

Extent of linkage disequilibrium in chicken

J. Aerts^{a,b} H.J. Megens^a T. Veenendaal^a I. Ovcharenko^c R. Crooijmans^a
L. Gordon^c L. Stubbs^c M. Groenen^a

^aAnimal Breeding and Genetics Group, Wageningen University, Wageningen (The Netherlands)

^bBioinformatics Department, Roslin Institute, Roslin, Scotland (UK)

^cEEBI Division and Genome Biology Division, Lawrence Livermore National Laboratory, Livermore, CA (USA)

Manuscript received 11 July 2006; accepted in revised form for publication by J. Smith, 25 October 2006.

Abstract. Many of the economically important traits in chicken are multifactorial and governed by multiple genes located at different quantitative trait loci (QTLs). The optimal marker density to identify these QTLs in linkage and association studies is largely determined by the extent of linkage disequilibrium (LD) around them. In this study, we investigated the extent of LD on two chromosomes in a white layer and two broiler chicken breeds. Pairwise levels of LD were calculated for 33 and 36 markers on chromosomes 10 and 28, respectively. We found that useful LD (i.e. an r^2 value higher than 0.3) in Nutreco chicken breed E5 (inbred) can extend to around 1 cM on chromosomes 10 and

28, although in a second region on chromosome 28 it extends to about 2.5 cM. The extent in breed Nutreco E3 (outbred) was very short in chromosome 10 (15 kb) but very much larger on chromosome 28, particularly in one region of depressed heterozygosity. The layer breed E2 (inbred) showed an extent of useful LD up to 4 cM on chromosome 10; the extent on chromosome 28 could not be assessed due to an erratic pattern of LD on that chromosome, although in one region LD appears to be in the order of 0.8 cM. This indicates that there may be very large differences in patterns of LD between different chicken breeds and different genomic regions.

Copyright © 2007 S. Karger AG, Basel

In the past few years, a wealth of genomic information has become available for chicken (International Chicken Genome Sequencing Consortium, 2004; International Chicken Polymorphism Map Consortium, 2004; Wallis et al., 2004; Schmid et al., 2005). Although these resources are indispensable tools for further genetic and genomic research in chicken, additional information is required to optimize their use in the ultimate aim of genomic research, which is to locate genes of interest and to discern their mode of operation.

An important piece of that information is the level of population-wide linkage disequilibrium (LD) in the region. Linkage disequilibrium refers to the non-random association of alleles at different loci and quantifies the level of informativeness between different markers. As such, it dictates marker density in marker-based studies (Goldstein and Weale, 2001; Weiss and Clark, 2002). A long extent of LD may be beneficiary for an initial survey on the genomic location of a quantitative trait locus (QTL) because only a limited number of markers has to be tested. On the other hand, a long extent of LD may be problematic for fine-mapping of a QTL, because association may be detected at a distance far from the causative locus.

The average extent of LD (i.e. the dependence of LD on the separation distance between markers) has been investigated for several organisms, including human (Reich et al., 2001) and some domesticated animals (Farnir et al., 2000; Nsengimana et al., 2004; Sutter et al., 2004; Jungerius et al.,

Request reprints from H.J. Megens

Animal Breeding and Genetics Group, Wageningen University
PO Box 338, 6700 AH, Wageningen (The Netherlands)
telephone: 0317 482 335; fax: 0317 483 929
e-mail: hendrik-jan.megens@wur.nl

Table 1. Minor allele frequencies (MAF; average \pm standard deviation) and number of polymorphic SNPs per breed and region. Region 1 on GGA28 is bordered by markers 1008152 and 1008995; region 2 on GGA28 is bordered by markers 1009135 and 1011434 (see Results and Appendix A).

		Breed					
		E5		E3		E2	
		MAF	Number of SNPs polymorphic	MAF	Number of SNPs polymorphic	MAF	Number of SNPs polymorphic
Region	GGA10	0.27 \pm 0.13	27	0.20 \pm 0.12	28	0.19 \pm 0.10	14
	GGA28 (region 1)	0.33 \pm 0.09	3	0.28 \pm 0.08	9	0.05 ^a	1
	GGA28 (region 2)	0.27 \pm 0.12	25	0.04 \pm 0.005	6	0.13 \pm 0.06	14

^a Only one marker in this region. Therefore standard deviation could not be calculated.

2005). Linkage disequilibrium studies in chicken have so far been sparse; an example where microsatellites were used is provided by Heifetz et al. (2005). There is clearly a need for better assessment of LD in chicken across populations and using markers that allow a spacing of loci on the order of kilobases (i.e. SNPs), particularly since chicken appears to be a species with high sequence variability (International Chicken Polymorphism Map Consortium, 2004).

In this paper, we present results relating to SNP-based LD in chicken. We sampled two chromosomes, one large (GGA10) and one small microchromosome (GGA28), which have a three-fold difference in recombination rate (Groenen et al., 2000). We compared three populations with different breed histories constituting both long separated breeds (i.e. a white layer versus broilers) and recently separated breeds (i.e. two broiler breeds).

Material and methods

Populations

The extent of linkage disequilibrium was investigated in three commercial chicken breeds, one white layer (designated E2) and two broilers. Both broiler breeds (designated E5 and E3) were selected for meat productivity. The number of nuclear families (father – mother – one offspring) was 37 and 43 for breeds E5 and E3, respectively. Families were selected to be as genetically different as possible. Breed E5 is a broiler-dam line and is considered to be a 'closed' breed (i.e. no external animals have been added for more than 30 years); breed E3, a broiler-sire line, is an 'open' breed, with an expected rate of inbreeding half that of breed E5 (Addie Vereijken, Nutreco, personal communication). For both breeds, approximately 700 dams and 50 to 60 sires are selected per generation to maintain it. Breed E5 has been selected towards reproduction, while breed E3 was selected towards growth, feed conversion and breastmeat. Breed E2 is a layer breed and is considered closed and relatively inbred (Jeroen Visscher, HPB, personal communication). An extended family structure was used using 17 males and 62 females. Families that showed Mendelian errors were discarded for further analysis.

