] Institute for Animal Health, Compton, Berkshire RG20 7NN, United Kingdom
[c] Station de Pathologie Infectieuse et Immunologie, INRA, 37380 Nouzilly, France
[d] Laboratoire de Génétique Cellulaire, INRA, Chemin de Borde-Rouge, Auzeville BP27, 31326 Castanet-Tolosan, France
[e] Station de Recherches Avicoles, INRA, 37380 Nouzilly, France

(Received 5 August 2004; accepted 2 May 2005)

Abstract – Selection for increased resistance to *Salmonella* colonisation and excretion could reduce the risk of foodborne *Salmonella* infection. In order to identify potential loci affecting resistance, differences in resistance were identified between the N and 6_1 inbred lines and two QTL research performed. In an F2 cross, the animals were inoculated at one week of age with *Salmonella enteritidis* and cloacal swabs were carried out 4 and 5 wk post inoculation (thereafter called CSW4F2 and CSW4F2) and caecal contamination (CAECF2) was assessed 1 week later. The animals from the $(N \times 6_1) \times N$ backcross were inoculated at six weeks of age with *Salmonella typhimurium* and cloacal swabs were studied from wk 1 to 4 (thereafter called CSW1BC to CSW4BC). A total of 33 F_2 and 46 backcross progeny were selectively genotyped for 103 and 135 microsatellite markers respectively. The analysis used least-squares-based and non-parametric interval mapping. Two genome-wise significant QTL were observed on Chromosome 1 for CSW2BC and on Chromosome 2 for CSW4F2, and four suggestive QTL for CSW5F2 on Chromosome 2, for CSW5F2 and CSW2BC on chromosome 5 and for CAECF2 on chromosome 16. These results suggest new regions of interest and the putative role of *SAL1*.

fowl / genetic resistance / *Salmonella* / carrier-state / *SLC11A1*

* Corresponding author: Beaumont@tours.inra.fr

1. INTRODUCTION

In France as in other countries, *Salmonella* remains a major cause of human disease related to food consumption [5]. In one third of the cases, the serotype responsible for human food poisoning is *Salmonella enteritidis*. Poultry products are the main source of human *Salmonella* infections, mostly because chickens may be asymptomatic carriers (*i.e.* remain contaminated by *Salmonella* for several weeks without showing any symptom that could help their detection). Both caecal and ovarian *Salmonella* carrier-states may be involved in human contamination. While the latter may result in vertical transmission of *Salmonella* and in yolk contamination, the former is responsible for horizontal transmission of the bacteria and for human disease through contamination of the egg shell at the oviposition and of the carcass during evisceration. In both cases, the existence of asymptomatic carriers dramatically complicates the prophylaxis of this disease.

Food safety could potentially benefit from an increase in the genetic resistance of fowls to the *Salmonella* carrier-state (*i.e.* a better ability of the animals to clear *Salmonella*), which can be measured by the persistency of the bacterial infection after inoculation. In order to address this question, experimental models of infection were defined in chicks [15] and adult hens [31]. By using these models, the heritability of resistance was estimated at 0.20 in young birds [4] and more than 0.35 in laying hens [2]. These results show that the *Salmonella* carrier-state is partly genetically controlled. Selection for the reduced carrier state should be possible but would require experimental infection of animals unless the underlying genes responsible for resistance can be identified.

So far, two major genes of resistance to infection have been identified in mice and fowls: *NRAMP1* (natural resistance associated macrophage protein 1; now renamed *SLC11A1* for solute carrier family 11, member 1) and *TLR4*. Both are involved in resistance to mortality in 1 day-old chicks after intravenous inoculation [19, 27], and in bacterial replication in the spleen after oral inoculation [28]. At least the former is also involved in the control of resistance at older ages: the *SLC11A1* region also had a significant effect on the number of cfu (colony-forming units) in the spleen 3 days after inoculation of older pullets [16] and *SLC11A1* and probably *TLR4* are involved in the resistance of adult hens after oral contamination [3]. A third gene, named *SAL1*, was detected in fowls in a QTL research. The latter was achieved by a genome scan where susceptibility to acute visceral infection was appreciated as the spleen contamination five days after intravenous inoculation of two-week old chicks with *Salmonella typhimurium*. It was achieved in crosses between resistant and susceptible inbred lines of chickens [29].

However, the currently known genes which play a role in resistance to infection do not explain all the genetic variability in resistance to the carrier-state and the first goal of the work described in the present paper was to perform QTL mapping of genes affecting the carrier-state. Therefore differences in resistance to the *Salmonella enteritidis* carrier-state after inoculation at one week of age between four inbred lines were first observed. In parallel, additional studies on resistance to *Salmonella typhimurium* after inoculation at six weeks led to the evidence of large differences between the same lines but in a different order [1]. This led us to the hypothesis that genes controlling both traits could differ. It was therefore decided to perform two QTL research on both models of inoculation.

In both cases, bacterial counts in cloacal swabs or organs of slaughtered animals were used to measure *Salmonella* clearance in animals. Such phenotypes can present a distribution far removed from the normal distribution required by classical statistical methods. Therefore, a second objective of this paper was to compare the results obtained by a least-squares-based QTL mapping method with those obtained by a non-parametric method. Statistical aspects such as the distributional properties of the phenotypes related to the carrier-state and such as the size of the selectively genotyped sample will be discussed.

2. MATERIALS AND METHODS

2.1. Choice of poultry lines

The resistance of four inbred White Leghorn chicken lines was compared. IAH lines 6_1,15I and N originally derived from stock provided by the USDA Avian Disease and Oncology Laboratory, East Lansing, MI. Line C was the highly inbred line received by IAH from the Wellcome Research Laboratories, Beckenham. Parent birds were maintained under SPF conditions and were confirmed to be free of *Salmonella*.

Resistance at one week of age of a total of 99 one-week-old-chicks was assessed 5 weeks after oral inoculation of one-week-old-chicks with 5×10^4 cfu of *Salmonella enteritidis* (SE) strain 1009 [15]. Between 5 and 7 chicks per line were slaughtered each week until 7 weeks post inoculation and their total caeca cultured for Salmonella, before and, when the results were negative, after enrichment. Caecal contamination was assessed by the logarithm of the number of colony forming units (cfu) per gram of caecum ($\log_{10}$(cfu+1)).

