Precision Of Distances And Orders Of Microsatellite Markers In A Chicken Genetic Linkage Map Of Chromosome 1 Using Bootstrap Resampling

M.F. Rosário[*], G.R.A. Margarido[†], C. Boschiero[*], A.S.A.M.T. Moura[‡], M.C. Ledur[§], A.A.F. Garcia[†] and L.L. Coutinho[*]

Introduction

Linkage maps have been the source for QTL mapping, gene discovery and BAC and radiation hybrid physical mapping for the assembly of the whole chicken genome sequence. However, several factors can influence the construction of linkage maps and their results: experimental population design, genetic background, population size, number of molecular markers, number of informative meioses, segregation pattern of loci, genotyping errors, number of missing genotypes and gender (Hackett and Broadfoot (2003)). Adopting measures of precision, such as confidence intervals and frequency distributions, based on the bootstrap resampling might be a useful tool for measuring the uncertainty in each map and in their comparisons (Matise *et al.* (2007)). This work reports the precision of distances and orders of microsatellite markers in the EMBRAPA chromosome 1 consensus linkage map from two Brazilian F_2 reciprocal chicken populations using bootstrap resampling.

Material and methods

Experimental populations. The TCTC population was obtained from crosses between seven males from a broiler line (TT) with seven females from a layer line (CC), whereas the CTCT population was obtained in a similar way, but using the reciprocal cross. Details are in Rosário *et al.* (2009). For TCTC seven families (648 F_2) were selected, according to a selective genotyping step (Nones *et al.* (2005)) and for CTCT, four families (356 F_2) were selected, according to the several genotypic parameters.

Genotyping. DNA extraction, PCR reactions and genotyping steps were applied according to Rosário *et al.* (2009). A total of 31 (14) microsatellite markers were used in TCTC (CTCT) populations. For CTCT population, markers were selected based on previous association with body weight at 42 days in the TCTC population (Nones *et al.* (2006)). Primer sequences (forward and reverse) are available at ArkDB (http://www.thearkdb.org/).

[*]USP/ESALQ, Department of Animal Science, Pádua Dias 11, Piracicaba, SP, 13418-900, Brazil, mfrosari@esalq.usp.br, millorfernandes@gmail.com
[†] USP/ESALQ, Department of Genetics, Pádua Dias 11, Piracicaba, SP, 13418-900, Brazil
[‡] UNESP/FMVZ, Fazenda Experimental Lageado, PO 560, Botucatu, SP, 18618-000, Brazil
[§] EMBRAPA Suínos e Aves, BR 153, km 110, Concórdia, SC, 89700-000, Brazil

Linkage map construction. Single maps were constructed for each population separately. Consensus map was also obtained, combining the genotype datasets from both populations, resulting in a total sample size of 1,004 F_2 offspring. We used the CRI-MAP software (http://linkage.rockefeller.edu/soft/crimap/) with TWOPOINT, BUILD, ALL, FLIPS2 options and LOD=3. The number of double and triple recombinations was obtained with the CHROMPIC. These procedures resulted in an averaged map for both sexes.

Precision of marker distances and orders. The procedures were based on the bootstrap with resampling method (Efron and Tibshirani (1993)) only for the Brazilian consensus linkage map. For loci distances, 1,000 independent random samples with replacement were generated using the R software (http://www.R-project.org); for each bootstrap sample a linkage map was constructed using the FIXED function of CRI-MAP, assuming that the order among the microsatellite loci was known previously. Scripts were produced in SHELL language to process the analyses using the CRI-MAP with LOD=3 and the 95% confidence intervals were estimated by the percentile method (Liu (1998)). For loci order, bootstrap samples were generated, and pair-wise recombination fraction estimates were obtained using the TWOPOINT function in the CRI-MAP with LOD=0. Markers were ordered using the seriation algorithm (Buetow and Chakravarti (1987)). Loci frequency distributions were plotted in a graphic. If the estimated order for a linkage group is reliable, the diagonal elements of this matrix will show large frequencies, with points concentrated on the diagonal (Liu (1998)).