Genomic regions

Two regions were investigated for extent of linkage disequilibrium. The first region encompasses 5.4 Mb on chicken chromosome 10 (GGA10) between nucleotide positions 6,313,704 and 11,716,327 (sequence build WASHUC1; International Chicken Genome Sequencing Consortium, 2004). This corresponds to 30.79 cM.

The second region consists of 1.4 Mb on chicken chromosome 28 (GGA28). The underlying sequence used for that chromosome is based on a BAC-based sequencing effort performed at the Lawrence Livermore National Laboratory (LLNL, California, USA; Gordon et al., in preparation) rather than the sequence build from the International Chicken Genome Sequencing Consortium (2004), because the former is believed to be of higher quality. This region is located between positions 2,213,089 and 3,699,264 on the LLNL sequence, which corresponds to the region between 2,964,842 and 4,083,891 on the WGS version. This corresponds to 19.8 cM.

Polymorphism positions are available as DAS files (Dowell et al., 2001) that can be loaded into the Ensembl genome browser for investigation of the genomic environment (Hubbard et al., 2002).

SNPs

Three different SNP sets were used, originating from three different resources. For GGA10, the list of SNPs contains both polymorphisms made available by the International Chicken Polymorphism Map Consortium (2004; names starting with 'snp.:' set 1) and SNPs discovered by resequencing of end sequences from BAC clones of the Wageningen BAC library (Crooijmans et al., 2002; names starting with 'SNP_'; set 2) that are known to be located within that region. For GGA28, SNPs were identified by comparing DNA sequences of BACs from two different chicken strains: a Red Jungle Fowl and a White Leghorn chicken (Gordon et al., in preparation). These SNPs formed SNP-set 3.

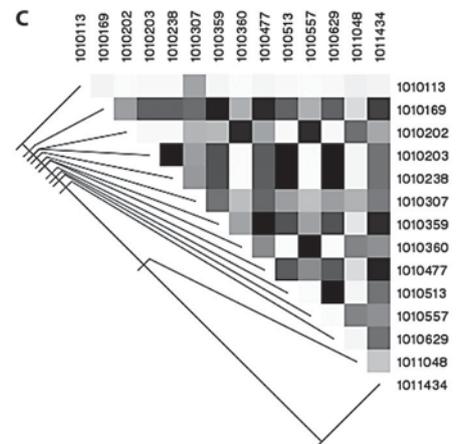
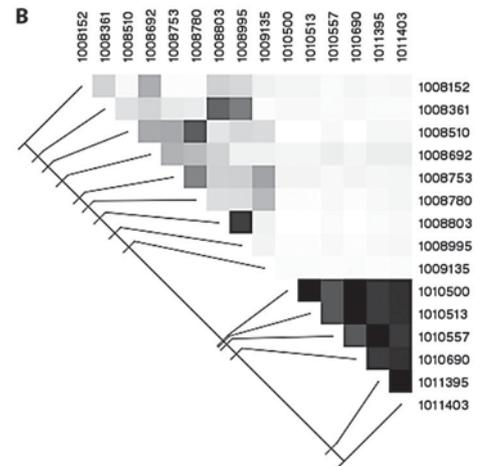
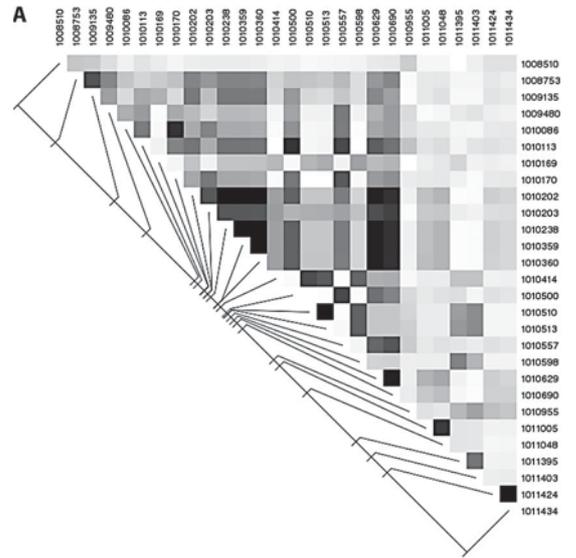
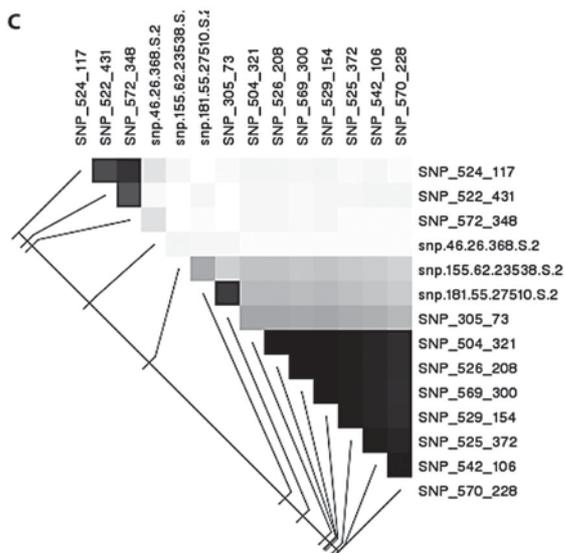
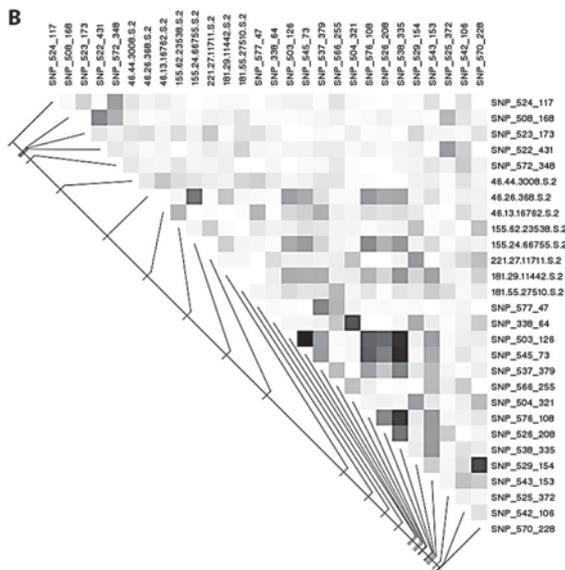
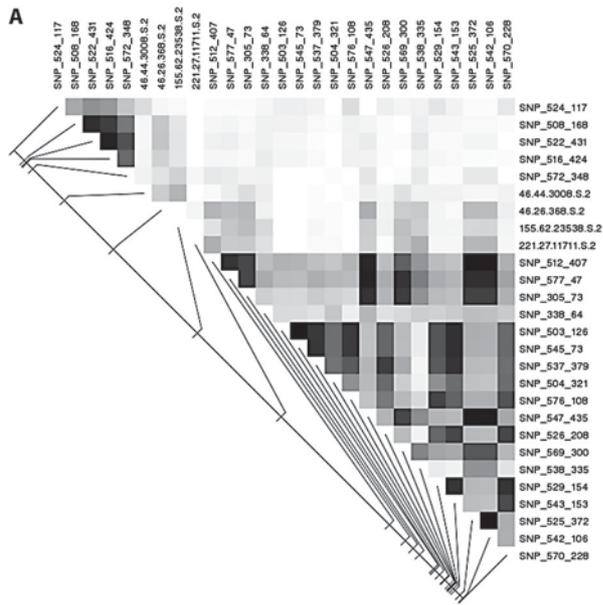
An overview of minor allele frequencies and number of usable SNPs per region per breed is presented in Table 1. A complete list of SNPs that complied to these criteria is presented as Appendix A (Supplementary material, www.karger.com/doi/10.1159/000103196).