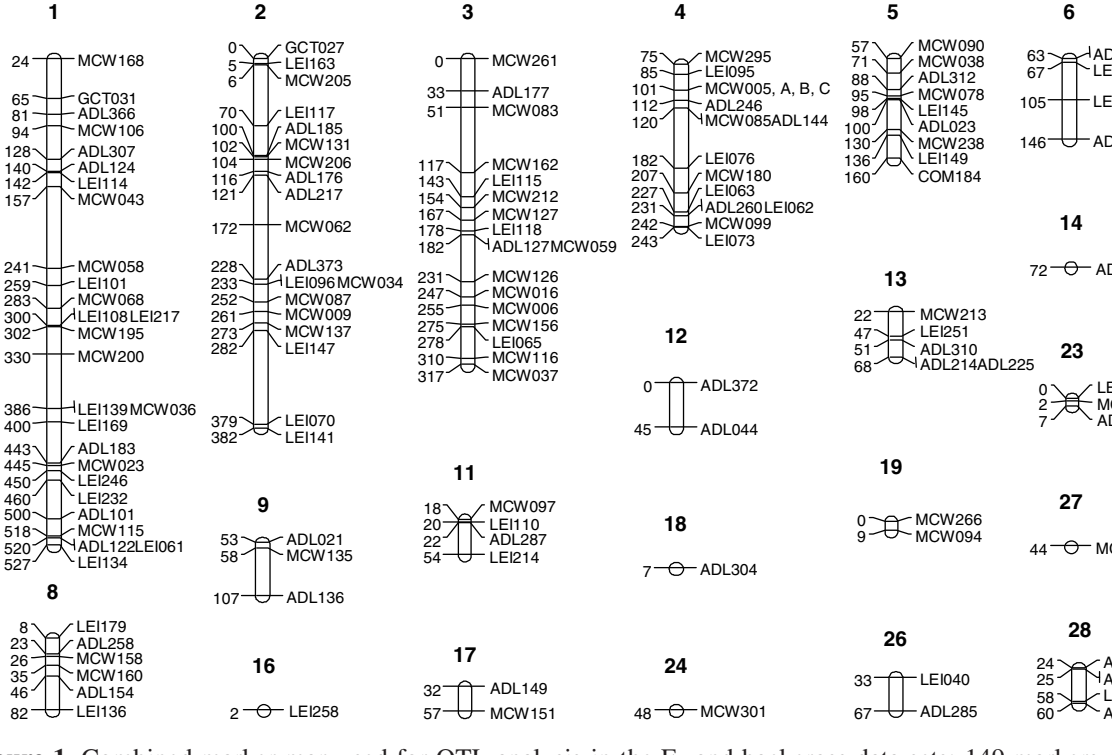

Figure 1. Combined marker map used for QTL analysis in the F_2 and backcross data sets; 149 markers c locations (cM) are expressed relative to the first marker on the linkage group.

2.2. QTL research

Two data sets were generated for this analysis: an F_2 and a backcross for resistance at one and six weeks of age respectively. In both experiments the birds of parental lines were included as controls.

In the F_2 data set, a total of 186 progeny were reared from a $(N \times 6_1) \times (N \times 6_1)$ cross and inoculated as formerly described ([15] and Sect. 2.1).Three phenotypic measurements were used in QTL mapping analyses: CSW4F2 and CSW5F5, *i.e.* the number of bacteria in the cloacal swabs performed four and five weeks post infection respectively and CAECF2, the number of cfu per gram of caecum after logarithmic transformation $[\log_{10}(x+1)]$. Codes were used for the assessment of cloacal swab counts (CSW4F2, CSW5F2) and were digitised as 0, 0.5, 1, 2, 4 and 8 to mimic $\log_{10}$ transformed counts.

Since only one third of the animals were contaminated (*i.e.* 41.3, 33.3 and 30.7% for CSW4F2, CSW5F2 and CAECF2 respectively), it was not possible to select the most resistant animals. It was therefore decided to choose the most susceptible (*i.e.* those with contaminated swabs and caeca) while the resistant animals were randomly chosen, within the same dam families, among the non contaminated animals. Thirty-three individuals, *i.e.* twenty-two susceptible and eleven resistant chickens, were thus selected at the two tails of the phenotypic distribution. Differences between both resistant and susceptible groups were tested using a Student test with the Satterwaithe correction in order to account for unequal variances [33, 34].

In the $(N \times 6_1)F1 \times N$ backcross experiment, a total of 80 progeny were reared and orally inoculated at six weeks of age with *Salmonella typhimurium* as formerly described [1]. Cloacal swabs were studied from wk 1 to 4 but only two phenotypic measurements, *i.e.* the number of cfu in the cloacal swab performed 1 and 2 weeks post infection were relevant since all individuals had zero counts after 3 weeks. These counts were logarithmically transformed $[log_{10}(x + 1)]$ in order to normalise their distribution and called CSW1BC and CSW2BC respectively. Fourty-six individuals (twenty-three at each tail of distribution) were selected for genotyping according to CSW1BC.

Pedigree information was obtained for the F_2 experiment but not for the backcross. The 33 F_2 individuals were shared among 11 full-sib families ranging between 2 and 6 full-sibs while the backcross experiment was treated as a single family of 46 full-sibs.

For both populations, sequences of the microsatellites previously described [10, 12, 17] were used in this study. For the F_2 cross and the backcross, selected individuals were respectively genotyped for a total of 199 and

Table I. Overall means and phenotypic standard deviations on the log_{10} scale for the traits studied in the F_2 an

Trait	Design	N	Age at inoculation (wk)	non-zero[a]	% zeros	Mean (SD) for genotyped animals	Mean (SD) for resistant animals	Mean (SD) for susceptib animals
Number of bacteria in cloacal swab 4 weeks p.i	F_2	33	1	14	57.6	0.62 (1.46)	0.07 (0.18)	1.73 (2.17)
Number of bacteria in cloacal swab 5 weeks p.i.	F_2	33	1	14	57.6	0.50 (1.00)	0.07 (0.18)	1.36 (1.38)
Number of bacteria per gram of caecum	F_2	33	1	11	66.7	1.11 (1.83)	0 (0)	3.34 (1.59)
Number of bacteria in cloacal swab 1 weeks p.i.	BC	46	6	23	50.0	0.82 (0.94)	0 (0)	1.64 (0.61)
Number of bacteria in cloacal swab 2 weeks p.i.	BC	46	6	5	89.1	0.07 (0.22)	0.09 (0.27)	0.07 (0.18)

[a] Number of selectively genotyped individuals with non-zero values.

177 markers. However, since only 81 informative markers were common to both data sets, and in order to have comparable results between both crosses in the interval mapping analysis, only informative markers belonging to the consensus linkage map of the chicken genome (http://www.thearkdb.org/ [35]) were used: *i.e.* 103 markers for the F_2 cross and 135 markers for the backcross for a total of 159 markers. In order to avoid the estimation of spurious linkage between markers located on different chromosomes, the marker map was not estimated using the selectively genotyped individuals [30]. Instead, the positions of markers on the consensus map were used. One-hundred and forty-nine markers were positioned on one of 25 linkage groups, the remaining 10 markers remain unlinked in the consensus map (Fig. 1).

2.3. Statistical analyses of QTL research

2.3.1. QTL analysis

Markers that were either isolated representative of linkage groups or unlinked were analysed individually using a single-marker test. In one or both data sets, chromosomes (Chr.) 9, 14, 16, 17, 18, 19, 24, 26, 27, E47W24 and E22C19W28 had only a single marker.

Interval mapping analysis was performed on 90 markers for the F_2 (over 16 linkage groups) and 119 markers for the backcross (over 17 linkage groups). The phenotypic measurements from the individuals that were not selected for genotyping were not used in the interval mapping analysis.

In order to avoid numerical problems due to recombinants between markers with zero separation, when two markers were located on the same position, markers were artificially separated by a distance of 1-cM and their distance from the next marker was corrected accordingly. The characteristics of both experimental designs are summarised in Table I.

For the analysis of the F_2 data set, the influence of sex (fixed) and cage (random) for every trait was initially estimated by analysis of variance using SAS® statistical software [33]. Both factors were found not to be significant and were therefore not included as effects in the linkage analysis. For the analysis of the backcross data set, no information was available about any possible fixed or random effects.