Results and discussion

This study was motivated by the difficulty in comparing chicken genetic maps from different populations. Therefore, confidence intervals for distances and frequency distributions for orders of microsatellite markers were estimated. Confidence intervals for SNP positions had been already done for human (Matise *et al.* (2007)) and the bovine (Snelling *et al.* (2005)).

For the TCTC and CTCT populations, on average, 568 and 342 F_2 individuals were genotyped, respectively, showing 3.4 and 3.5 alleles *per* locus and 596 and 353 phase-known informative meioses *a posteriori*. Lengths of linkage maps were 425.1, 231.6 and 433.1 cM for TCTC, CTCT and the consensus, respectively (Figure 1A). Distances between two adjacent loci presented, on average, 13.7, 17.8 and 13.5 cM in TCTC, CTCT and consensus maps, respectively. Only *MCW0145* and *LEI0079* presented inversions in the CTCT map relative to the consensus.

Locus *MCW0020* in the TCTC map and loci *ADL0234*, *MCW0297*, *LEI0174*, *MCW0112*, *ADL0150* and *LEI0079* in the CTCT map fell out the confidence intervals of distances between loci, indicating significant position shifts relative to the consensus map (Figure 1A). Figure 1B shows a concentration of points on and around the diagonal suggest that the orders formerly proposed for the consensus map were partially confirmed. Moreover, another concentration of points out of the diagonal was observed approximately between *MCW0208* and *LEI0169* (loci 0 to 25 in the "markers" axis), revealing a possible inversion between two linkage groups: one between *MCW0208* and *LEI0169*, which could also be in the opposite direction and another between *LEI0107* and *ROS0025*.

The greater similarity between TCTC and the consensus maps, than between CTCT and the consensus maps (Figure 1A) can be due to the higher number of phase known informative meioses *a posteriori*, as well as of markers in TCTC, considering that only 14 loci were genotyped in CTCT, whereas 31 were genotyped in TCTC. In the consensus map a total of 32 loci were used, whereas only 12 loci were simultaneously positioned in the TCTC and CTCT maps, 19 loci were genotyped exclusively in TCTC and one exclusively in CTCT.

The three largest confidence intervals (represented by **I**) were found for distances 2, 24 and 31 (60.9 cM on average) and the three smallest (represented by ▌) were reported for distances 8, 10 and 13 (0.1 cM on average) (Figure 1A). It should be noted that markers flanking distances 2, 10, 13 and 31 were restricted to TCTC, whereas for the distances 8 and 24, *MCW0297* and *LEI0169* were genotyped in both populations. When the intersection of orders between "markers" and "positions" axes were considered in the resampling, 22 loci did not exceed the 10% threshold, that is, in less than 100 resamplings no correspondence between these two axes was found. These loci were located between *ADL0188* and *LEI0107*, except for *MCW0289* and *LEI0146*.

For the TCTC population, Nones *et al.* (2005) had already constructed a 464.1 cM linkage map for GGA1 using 26 microsatellite markers, with 15.0 cM average distance between adjacent loci. In the present study, the TCTC map for GGA1 was reanalyzed using 31 microsatellite markers and showed total length of 491.1 cM, with 13.7 cM average distance between loci. Two inversions were observed comparing the results showed here with those from Nones *et al.* (2005): between *MCW0208* and *MCW0010* and between *ADL0183* and *LEI0106*. According to Matise *et al.* (2007) linkage maps represent resampling trials, and therefore, bias associated to loci position and order estimates could be detected. This is important because, in general, the uncertainty effects associated to the construction of linkage maps is ignored.

Nineteen position inversions were observed in the comparison between the Brazilian (BCM) and the International (ICM) (Schmid *et al.* (2005)) consensus maps and seven between BCM and Genome Sequence (GS) (Hillier *et al.* (2004)). Locus *LEI0138* showed a relevant position shift in BCM compared to ICM; its position in GS, however, was similar to BCM. Twenty-one loci fell out the confidence intervals of distances in the comparison between BCM and ICM, and 13 between BCM and GS. A search for QTLs flanked by *LEI0138* in the Chicken QTL Database (http://www.animalgenome.org/QTLdb/chicken.html) revealed that only Nones *et al.* (2006) mapped QTLs for body weight at 35 days of age (in the linecross analysis) and for liver weight adjusted for body weight at 42 days of age (in the half-sib analysis) using the TCTC population. Therefore, it is suggested that the position of this locus be revised in order to allow new QTL to be mapped.