Genotyping

Multiplex reactions of up to six PCRs with the same annealing temperature were performed in 12 μ l and contained 30 ng template DNA, 6 μ l AccuPrime (Invitrogen) SuperMix II or ABgene ReadyMix and 0.2 μ M of each primer. PCR conditions were 94°C for 10 min, 41 cycles of 94°C for 30 s, annealing temperature for 30 s and 68°C for 3 min, followed by 68°C for 2 min. PCR products were then pooled based on single base extension (SBE) primer length into six super-pools of up to 17 assays.

Genotyping was performed using the standard SnaPshot Multiplex Kit (Applied Biosystems) with the following modifications. For the *Exo1* treatment, 0.4 μ l *Exo1* (10 U/ μ l) was used instead of 0.2 μ l. For the SnaPshot SBE reaction, 4 μ l Magic Dye (RedRabbit) and 1 μ l SnaPshot Ready Reaction Mix were used. This reaction involved 40 cycles.

Genotype detection was achieved using the ABI Prism 3100 Genetic Analyzer (Applied Biosystems). The sample preparation scheme was modified, using 2 μ l SnaPshot product, 8 μ l Hi-Di formamide and 0.25 μ l GeneScan-120 LIZ size standard. Genotypes were scored using Genemapper v3.0 (Applied Biosystems).