Two types of QTL mapping software were used. Firstly, the QTL express program [36] was used to perform least-squares-based interval mapping [18]. This approach is based on the regression of phenotypes on probabilities of inheriting the QTL at the position being tested and assumes that the distribution of the phenotype is normal. Two strategies were used to cope with this

issue. On the one hand, a logarithmic transformation was applied to all traits before least-squares-based analysis. On the other hand, since logarithmically transformed phenotypes do not necessarily satisfy the normality assumption, non-parametric (NP) interval mapping was performed using R/qtl [7].

In F_2 designs, the non-parametric interval mapping test was based on an extension of the Kruskal-Wallis test which is similar to the method described by Kruglyak and Lander [24] for such designs. In the case of incomplete genotype information (such as at locations between genetic markers), the Kruskal-Wallis statistic is modified so that the rank for each individual is weighted by the genotype probabilities (for more details, see [6]). In backcross designs, the extension of the Wilcoxon rank-sum test [24] was used. Due to the high number of ties for all phenotypes in both crosses, tied phenotypes were given midranks rather than random ranks [39].

2.3.2. Significance thresholds

First, the significance level (*i.e.* the *P*-value) of each test statistics value at every position was determined empirically by chromosome-wide permutations [11]. A total of 10 000 permutations were performed for each chromosome × trait combination.

Then, threshold *P*-values for *suggestive* and genome-wide *significant* linkage were obtained, based on the actual marker density and not on an infinitely dense marker map [26]. Suggestive linkage was first obtained as the probability of obtaining, by chance, one significant result per genome analysis [26]. In order to take into account the differences between chromosome lengths (macro- and micro-chromosomes), the chromosome-wide *P*-value for suggestive linkage (P_c) of a specific chromosome was computed as the contribution (r) of that chromosome to the total genome length, which was obtained by dividing its length by the total autosomal length of the chicken genome, both given by the consensus map ([35] and Tab. II). This approach was used as an alternative to the classical suggestive threshold computed as $1/N$ where N is the number of autosomes [22]. Instead of using the consensus lengths (for the ratio of chromosomes and genome lengths), Tuiskula-Haavisto *et al.* [40] used the length of the chromosome covered by markers as the numerator of this ratio, and the length of the genome covered by the markers they analysed as the denominator. Threshold *P*-values obtained using this method are also given in Table II but were not used in the analyses.

In order to derive chromosome-wide significant *P*-values (P_c) corresponding to a 5% genome-wide significance level (P_g) [26], the following Bonferroni

Table II. Thresholds for *suggestive* and genome-wide *significant* linkage computed using consensus lengths of chromosomes and of the genome.

Chr.	Map length (cM) [a]		Threshold value [b] %		Tuiskula-Haavisto *et al.* [40] [c] %	
	consensus	covered	*suggestive*	significant	*suggestive*	significant
1	566	504	0.14513	0.00742	0.21884	0.01116
2	482	383	0.12359	0.00632	0.16630	0.00849
3	318	318	0.08154	0.00417	0.13808	0.00706
4	271	169	0.06949	0.00356	0.07338	0.00376
5	199	104	0.05103	0.00261	0.04516	0.00231
6	147	84	0.03769	0.00193	0.03647	0.00187
7	166	118	0.04256	0.00218	0.05124	0.00262
8	106	75	0.02718	0.00139	0.03257	0.00167
9	133	55	0.03410	0.00175	0.02388	0.00122
11	89	37	0.02282	0.00117	0.01607	0.00082
12	91	46	0.02333	0.00120	0.01997	0.00102
13	75	47	0.01923	0.00099	0.02041	0.00105
14	78	11	0.02000	0.00103	0.00478	0.00024
15	72	50	0.01846	0.00095	0.02171	0.00111
16	61	11	0.01564	0.00080	0.00478	0.00024
17	71	26	0.01821	0.00093	0.01129	0.00058
18	48	11	0.01231	0.00063	0.00478	0.00024
19	41	10	0.01051	0.00054	0.00434	0.00022
23	14	8	0.00359	0.00018	0.00347	0.00018
24	59	11	0.01513	0.00078	0.00478	0.00024
26	68	35	0.01744	0.00089	0.01520	0.00078
27	70	11	0.01795	0.00092	0.00478	0.00024
28	76	37	0.01949	0.00100	0.01607	0.00082
E22C19W2	51	11	0.01308	0.00067	0.00478	0.00024
E47W24	27	21	0.00692	0.00036	0.00912	0.00047
unlinked	11	11	0.00282	0.00014	0.00478	0.00024
Total	3900	2303				

[a] Lengths of chromosomes according to the consensus map or to the distances covered by markers; the contribution of unlinked markers ($n = 10$) to the total genome length taken as 1 ± 5 cM (11 cM); this length of segment was taken as a reasonable estimate of the extent of linkage disequilibrium around a marker.

[b] *P*-values for *suggestive* linkage computed as the contribution of a chromosome to the total autosomal length of the chicken genome, both given by the consensus map; *P*-values for *significant* linkage at the chromosome level computed to guarantee a 5% genome-wide significance level using a Bonferroni correction (see Eq. (1)).

[c] *P*-values computed using chromosome lengths and genome length covered by markers as applied by Tuiskula-Haavisto *et al.* [40].

correction [40] was applied for each chromosome separately:

$$P_g = 1 - (1 - P_c)^{1/r}. \quad (1)$$

In order to make the comparison possible between different studies, significance thresholds did not take the testing of multiple traits into account [14]. However, within an experiment (F_2 or backcross), among all traits, the permutation distribution with the most extreme 5% threshold was taken to compute the $P-$values and *suggestive* and *significant* threshold values for all traits in that experiment.

Single marker tests were carried out on the isolated markers by setting all markers evenly spaced (every 10 cM) on a single chromosome and by using both approaches in order to obtain test statistics values for each isolated marker. Single-position permutations were performed to obtain significance levels of single marker tests, and the threshold P-value to apply for those tests was computed as for interval mapping tests, taking 1 ± 5 cM (11 cM) as the contribution of an isolated or unlinked marker to the total genome length. This length of the segment was taken as a reasonable estimate of the extent of linkage disequilibrium around a marker.

3. RESULTS

3.1. Choice of the poultry lines

Significant differences between lines ($P < 1.10^{-4}$) as well as the interaction between interval post inoculation and lines could be observed. When restricting the analysis to the weeks 5 to 7 (*i.e.* long-term carrier-state), lines 6_1 and C appeared to be the most susceptible (with least squares means equal to 4.94 and 4.63 respectively) and lines N and 15 the most resistant (with estimated least squares means at 3.59 and 3.26). The second comparison confirmed previous results: the line 6_1 appeared to be significantly more susceptible. A difference in the rate of excretion could also be observed, with line 6_1 excreting more often than line N. Since the degree of dominance of resistance seemed to vary with the post inoculation interval, F2 crosses were chosen for the detection of QTL.

When considering resistance to oral inoculation at six weeks of age, large differences were observed [1]. The pattern of faecal shedding of the F1 birds closely resembled those of line 6_1 with rapid elimination of the challenge strain and low rates of excretion. The dominant resistance was expressed fairly early after challenge and it was this rapid effect that was inherited. A backcross was therefore chosen for QTL detection.