Acknowledgements

We are grateful to EMBRAPA/PRODETAB, FAPESP (04/02080-2) and CNPq (152500/2007-3 and 480074/2008-9) by financial support and to K. Nones for generation of the genotypic data.

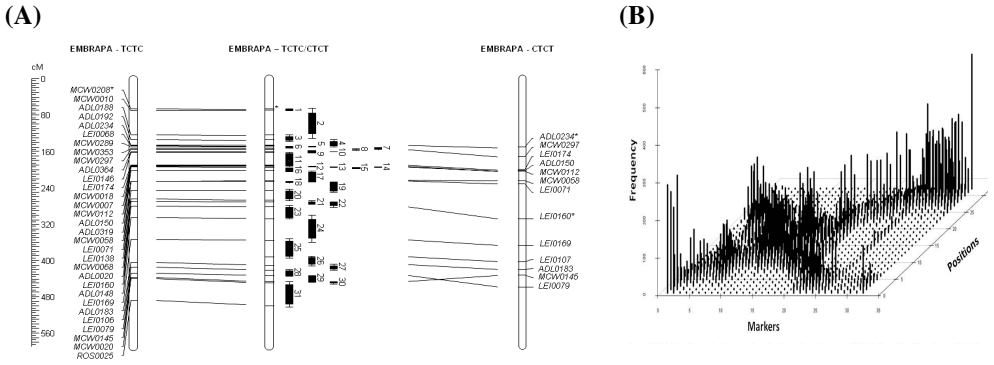

Figure 1: (A) Representation of the single and consensus linkage maps constructed from two Brazilian reciprocal chicken populations. Each distance (numbered from 1 to 31) in EMBRAPA – TCTC/CTCT consensus linkage map has a confidence interval represented by ▌ (lower limit) and I (upper limit), using bootstrap resampling; (B) Frequency distributions between markers and positions using seriation method. Axis x is position of the markers on initial linkage map, axis Y is the frequency on the bootstrap resampling and axis Z is the loci position on each resampling

Conclusion

The chicken genetic linkage map of chromosome 1 with precision measures of microsatellite loci distances and orders, using bootstrap resampling, shown here allows understanding part of the discrepancies found among different studies. The extension of these analyses to the remaining chromosomes may contribute to the comparison among several genetic maps, allowing QTL to be mapped with increased precision regarding their estimates of positions and effects.

References

Buetow K.H., and Chakravarti A. (1987). *Am. J. Hum. Genet.*, 41:180-188.

Efron B., and Tibshirani R.J. (1993). *An Introduction to the Bootstrap.* Chapman & Hall.

Hackett C.A., and Broadfoot L.B. (2003). *Heredity*, 90:33-38.

Hillier L.W., Miller W., Birney E. *et al.* (2004). *Nature*, 432:695-716.

Liu B.H. (1998). *Statistical Genomics: Linkage, Mapping and QTL Analysis.* CRC Press.

Matise T.C., Chen F., Chen W. *et al.* (2007). *Genome Res.*, 17:1783-1786.

Nones K., Ledur M.C., Ruy D.C. *et al.* (2005). *Sci. Agri.*, 62:12-17.

Nones K., Ledur M.C., Ruy D.C. *et al.* (2006). *Anim. Genet.*, 37:95-100.

Rosário M.F., Ledur M.C., Moura A.S.A.M.T. *et al.* (2009). *Sci. Agri.*, 66:150-158.

Schmid M., Nanda I., Hoehn H. *et al.* (2005). *Cytogenet. Genome Res.*, 109:415-479.

Snelling W.M., Casas E., Stone R.T. *et al.* (2005). *BMC Genomics*, 6:78.