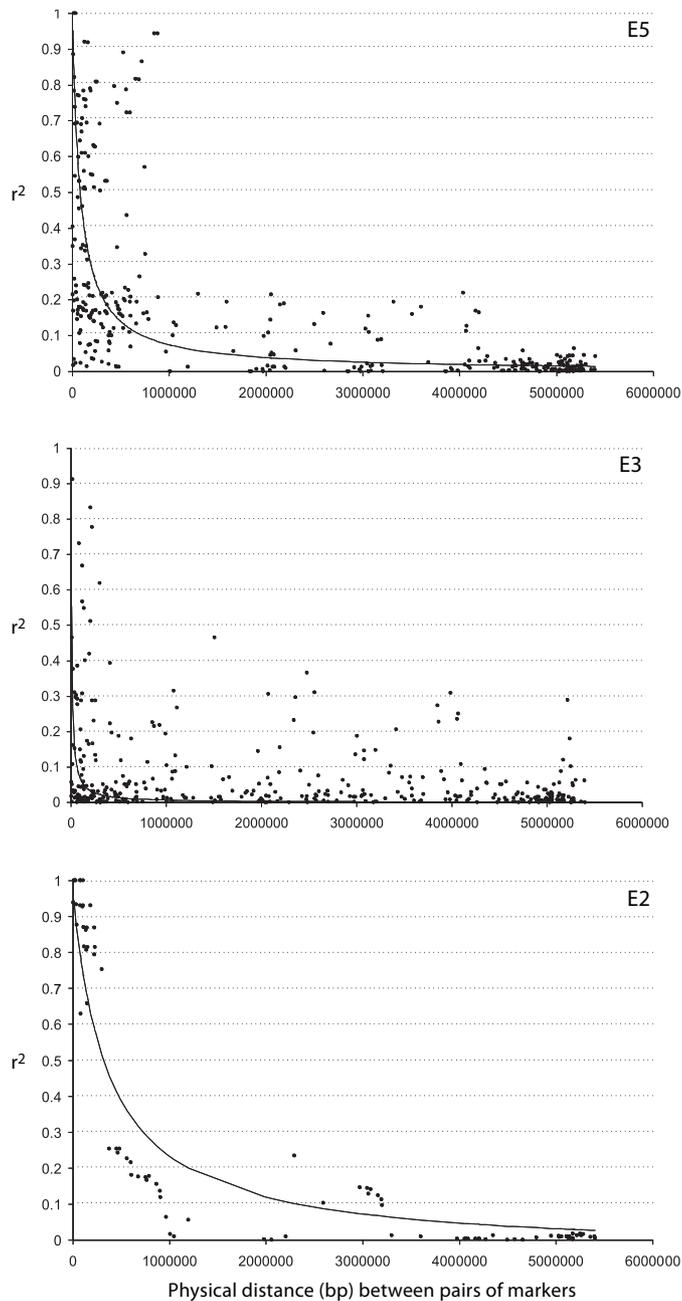


Fig. 3. Extent of LD on GGA10 for breeds E5, E3 and E2. Dots represent the r^2 value between two markers at the given inter-marker distance. The line represents the expected value based on the Sved (1971) equation.

Fig. 1. Pairwise LD on GGA10 for breeds E5 (A), E3 (B) and E2 (C). Cells represent the r^2 value for each pairwise comparison, ranging from 'no LD' (white) to 'full LD' (black). Relative marker locations are shown on the line parallel to the diagonal of the matrix.

Fig. 2. Pairwise LD on GGA28 for breeds E5 (A), E3 (B) and E2 (C). Cells represent the r^2 value for each pairwise comparison, ranging from 'no LD' (white) to 'full LD' (black). Relative marker locations are shown on the line parallel to the diagonal of the matrix.

Calculation of linkage disequilibrium

Linkage disequilibrium parameters were calculated using the haploxt program (Abecasis and Cookson, 2000) on the parent haplotypes as produced by simwalk2 (Sobel and Lange, 1996). Only SNPs that complied to the following criteria were used in further analysis: (1) the minor allele frequency had to be higher than 0.025 and (2) at least half of the genotypes could be assessed. Markers were said to be in useful LD if the pairwise r^2 statistic was higher than the threshold of 0.3, as described in Ardlie et al. (2002).

Presentation of results

Two aspects of the results are presented: (1) pairwise LD among all pairs of markers (Figs. 1 and 2), and (2) relation of LD to distance separating the markers making up the pair (Figs. 3 and 4). Pairwise LD is shown as a heatmap using custom perl and ruby scripts, where cell shading reflects the level of LD for each pair of markers. Plots representing the relation of LD to marker separation were fitted with a curve based on the well known equation of Sved (1971; c.f. Heifetz et al., 2005). The extent of LD was considered to be the distance where the moving average of LD drops below an r^2 value of 0.3 (Abecasis et al., 2001) based on expected values of r^2 using the Sved equation.

Results

SNP rate

Of the markers on chromosome 10 from set 2, all were polymorphic in either breed E5, E3 or E2. In contrast, only nine out of seventeen SNPs discovered by the International Chicken Polymorphism Map Consortium (set 1) could be used as markers in this study. Similarly of the 72 putative SNPs in the region on chromosome 28 (set 3), 47 showed polymorphism in at least one of the breeds that were studied. Only 36 of the latter complied to our criteria for minimal minor allele frequency for at least one breed.

Table 1 displays the minor allele frequencies (MAF) for all markers within the different breeds. A full list of markers and minor allele frequencies is available in Appendix A. For breed E3, chromosome 28 could be clearly divided into two subregions based on MAF: in the first subregion (markers 1008152 through 1008995), minor allele frequencies were of the same order of magnitude as those for breed E5; in the second subregion (markers 1009135 through 1011434) the markers tested were either not polymorphic, or had a very low minor allele frequency (0.04 ± 0.005).

Linkage disequilibrium

The r^2 pairwise LD statistics between markers for each chromosome and breed are represented in Figs. 1 and 2 for the loci on chromosomes 10 and 28, respectively. Pairwise LD values for chromosome 10 in breed E2 (Fig. 1C) illustrate the relation between physical distance and linkage disequilibrium: high pairwise LD exists between markers that lie close to each other, while almost no LD exists over large distances. Pairwise LD for chromosome 10 in breed E5 (Fig. 1A) also shows higher LD between markers that are located close to each other, but these values are generally much lower than those in breed E2. By contrast, E3 displays much lower degrees of LD even between markers fairly close together.