3.2. QTL analysis

The total length of the chicken genome covered by markers was 2303 cM or about 59% of the consensus linkage map. This figure includes an arbitrary length of 11 cM for each unlinked marker. In the sample of 140 markers belonging to the 19 linkage groups with more than one marker, the average distance between adjacent markers (± S.D.) was 17.0 ± 18.6 cM. The information content obtained across the chicken genome averages 68.8% for the F2 and 70.1% for the backcross, ranging from 20% to 100% (data not shown).

The overall means, phenotypic standard deviations and proportion of zeros in selected animals for the five phenotypic traits are given in Table I. Highly significant differences were observed between resistant and susceptible animals for all traits but CSW2BC (Tab. I). Due to the selective genotyping approach used for both crosses, the proportion of individuals with a zero value was always higher or equal to 50% (see Tab. I). All distributions were therefore asymmetric, although phenotypic values were a priori log-transformed (Fig. 2).

All regions of the genome that provided support for segregating QTL at the *suggestive* and *significant* levels are reported in Table III. While it is probable that some of suggestive QTL are false positives, it is generally regarded as informative to the mapping community to report all regions that offer any evidence of linkage [26].

A total of six regions (QTL) of the chicken genome which could potentially affect the clearance of *Salmonella* were identified by using the least-squares-based approach (Tab. III). The QTL on Chr. 2 around position 87 for CSW4F2 was genome-wide significant ($P = 0.0032$, marker interval GCT027–ADL185) and the QTL for CAECF2 on Chr. 16 at position 2 was nearly significant ($P = 0.0012$, marker LEI258). Five of the six QTL were identified either in the F_2 data set or in the backcross. One suggestive QTL on Chr. 5 (position 100–111) was confirmed in both data sets. On Chr. 1, two QTL were identified at position 85 in the F_2 and at position 207 in the backcross. In the backcross, the two QTL identified on Chr. 1 (position 206–207 cM) and on Chr. 5 (position 100–106 cM, around marker ADL023) were confirmed by the non-parametric analysis. Furthermore, the QTL on Chr. 1 was *significant* ($P = 0.0060$) in the non-parametric analysis rather than *suggestive* in the least-squares-based analysis.

Among the six identified QTL, three (or four if there are two QTL on Chr. 1) QTL explained more than 7% of phenotypic variance and could therefore be substantially involved in the determination of resistance to *Salmonella* in chickens. The QTL identified on Chr. 1 was estimated to explain 7.7% of the phenotypic variance of CSW5F2 in F_2 and 37.5% of the phenotypic variance

Table III. QTL identified in the interval mapping analysis performed on 25 of the 38 chicken autosomes.

Trait[a]	Chr.	Threshold *P*-values	Design	Least-squares based mapping						Non
				Position, cM	F_{MAX}	*P*-value[b]	Additive effect ± s.e.	Dominance effect ± s.e.	% var.[c]	Posi cM
CSW5F2	1	0.00742	F_2	85	11.56	0.0189 *	2.0 ± 0.4	−1.7 ± 0.5	7.7	50
CSW2BC	1	0.00742	BC	207	8.51	0.0541 *	−0.9 ± 0.3	–	37.5	20
CSW4F2	2	0.00632	F_2	87	15.15	0.0032 **	2.1 ± 0.4	−4.7 ± 0.9	7.0	9
CAECF2	5	0.00261	F_2	111	4.64	0.0746 ns	1.4 ± 0.5	−1.6 ± 0.9	0.8	11
CSW5F2	5	0.00261	F_2	111	9.01	0.0071 *	0.9 ± 0.3	−1.2 ± 0.4	2.3	11
CSW2BC	5	0.00261	BC	100	10.36	0.0084 *	0.7 ± 0.2	–	21.6	10
CAECF2	11	0.00117	F_2	18	4.17	0.0614 ns	−0.9 ± 0.4	−1.2 ± 0.6	0.4	1
CSW5F2	11	0.00117	F_2	18	7.02	0.0122 *	−0.7 ± 0.2	−0.7 ± 0.3	1.0	1
CAECF2	16	0.0008	F_2	2	9.76	0.0012 *	−0.5 ± 0.4	−2.3 ± 0.5	0.7	2

[a] See Table I for the definition of the traits.

[b] *P*-values obtained by 10000 chromosome-wide permutations; threshold *P*-values for *suggestive* (*) and genome-wide *signific* the length of the chromosome relative to the total genome length studied; ns = not significant.

[c] Genetic variance explained by the QTL based on estimated additive and dominance effects and allele frequencies 0.5, as a perce in the entire F_2 or backcross population (*i.e.* before selective genotyping to moderate overestimation).

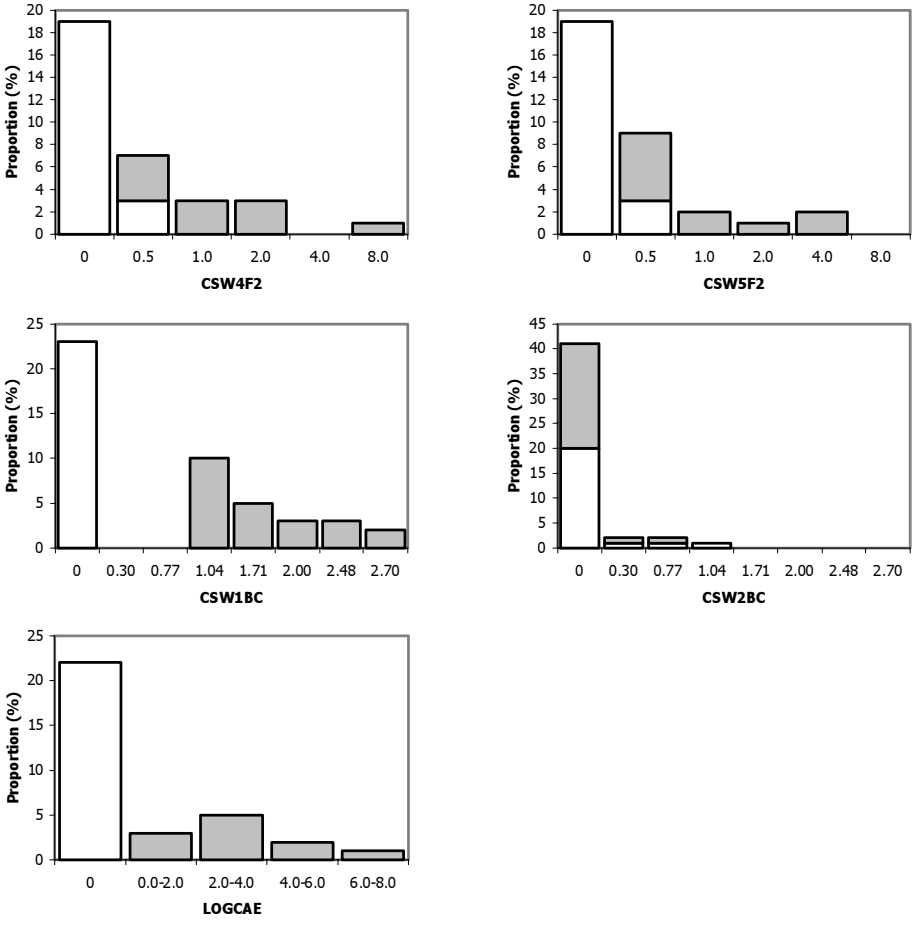

Figure 2. Phenotypic distributions, for animals selectively genotyped, on the log_{10} scale for the traits studied. Their numbers were equal to 33 and 46 in the F2 and in the backcross respectively. Grey areas correspond to the susceptible group.

of CSW2BC in the backcross. The significant QTL on Chr. 2 in the F_2 explained 7.0% of the phenotypic variance of CSW4F2. The QTL positioned on Chr. 5 explained 21.6% of the phenotypic variance of CSW2BC in the backcross, but only 2.3% of CSW5F2 in the F_2. Although it has the smallest *P*-value ($P = 0.0012$), the QTL identified on Chr. 16 (marker LEI258) explained only 0.7% of the phenotypic variance of the number of bacteria per gram of caecum (CAECF2).

Figures 3 and 4 show QTL on chromosomes 1 and 2, respectively, for the five disease resistance traits using the LS approach (a) as well as the

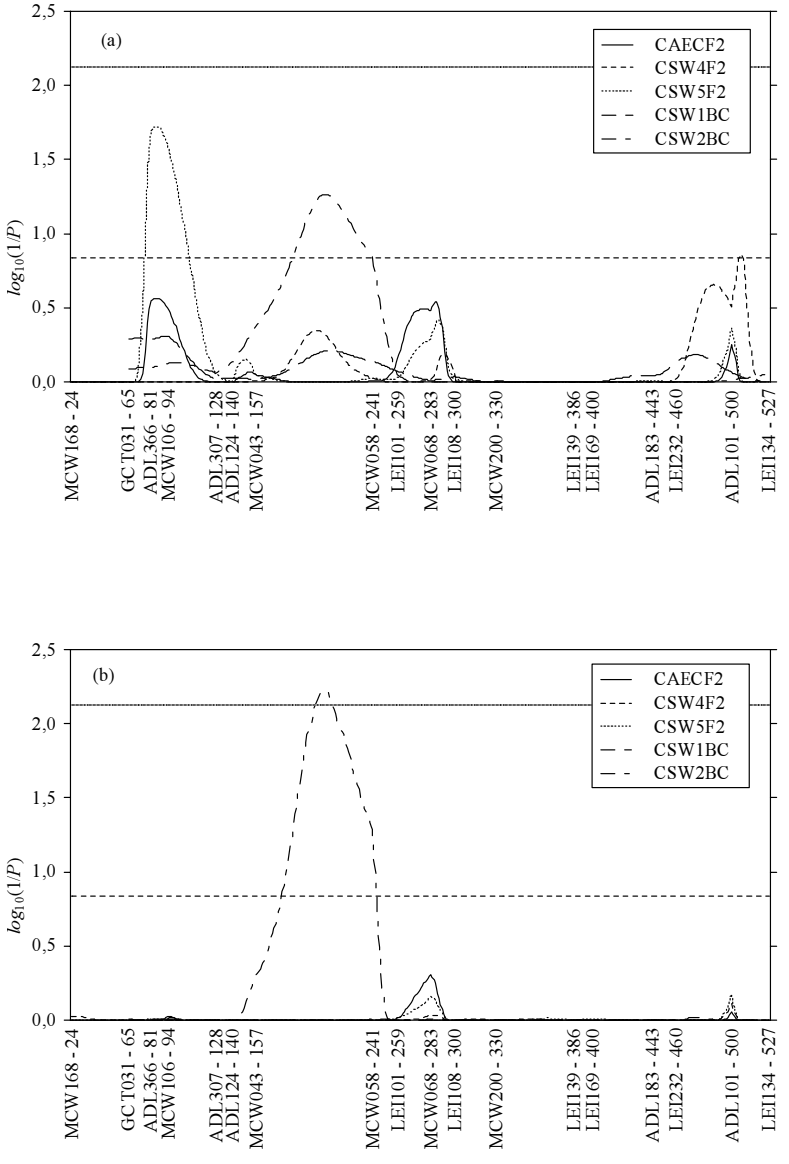

Figure 3. Location scores (in cM) obtained from the least-squares (a) and non-parametric (b) interval mapping analyses of logarithmically transformed number of bacteria per gram of caecum and cloacal swab counts on chicken Chr. 1 for the F_2 and BC populations. Evidence in favour of the presence of a QTL (y-axis) is measured as $log_{10}(1/P)$, where P is the associated chromosome-wide P-value determined by phenotype permutations. Horizontal dashed lines (large dots and small dots) are respectively *suggestive* and genome-wide *significant* thresholds. Note that some markers have been omitted from the x-axis for clarity.

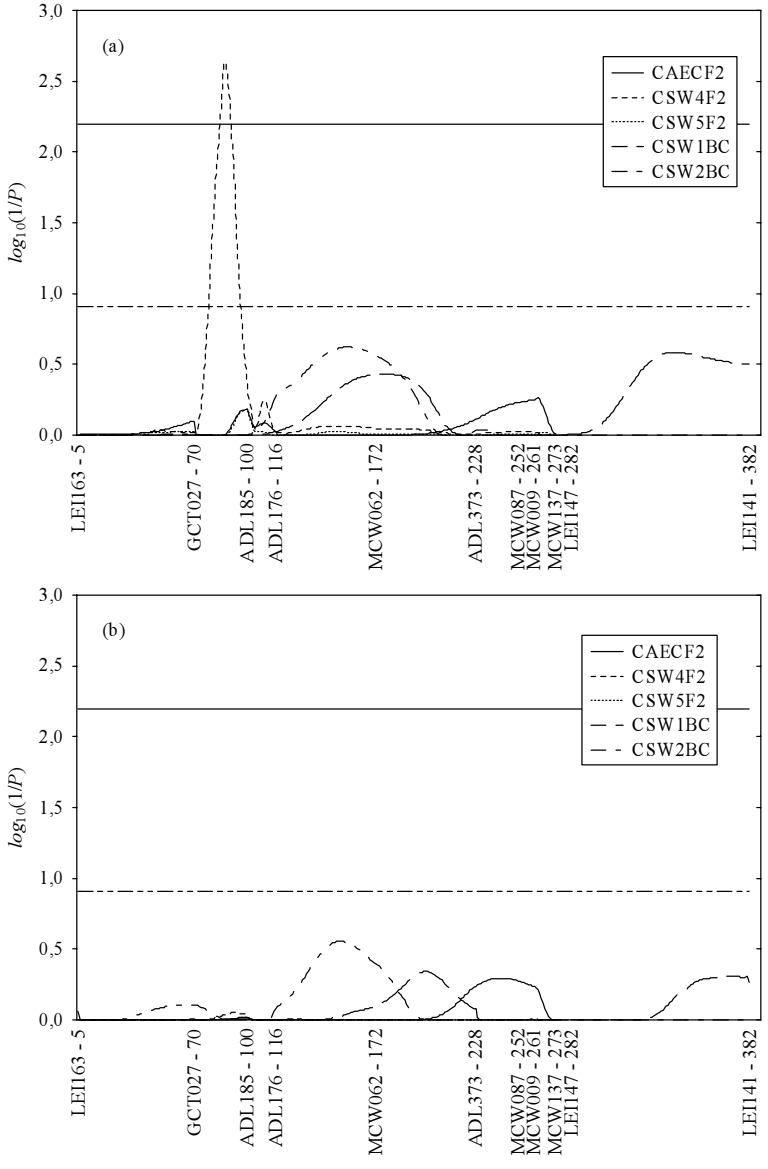

Figure 4. Location scores (in cM) obtained from the least-squares (a) and non-parametric (b) interval mapping analyses of logarithmically transformed number of bacteria per gram of caecum and cloacal swab counts on chicken Chr. 2 for the F_2 and BC populations. Evidence in favour of the presence of a QTL (*y*-axis) is measured as $log_{10}(1/P)$, where P is the associated chromosome-wide P-value determined by phenotype permutations. Horizontal dashed lines (large dots and small dots) are respectively *suggestive* and genome-wide *significant* thresholds. Note that some markers have been omitted from the *x*-axis for clarity.

non-parametric approach (b). The curves show the log_{10} of the inverse of the chromosome-wide *P*-values of the data under the null hypotheses of no QTL at the corresponding position estimated from 10 000 permutations. Because of the *F* ratio curves (data not shown) and the low proportions of phenotypic variance explained, the QTL identified on chromosome 11 was considered as probably being spurious.