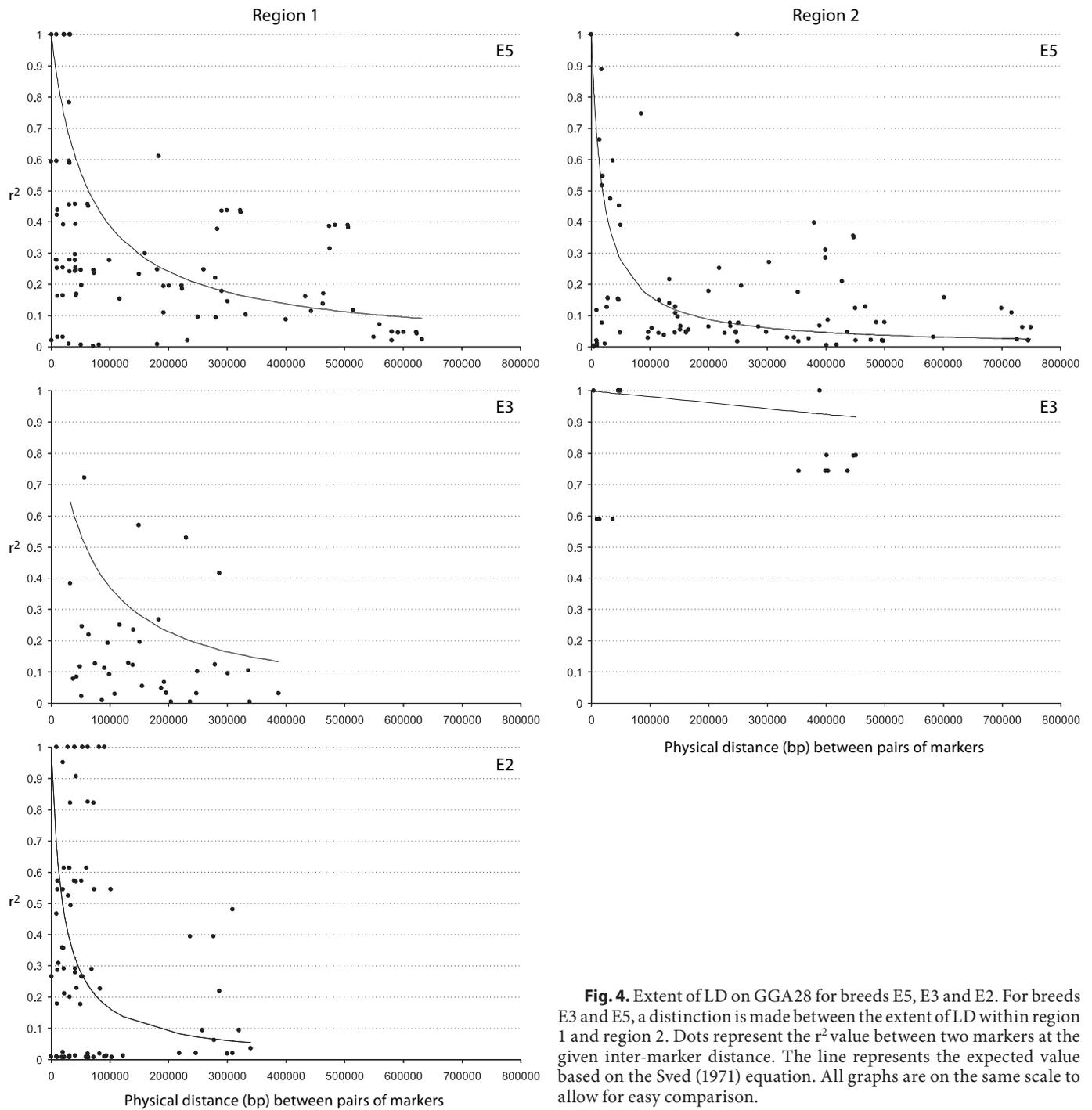


Fig. 4. Extent of LD on GGA28 for breeds E5, E3 and E2. For breeds E3 and E5, a distinction is made between the extent of LD within region 1 and region 2. Dots represent the r^2 value between two markers at the given inter-marker distance. The line represents the expected value based on the Sved (1971) equation. All graphs are on the same scale to allow for easy comparison.

For chromosome 28, a relatively high amount of linkage disequilibrium can be found between markers 1008753 and 1010690 within breed E5 (pairwise $r^2 = 0.31 \pm 0.26$; Fig. 2A). In contrast, almost no LD exists between any marker 5' from 1010629 with any marker 3' from 1010690, even though these markers are only separated by 17.6 kb. For breed E3, the two subregions that are distinguished by their difference in minor allele frequencies (see above) show a significant difference in average values for LD (Fig. 2B): for markers

1008152 through 1008995, only minor linkage disequilibrium could be detected; in contrast, markers 1009135 through 1011403 showed very high values. Breed E2, in contrast to the situation in chromosome 10, displays very erratic LD within the region in chromosome 28 (Fig. 2C). Pairs of SNPs close together and far apart can have either high or low LD. Interestingly, 1010113 displays low LD with almost all other markers, while 1011434 shows substantial LD with all other SNPs, even though it is far apart from all of them.

Extent of linkage disequilibrium

Figures 3 and 4 display the pairwise LD between markers as a function of pairwise physical distance. In the region on GGA10, the extent of useful LD for breed E5 can be derived to be around 200 kb or 1.2 cM. The average extent of useful LD for breed E3 was estimated to be around 15 kb. For breed E2, Fig. 3 shows a clear decrease of LD with increasing distance between markers. The extent of useful LD can be inferred as being 700 kb or about 4 cM.

For chromosome 28, a distinction was made between region 1 (markers 1008152 through 1008995) and region 2 (markers 1009135 through 1011403) based on the difference in average minor allele frequency for the SNPs within breed E3 (see Table 1 and Appendix A). For breed E5, the extent of useful LD was found to be 150 kb or 2.5 cM for region 1 and only 50 kb or 0.8 cM for region 2. For breed E3, the pairwise LD between markers in region 1 was similar compared to E5 (130 kb or 2.2 cM). In contrast, pairwise LD was very high in region 2 and extent of useful LD is much longer than the physical size of the whole region (823 kb). The extent of LD for breed E2 in region 1 was estimated at 50 kb (0.8 cM), but could not be determined for region 2, because for this region only one polymorphic marker was genotyped.