4. DISCUSSION

Significant differences could be observed between inbred poultry lines. However, the relative resistance of inbred poultry lines differed. Indeed, line N was resistant when the poultry were inoculated at one week of age with *Salmonella enteritidis* and when faecal excretion was measured 4 and 5 weeks later. However, the line was susceptible after inoculation at 6 weeks of age with *Salmonella typhimurium* and when excretion was assessed one and two weeks after inoculation. This led us to the hypothesis of a difference in the genetic control of both traits. This can originate from a difference in mechanisms of resistance. Indeed they may result from differences in inoculated *Salmonella* (according to the importance of the host-pathogen dialog), in age at inoculation (that may be important, especially since the immune response develops during the first weeks of age, as shown on one of these lines [32]) and in interval post inoculation, in accordance with results strongly suggesting that the wild allele for *Nramp1* gene could be favourable with regards to resistance to inoculation (a few days after inoculation) but not to bacterial clearance 7 weeks post inoculation [9].

However, whatever the origin of these differences, this between line variability could be used for QTL research. The use of inbred lines allowed to take advantage of the lower genetic variability of those lines and of the knowledge of their informativity at a large range of genetic markers. Indeed the size of the data sets was rather low which results in a higher probability of false negative results. The latter was still increased due to the non-normality of the phenotypes and to the high number of ties [39], which is related to the high proportion of animals having cleared all the bacteria. However, significant or nearly significant QTL could be observed in both data sets. Most of the QTL identified in these experiments had an effect on only one trait, which can be explained by a lack of statistical power but also by differences in genetic control. This result enhances one of the hypotheses that underlined this experiment, *i.e.* that

resistance to colonisation (number of bacteria per gram of caecum) and excretion (cloacal swabs) might be affected by different genes. This result will have to be considered for any practical application. Only two of the QTL were identified in both crosses: while both age at inoculation and *Salmonella* serotypes differed between the two crosses, the former seemed to be more responsible for between line differences, since similar relative resistance of lines to caecal excretion were observed after inoculation with *Salmonella typhimurium*, *Salmonella enteritidis* or *Salmonella infantis* [1]. Moreover measures in the backcross were taken earlier (because of quicker clearance of *Salmonella*) and could partly involve different genes.

As opposed to the QTL identified in the backcross, all QTL identified in the F_2 using least-squares-based interval mapping were not confirmed by the non-parametric analysis. This can be explained by the small number of genotyped individuals ($n = 33$) in the F_2. Indeed, in the F_2, the 33 genotyped individuals are distributed among three genotypic classes (*e.g.* NN, $N6_1$ or 6_16_1), as opposed to two genotypic classes in the backcross (NN or $N6_1$) in which 46 individuals were genotyped. This aspect reduces the power of the Kruskal-Wallis test (analogue of an ANOVA test) performed in the non-parametric analysis. Furthermore, since the non-parametric test uses ranks instead of actual values, its power is reduced compared to the parametric test. Since the least-squares-based interval mapping method is more powerful than the non-parametric test and is quite robust against non-normality (see [38, 39]), it may have greater value in this situation.

The QTL effects estimated in this study were overestimated due to the selective genotyping approach used in both F_2 and backcross data sets. Formulae are given [13] to convert observed QTL effects under selective genotyping to actual gene effects, but these are based on the assumption that the phenotype is normally distributed and were therefore not applicable in the context of this study. In order to moderate overestimation, the proportions of variance explained by QTL were computed using the observed variance in the full population before selective genotyping rather than that of the selected sample.

The concordance of the observed maximum test statistic values with the observed phenotypes and genotypes was checked (data not shown) by comparing in terms of phenotypic values the genotype classes of the markers either at the position of the identified QTL or for the pair of markers bracketing a QTL position. This analysis confirmed the QTL on chromosomes 1 and 5. In addition, these QTL were in accordance with the pattern expected for inheritance of 6_1

and N alleles. However, phenotypic and inheritance patterns of the QTL identified on chromosomes 2, 11 and 16 were less clear. This result, in addition to the *F* ratio curve of chromosome 11 and the low proportion of phenotypic variance explained, led to the conclusion that the QTL observed on chromosome 11 was probably spurious, while the QTL on chromosomes 2 and 16 should be further tested in a larger population in order to obtain better estimates of their effects.

The QTL on Chr. 1 was the QTL explaining the greatest proportion of variance of cloacal swabs (7.7% in the F_2 and 37.5% in the backcross). The difference in their positions of 122 cM could reflect an error in position estimates for a single QTL due to the low information content in that region in both data sets or to different numbers of markers in that region between data sets (3 and 5 markers respectively for the F_2 and backcross). At least one of these QTL may relate to that identified in mice on chromosome 7 between microsatellites D7Mit83 and D7mit62 [9] (*i.e.* at positions 26.5 and 42.6 respectively, http://www.informatics.jax.org, Mouse Genome Informatics). Indeed, from the review of comparative mapping [35], chicken chromosomal regions corresponding to mouse Chr. 7 are split between chromosomes 10, 6 and proximal Chr. 1 (between positions 157 and 163) (mouse position 4), *i.e.* at about 50 cM from the estimated position of this chicken QTL. Kaiser *et al.* [21] observed by bulked segregant analysis, two unlinked microsatellites of Chr. 1 associated with the response to the *Salmonella* vaccine; however these are also at a more distal position, at 286 and 456 cM. The sire allele for those microsatellites also had significant effects on the spleen and caecal *Salmonella* counts one week after oral inoculation of one day-old chicks for two of the sires tested [20]. These coherent results reinforce the interest of those QTL even if no significant difference could be observed for CSW2BC between resistant and susceptible animals.

No QTL involved in the resistance to *Salmonella* has been previously detected on Chr. 2 and up to now few genes have been identified in this region in chickens. The MIFL1 gene (macrophage migration inhibitory factor, like 1) which lies close to LEI117 is a possible candidate gene, but it is not currently possible to relate this region to human or murine chromosomal regions.

SAL1 which is located at about 150 cM on chicken Chr. 5 could potentially contribute to the QTL observed on this chromosome at about 100–111 cM. An effect of the microsatellite ADL298 (at 198 cM, *i.e.* further distal on this chromosome) on the response to the *Salmonella* vaccine and on caecal and spleen bacterial burden 1 week after oral inoculation of day-old chicks respectively

was observed [20, 21], like the QTL detected in this region in the present study which have an effect on early infection. Recent studies [23] have suggested a possible role of transforming growth factor β3 (TGFB3) mapped on Chr. 5 at position 113 next to marker ADL023 [37] on spleen contamination, but no effect could be detected on caecal or liver contamination. The QTL identified in both data sets on Chr. 5 is very close to the location of TGFB3. If confirmed, this result suggests a possible role of TGFB3 in the mechanism underlying the resistance of chickens to *Salmonella* infection.

The QTL on Chr. 11 was weakly significant and accounted for little of the variance and hence may be spurious. No other observations suggesting a role of this chromosome in *Salmonella* resistance has been noted yet.

Even if its effect is very small, the LEI258 marker on Chr. 16 (position 2) is of particular interest since it lies very close to the major histocompatibility complex (MHC). Indeed, a significant effect of SNP within MHC was observed on splenic (but not caecal) bacterial burden 1 week after oral inoculation of one-day-old chicks [25], even if no MHC-associated effect could be noted on resistance to infection [8]. However it seems unlikely that MHC could be involved in this cross since lines 6_1 and N are of the same MHC haplotype.

As a final result of this paper, we propose a new method to compute *suggestive* and genome-wide *significant* thresholds for QTL analyses in the chicken taking into account the differences in chromosome lengths. Classical suggestive thresholds are computed as $1/N$ where N is the number of autosomes [22]. The length of the chromosome and of the genome covered by markers used for genotyping may be considered [40] but this method does not allow comparing the results from different analyses. Our approach has the considerable merit of guaranteeing the comparability of the results between different QTL mapping studies in the chicken. The same threshold *P*-values could be used for both F_2 and backcross data sets although the number of markers genotyped in each data set was different (Tab. I). As shown by the small number of genome-wide significant results obtained from this study, the stringency of the threshold *P*-values obtained by this approach is quite high, and is higher for micro-chromosomes than for macro-chromosomes. This can be explained by the fact that since they contain more mapped positions than micro-chromosomes, macro-chromosomes have a higher chance of hosting a random false positive result in the course of a genome scan. This chance corresponds to the suggestive level [26]. When using the length of the chromosome [40] to compute threshold *P*-values, for macro-chromosomes this chance is even higher. Threshold *P*-values are therefore less conservative and, moreover, cannot be generalised to every QTL mapping study in the chicken.

Finally, this study has shown that the phenotypes related to the carrier-state (cloacal swab counts or bacterial colonisation in the caecum) have specific distributions (far from normality) that are very dependent of the time at which the sampling is performed. If the course of the infection is not followed carefully, the peak of infection may be missed and differences between animals in terms of resistance may have nearly disappeared at the time of sampling. Although this issue was addressed in the present experiments, in future experiments, even greater attention should be paid to the time of sampling to obtain phenotypic measurements representing the greatest differences between individuals in terms of resistance. Furthermore, this study has shown that a selective genotyping approach applied in the context of the carrier-state implies that a high proportion of individuals will have zero values, since resistant and susceptible animals need to be selected. In this study for example, 50% of the individuals had zero values, which considerably skewed the distribution of the phenotypes, and therefore decreased the expected power of QTL detection. In such a context, if selective genotyping is chosen, a high proportion of selection should be preferred. However, even with these drawbacks, this study showed the existence of several QTL.

In conclusion, this study showed the existence of several QTL affecting caecal colonisation in chickens. If confirmed in commercial populations, these will be of great help in practical application to reduce the possibility of food contamination and to increase our understanding of the mechanisms controlling *Salmonella* colonisation in chickens.

ACKNOWLEDGEMENTS

The authors dedicate this publication to the memory of Nat Bumstead. They are grateful to George Seaton and Karl Broman for helping us in using QTL express and R/qtl software respectively, and to Nigel Salmon for his excellent technical assistance in genotyping the F_2 and backcross panels. We thank Steve Bishop, Dirk-Jan de Koning and Michel Georges for helpful suggestions on statistical aspects. The authors are grateful to Axel Boucard, Nicole Millet, Nadine Sellier and to the experimental unit of the Station de Recherches Avicoles for realization of the protocol. They thank the referees and editor for their interesting comments and questions. Financial support is acknowledged from the Université catholique de Louvain, from the BBSRC and European Union (FAIR CT-98-4311).

REFERENCES

[1] Barrow P.A., Bumstead N., Marston K., Lovell M.A., Wigley P., Faecal shedding and intestinal colonization of Salmonella enterica in in-bred chickens: the effect of host-genetic background, Epidemiol. Infect. 132 (2004) 117–126.

[2] Beaumont C., Protais J., Guillot J.F., Colin P., Proux K., Millet N., Pardon P., Genetic resistance to mortality of day-old chicks and carrier-state of hens after inoculation with *Salmonella enteritidis*, Avian Path. 28 (1999) 131–135.

[3] Beaumont C., Protais J., Pitel F., Leveque G., Malo D., Lantier F., Plisson-Petit F., Colin P., Protais M., Le Roy P., Elsen J.M., Milan D., Lantier I., Neau A., Salvat G., Vignal A., Effect of two candidate genes on the *Salmonella* carrier state in fowl, Poult. Sci. 82 (2003) 721–726.

[4] Berthelot F., Beaumont C., Mompart F., Girard-Santosuosso O., Pardon P., Duchet-Suchaux M., Estimated heritability of the resistance to cecal carrier state of *Salmonella enteritidis* in chickens, Poult. Sci. 77 (1999) 797–801.

[5] Bouvet P.J.M., Fougerat I., Guesnier F., Guibert F., K'Ouas G., Lenormand P., Metz L., Ruckly C., Grimont P.A.D., Human salmonellosis surveillance in France: recent data from the national referee center, in: Proceedings of the International Symposium on Salmonella and Salmonellosis, 29–31 May 2002, Institut Supérieur des Productions Animales et Industries Agro-Alimentaires, pp. 411–416.

[6] Broman K.W., QTL mapping in the case of a spike in the phenotype distribution, Genet. 163 (2003) 1169–1175.

[7] Broman K.W., Wu H., Sen S., Churchill G.A., R/qtl: QTL mapping in experimental crosses, Bioinform. 19 (2003) 889–890.

[8] Bumstead N., Barrow P., Genetics of resistance to *Salmonella* as measure of inbreeding in domestic fowls, Brit. Poult. Sci. 28 (1988) 717–726.

[9] Caron J., Loredo-Osti J.C., Laroche L., Skamene E., Morgan K., Malo D., Identification of genetic loci controlling bacterial clearance in experimental *Salmonella enteritidis* infection: an unexpected role of Nramp1 (Slc11a1) in the persistence of infection in mice, Genes Immun. 3 (2002) 196–204.

[10] Cheng H.H., Levin I., Vallejo R.L., Khati H., Dodgson J.B., Crittenden L.B., Hillel J., Development of a genetic map of the chicken with markers of high utility, Poult. Sci. 74 (1995) 1855–1874.

[11] Churchill G.A., Doerge R.W., Empirical threshold values for quantitative trait mapping, Genet. 138 (1994) 963–971.