Discussion

SNP rate

Although all markers that were discovered by resequencing the BAC ends of the Wageningen BAC library (i.e. set 2) were polymorphic in either breed E5, E3 or E2, this was not the case for those discovered by the International Chicken Polymorphism Map Consortium (2004) or by comparing Red Jungle Fowl and White Leghorn chromosomes 28 (i.e. sets 1 and 3). This can probably to a large extent be attributed to the fact that both latter batches of SNPs were discovered by comparing one chromosome each from two populations other than the populations eventually genotyped for assessment of LD. In contrast, the SNPs from set 2 were discovered by resequencing of samples from the same populations that were eventually used for genotyping. Nevertheless, a higher success rate was expected for sets 1 and 3, because a survey of 125 markers on nine divergent breeds performed by the International Chicken Polymorphism Map Consortium (2004) indicated that, on average, an inter-breed polymorphism has about 70% chance of being polymorphic within a breed (data not shown). Our results clearly showed a lower frequency of occurrence of these SNPs within the breeds tested.

Breed E2 showed the lowest degree of polymorphism. This is in line with previous biodiversity studies (e.g. Hillel et al., 2003) that showed that (commercial) white layer (leghorn derived) breeds are much less diverse than broiler breeds.

The low degree of polymorphism in the white layer appears throughout the genomic regions studied here. By contrast, the E3 breed was in general polymorphic for most SNPs, but interestingly, the markers on chromosome 28 be-

tween 1009135 and 1011434 were either not polymorphic in this breed, or had very low minor allele frequencies. This can be the result of selection by the breeding company for a QTL that influences one of the breeding goals of the E3 sire-line that are not the primary breeding goals for the E5 dam-line, i.e. growth, feed conversion or breastmeat. This could lead to a selective sweep lowering the heterozygosity at that locus (Innan and Kim, 2004; Kim and Nielsen, 2004).

This region of low heterozygosity was not detected in breed E5. Indeed, different selection goals were set out for each breed. The putative selective sweep in breed E3 (selected for feed conversion and breastmeat) might be connected to a QTL for fat that is found on chromosome 28 (Ikeobi et al., 2002). Closer investigation of this region showed that it contains 58 genes, a few being excellent candidates for influencing key production traits such as an insulin receptor.

It is of interest to note that in the region between 1010500 and 1011434 the number of SNPs that are recorded in dbSNP is markedly depressed compared to other regions on chromosome 28 (based on a survey of dbSNP, results not shown). As the SNP discovery program (International Chicken Polymorphism Map Consortium, 2004) was based on three chickens, two originating from commercial breeds (a broiler and a white layer), this pattern of depressed SNP rate could well represent the footprint of selection. However, it is equally possible that it has resulted from fundamental properties of this genomic region. Further investigations are needed to distinguish between various possibilities.

Linkage disequilibrium statistic

Although the principle of linkage disequilibrium is fairly simple (i.e. the non-random segregation of markers in close proximity), the complex interplay between all confounding factors such as population subdivision, bottlenecks and expansions, is not yet completely understood and makes interpretation of LD results not always straightforward. As a result, many different statistics have been developed to characterize the amount of linkage disequilibrium between any two markers of which Lewontin's D' (Lewontin, 1988) and r^2 (Pritchard and Prezowski, 2001) are widely used. Both range from 0 (no LD) to 1 (full LD), but differ in the interpretation of the intermediate values. Intermediate values for D' are not clearly interpretable and are known to be biased upwards (Ardlie et al., 2002; Ke et al., 2004). In addition, D' is affected by the number of animals used (Weiss and Clark, 2002). In contrast, intermediate values for r^2 give an indication of the power to detect association: to have the same power to detect association between a disease and a marker locus, the sample size must be increased by $1/r^2$ when compared with sample size for detecting the effect of the susceptibility locus itself. Therefore, 'useful LD' is often defined as an r^2 value higher than 0.3, which indicates that the sample size has to be increased 3-fold (Ardlie et al., 2002; Zondervan and Cardon, 2004). If, for example, 1,000 individuals would be needed for an association study given a 'perfect LD model' ($r^2 = 1.0$), $1,000/0.3$ individuals would

be needed if the r^2 statistic has a value of 0.3 in the region under study.

Even though the r^2 statistic gives a good impression of the level of LD between two markers, SNP discovery strategy and demographic history of the population can influence the actual values for the statistic. As a result, two markers that are very close together can exhibit a low level of LD, while markers that are very distant can show a higher than expected level of LD. It is known that linkage disequilibrium between SNPs with a low minor allele frequency is biased upwards (see for example Gaut and Long, 2003). In part this can be explained by statistical properties of the LD statistics (Dunning et al., 2000), but may also have a real meaning because low frequency SNPs have a higher probability of having arisen recently (Nordborg and Tavaré, 2002). A new SNP is in complete LD with all other loci, and the more recent the SNP the less time LD will have had the chance to break down. This is exemplified by the high r^2 values in the region between markers 1008995 and 1009135 on chicken chromosome 28 for breed E3 – which is probably a result of the selection history of the population (see above). Similarly, marker 1011343 shows very high LD with almost all other markers, regardless of distance. This marker has a minor allele frequency of 0.17, which is not low. However, marker frequencies could have been increased rapidly due to selection or drift.

Extent of LD

A common approach to summarize the distribution of LD in a genomic region is to plot r^2 against the physical distance between markers (Ke et al., 2004; see Figs. 3 and 4). These plots illustrate the rate at which linkage disequilibrium decays over distance in that region. To quantify the extent of useful LD, the distance was assessed at which the expected value of r^2 , based on the Sved equation (Sved, 1971) dropped below a value of 0.3. The choice for a cutoff value of 0.3 is basically arbitrary. Choosing a higher cutoff value (e.g. 0.5) would mean that the sample size only has to be increased two-fold instead of three-fold for association studies. However, markers would have to be closer to each other to be considered in useful LD, resulting in more markers that would have to be tested ('fewer animals, more markers'). On the other hand, a lower cutoff value for r^2 (e.g. 0.2) would mean that markers further apart would be considered to be in useful LD, but more animals would have to be tested in an association study ('more animals, fewer markers').

Comparing the extent of LD within the breeds for the different chromosomes, we found that the extent of LD for breed E5 is about four times longer on chromosome 10 compared to chromosome 28 in terms of physical distance. The recombination rate on chromosome 28 (17.7 cM/Mb) is about three times higher than that on chromosome 10 (5.7 cM/Mb), which is similar to this observed difference. In region 1, extent of LD is similar to chromosome 10 in terms of physical distance, but considering the recombination rate it is actually larger. This discrepancy might be due to population history, selection, or other factors and requires fur-

ther investigation to explore. For breed E2, the level of LD decreases clearly with inter-marker distance on chromosome 10, and the extent of useful LD can be estimated to be approximately 700 kb. On chromosome 28, the extent of useful LD is much lower (region 1). For region 2, however, the extent of useful LD can not meaningfully be deduced from our current data. The high level of LD between markers that are at the extremes of the whole region under consideration might indicate that the extent of useful LD is longer than 823 kb; however, that would be contradicted by the erratic levels of LD in the region.

Comparing the extent of LD on chromosome 10 between the breeds, it is shown in this study that the extent of LD for breeds E5 and E2 is considerably longer than that for E3. This was expected, as the E5 and particularly E2 breeds are considered relatively inbred (Jeroen Visscher, HPB, personal communication) and white layers are generally less genetically variable compared to broilers (Hillel et al., 2003). The E3 breed is more 'open' than the E5 breed, leading to shorter regions of identity-by-descent. The long extent of LD in region 2 on chromosome 28 for breed E3 follows from the low heterozygosity and hence upwards biased pairwise LD.

Investigations on LD in other livestock species such as cattle (Farnir et al., 2000; Tenesa et al., 2003) and pig (Nsenigimana et al., 2004; Jungerius et al., 2005) have revealed that considerable levels of LD can exist to over more than 10 cM. This has led to the assumption that haplotype mapping (International HapMap Consortium, 2005) may be rather straightforward in livestock breeds (Andersson and Georges, 2004). Unfortunately, comparison of our results with the LD studies in other livestock species is not easy. First, different studies use different LD statistics (e.g. D' or r^2) that behave differently, with r^2 typically being lower than D' for any chromosomal distance (Weiss and Clark, 2002). Indeed, recalculating LD using D' instead of r^2 showed that the extent of useful LD based on D' is larger than the region under investigation for all three breeds and regions (data not shown). Second, as population structure plays an important role in linkage disequilibrium, results from studies in human are not easily comparable with results from studies in highly structured populations as for example dog breeds (Sutter et al., 2004). And as breed management greatly differs with many other species, the population structure of these chicken breeds is significantly different from that in many cattle breeds (with a relatively small pool of very important sires) and dog breeds (with inbred subpopulations). Third, the choice of markers influences the extent of LD. For example, data using SNPs for studying LD are not easily merged with data using microsatellites. Microsatellite markers have a much higher mutation rate than SNPs (Vignal et al., 2002), significantly influencing the extent of LD around the markers. Therefore, these results can not be compared to data obtained from SNP genotyping surveys.

One important message from this paper is that in chicken, unlike perhaps in human but possibly like in other domesticated species (e.g. dog; Lindblad-Toh et al., 2005), there may be substantial differences in patterns of LD. These

differences are congruent with breed history (chromosome 10). However, we also observe substantial differences between genomic regions (chromosome 28). As a consequence, it will be very difficult to translate current results to general statements concerning LD in chicken. However, from our results it appears that for some breeds and genomic regions haplotype mapping may require too dense a marker sampling to be currently economical, while other breeds or genomic regions tentatively appear to display levels of LD that could allow for haplotype mapping using relatively low marker densities.

Conclusion

We showed that extensive linkage disequilibrium does exist around markers in the chicken genome. As expected,

linkage disequilibrium increases with degree of inbreeding and decreases with variability, which is most significantly displayed by the more 'closed' of the two broiler breeds in showing the highest degree of LD in chromosome 10. The finding of an anomalous region on chromosome 28 for the broiler E3 (and possibly E2) may be linked to selection, but this requires further investigation.

Acknowledgements

We wish to thank Addie Vereijken and Annemieke Rattink (Nutreco BV, The Netherlands) and Jeroen Visscher (Hendrix Poultry Breeding, HPB, Boxmeer, The Netherlands) for providing useful information on the chicken breeds. We also thank Johan Van Arendonk for his helpful comments.