[12] Crooijmans R.P.M.A., van Oers P.A.M., Strijk J.A., van der Poel J.J., Groenen M.A.M., Preliminary linkage map of the chicken (*Gallus domesticus*) genome based on microsatellite markers: 77 new markers mapped, Poult. Sci. 75 (1996) 746–754.

[13] Darvasi A., Soller M., Selective genotyping for determination of linkage between a marker locus and a quantitative trait locus, Theoret. Appl. Genet. 85 (1992) 353–359.

[14] de Koning D.J., Janss L.L.G., Rattink A.P., van Oers P.A.M., de Vries B.J., Groenen M.A.M., van der Poel J.J., de Groot P.N., Brascamp E.W.,

van Arendonk J.A.M., Detection of quantitative trait loci for backfat thickness and intramuscular fat content in pigs (*Sus scrofa*), Genet. 152 (1999) 1679–1690.

[15] Duchet-Suchaux M., Lechopier P., Marly J., Bernardet P., Delaunay R., Pardon P., Quantification of experimental *Salmonella enteritidis* carrier state in B13 leghorn chicks, Av. Dis. 39 (1995) 796–803.

[16] Girard-Santosuosso O., Lantier F., Lantier I., Bumstead N., Elsen J.M., Beaumont C., Heritability of susceptibility to *Salmonella enteritidis* infection in fowls and test of the role of the chromosome carrying the NRAMP1 gene, Genet. Sel. Evol. 34 (2002) 211–219.

[17] Groenen M.A.M., Crooijmans R.P.M.A., Veenendaal A., Cheng H.H., Siwek M., van der Poel J.J., A comprehensive microsatellite linkage map of the chicken genome, Genom. 49 (1998) 265–274.

[18] Haley C.S., Knott S.A., A simple regression method for mapping quantitative trait loci in line crosses using flanking markers, Heredity 69 (1992) 315–324.

[19] Hu J., Bumstead N., Barrow P., Sebastiani G., Olien L., Morgan K., Malo D., Resistance to salmonellosis in the chicken is linked to NRAMP1 and TNC, Genome Res. 7 (1997) 693–704.

[20] Kaiser M.G., Lamont S.J., Microsatellites linked to *Salmonella enterica* serovar *enteritidis* burden in spleen and cecal content of young F1 broiler-cross chicks, Poult. Sci. 81 (2002) 657–663.

[21] Kaiser M.G., Deeb N., Lamont S.J., Microsatellite markers linked to *Salmonella enterica* serovar *enteritidis* vaccine response in young F1 broiler-cross chicks, Poult. Sci. 81 (2002) 193–201.

[22] Knott S.A., Marklund L., Haley C.S., Andersson K., Davies W., Ellegren H., Fredholm M., Hansson I., Hoyheim B., Lundstrom K., Moller M., Andersson L., Multiple marker mapping of quantitative trait loci in a cross between outbred wild boar and large white pigs, Genet. 149 (1998) 1069–1108.

[23] Kramer J., Malek M., Lamont S.J., Association of twelve candidate gene polymorphisms and response to challenge with *Salmonella enteritidis* in poultry, Anim. Genet. 34, (2003) 339–348.

[24] Kruglyak L., Lander E.S., A nonparametric approach for mapping quantitative trait loci, Genet. 139 (1995) 1421–1428.

[25] Lamont S.J., Kaiser M.G., Liu W., Candidate genes for resistance to *Salmonella enteritidis* colonization in chickens as detected in a novel genetic cross, Vet. Immunol. Immunopath. 87 (2002) 423–428.

[26] Lander E.S., Kruglyak L., Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results, Nat. Genet. 11 (1995) 241–247.

[27] Leveque G., Forgetta V., Morroll S., Smith A.L., Bumstead N., Barrow P., Loredo-Osti J.C., Morgan K., Malo D., Allelic variation in TLR4 is linked to susceptibility to *Salmonella enterica* Serovar Typhimurium infection in chickens, Infect. Immun. 71 (2003) 1116–1124.

[28] Liu W., Kaiser M.G., Lamont. S.J., Natural resistance-associated macrophage protein 1 gene polymorphisms and response to vaccine against or challenge with *Salmonella enteritidis* in young chicks, Poult. Sci. 82 (2003) 259–266.

[29] Mariani P., Barrow P.A., Cheng H.H., Groenen M.A.M., Negrini R., Bumstead N., Localization to chicken Chromosome 5 of a novel locus determining salmonellosis resistance, Immunogenet. 53 (2001) 786–791.
[30] Martinez O., Spurious linkage between markers in QTL mapping, Molec. Breed. 2 (1996) 351–358.
[31] Protais J., Colin P., Beaumont C., Guillot J.F., Lantier F., Pardon P., Bennejean G., Line differences in resistance to *Salmonella enteritidis* PT4 infection, Brit. Poult. Sci. 37 (1996) 329–339.
[32] Sadeyen J.R., Trotereau J., Velge P., Marly J., Beaumont C., Barrow P.A., Bumstead N., Lalmanach A.C., Salmonella carrier-state in chicken: comparison of expression of immune response genes between susceptible and resistant animals, Mic. Inf. 6 (2004) 1278–1286.
[33] SAS Institute Inc., SAS/STAT® User's Guide, Version 6, SAS Institute Inc., Cary, 1989.
[34] Satterthwaite F.W., An approximate distribution of estimates of variance components, Biom. Bull. 2 (1946) 110–114.
[35] Schmid M., Nandra I., Guttenbach M., Steinlein C., Hoehn H., Schartl M., Haaf T., Weigend S., Fries R., Buerstedde J.M., Wimmers K., Burt D.W., Smith J., A'Hara S., Law A., Griffin D.K., Bumstead N., Kaufman J., Thomson P.A., Burke T., Groenen M.A.M., Crooijmans R.P.M.A., Vignal A., Fillon V., Morisson M., Pitel F., Tixier-Boichard M., Ladjali-Mohammedi K., Hillel J., Maki-Tanila A., Cheng H.H., Delany M.E., Burnside J., Mizuno S., First report on chicken genes and chromosomes 2000, Cytogenet. Cell Genet. 90 (2000) 169–218.
[36] Seaton G., Haley C.S., Knott S.A., Kearsey M.J., Visscher P.M., QTL Express: mapping quantitative trait loci in simple and complex pedigrees, Bioinformat. 18 (2002) 339–340.
[37] Smith J., Bruley C.K., Paton I.R., Dunn I., Jones C.T., Windsor D., Morrice D.R., Law A.S., Masabanda J., Sazanov A., Waddington D., Fries R., Burt D.W., Differences in gene density on chicken macrochromosomes and microchromosomes, Anim. Genet. 31 (2000) 96–103.
[38] Tilquin P., Coppieters W., Elsen J.M., Lantier F., Moreno C., Baret P.V., Statistical power of QTL mapping methods applied to bacteria counts, Genet. Res. 78 (2001) 303–316.
[39] Tilquin P., Van Keilegom I., Coppieters W., Le Boulengé E., Baret P.V., Nonparametric interval mapping in half-sib designs: use of midranks to account for ties, Genet. Res. 81 (2003) 221–228.
[40] Tuiskula-Haavisto M., Honkatukia M., Vilkki J., de Koning D.J., Schulman N.F., Maki-Tanila A., Mapping of quantitative trait loci affecting quality and production traits in egg layers, Poult. Sci. 81 (2002) 919–927.