References

- Abecasis GR, Cookson WO: GOLD – graphical overview of linkage disequilibrium. *Bioinformatics* 16:182–183 (2000).
- Abecasis GR, Noguchi E, Heinzmann A, Traherne JA, Bhattacharyya S, et al: Extent and distribution of linkage disequilibrium in three genomic regions. *Am J Hum Genet* 68:191–197 (2001).
- Andersson L, Georges M: Domestic-animal genomics: deciphering the genetics of complex traits. *Nat Rev Genet* 5:202–212 (2004).
- Ardlie KG, Kruglyak L, Seielstad M: Patterns of linkage disequilibrium in the human genome. *Nat Rev Genet* 3:299–309 (2002).
- Crooijmans RPMA, Dijkhof RJM, Veenendaal T, van der Poel JJ, Nicholls RD, et al: The gene orders on human chromosome 15 and chicken chromosome 10 reveal multiple inter- and intrachromosomal rearrangements. *Mol Biol Evol* 18:2102–2109 (2002).
- Dowell RD, Jokerst RM, Day A, Eddy SR, Stein LD: The distributed annotation system. *BMC Bioinformatics* 2:7 (2001).
- Dunning AM, Durocher F, Healey CS, Teare MD, McBride SE, et al: The extent of linkage disequilibrium in four populations with distinct demographic histories. *Am J Hum Genet* 67:1544–1554 (2000).
- Farnir F, Coppeters W, Arranz JJ, Berzi P, Cambisano N, et al: Extensive genome-wide linkage disequilibrium in cattle. *Genome Res* 10:220–227 (2000).
- Gaut BS, Long AD: The lowdown on linkage disequilibrium. *Plant Cell* 15:1502–1506 (2003).
- Goldstein DB, Weale ME: Population genomics: linkage disequilibrium holds the key. *Curr Biol* 11:R576–R579 (2001).
- Groenen MAM, Cheng HH, Bumstead N, Benkel BF, Briles WE, et al: A consensus linkage map of the chicken genome. *Genome Res* 10:137–147 (2000).
- Heifetz EM, Fulton JE, O'Sullivan N, Zhao H, Dekkers JCM, Soller M: Extent and consistency across generations of linkage disequilibrium in commercial layer chicken breeding populations. *Genetics* 171:1173–1181 (2005).
- Hillel J, Groenen MAM, Tixier-Boichard M, Korol AB, David L, et al: Biodiversity of 52 chicken populations assessed by microsatellite typing of DNA pools. *Genet Sel Evol* 35:533–557 (2003).
- Hubbard T, Barker D, Birney E, Cameron G, Chen Y, et al: The Ensembl genome database project. *Nucleic Acids Res* 30:38–41 (2002).
- Ikeobi CON, Woolliams JA, Morrice DR, Law A, Windsor D, et al: Quantitative trait loci affecting fatness in the chicken. *Anim Genet* 33:428–435 (2002).
- Innan H, Kim Y: Pattern of polymorphism after strong artificial selection in a domestication event. *Proc Natl Acad Sci USA* 101:10667–10672 (2004).
- International Chicken Genome Sequencing Consortium: Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432:695–716 (2004).
- International Chicken Polymorphism Map Consortium: A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* 432:717–722 (2004).
- International HapMap Consortium: A haplotype map of the human genome. *Nature* 437:1299–1320 (2005).
- Jungerius B, Gu J, Crooijmans R, van der Poel J, van Oost B, et al: Estimation of the extent of linkage disequilibrium in seven regions of the porcine genome. *Anim Biotech* 16:41–54 (2005).
- Ke X, Hunt S, Tapper W, Lawrence R, Stavrides G, et al: The impact of SNP density on fine-scale patterns of linkage disequilibrium. *Hum Mol Genet* 13:577–588 (2004).
- Kim Y, Nielsen R: Linkage disequilibrium as a signature of selective sweeps. *Genetics* 167:1513–1524 (2004).
- Lewontin RC: On measures of gametic disequilibrium. *Genetics* 120:849–852 (1988).
- Lindblad-Toh K, Wade CM, Mikkelsen TS, Karlsson EK, Jaffe DB, et al: Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* 438:803–819 (2005).
- Nordborg N, Tavaré S: Linkage disequilibrium: what history has to tell us. *Trends Genet* 18:83–90 (2002).
- Nsengimana J, Baret P, Haley CS, Visscher PM: Linkage disequilibrium in the domesticated pig. *Genetics* 166:1395–1404 (2004).
- Pritchard JK, Przeworski M: Linkage disequilibrium in humans: models and data. *Am J Hum Genet* 69:1–14 (2001).
- Reich DE, Cargill M, Bolk S, Ireland J, Sabeti PC, et al: Linkage disequilibrium in the human genome. *Nature* 411:199–204 (2001).
- Schmid M, Nanda I, Hoehn H, Scharl M, Haaf T, et al: Second report on chicken genes and chromosomes 2005. *Cytogenet Genome Res* 109:415–479 (2005).
- Sobel E, Lange K: Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *Am J Hum Genet* 58:1323–1337 (1996).
- Sutter NB, Eberle MA, Parker HG, Pullar BJ, Kirkness EF, et al: Extensive and breed-specific linkage disequilibrium in *Canis familiaris*. *Genome Res* 14:2388–2396 (2004).
- Sved JA: Linkage disequilibrium and homozygosity of chromosome segments in finite populations. *Theor Pop Biol* 2:125–141 (1971).
- Tenesa A, Knott SA, Ward D, Smith D, Williams JL, Visscher PM: Estimation of linkage disequilibrium in a sample of the United Kingdom dairy cattle population using unphased genotypes. *J Anim Sci* 81:617–623 (2003).
- Vignal A, Milan D, SanCristobal M, Eggen A: A review on SNP and other types of molecular markers and their use in animal genetics. *Genet Sel Evol* 34:275–305 (2002).
- Wallis JW, Aerts J, Groenen MAM, Crooijmans RPMA, Layman D, et al: A physical map of the chicken genome. *Nature* 432:761–776 (2004).
- Weiss KM, Clark AG: Linkage disequilibrium and the mapping of complex traits. *Trends Genet* 18:19–24 (2002).
- Zondervan KR, Cardon LR: The complex interplay among factors that influence allelic association. *Nat Rev Genet* 5:89–100 (2004).