Proc. Assoc. Advmt. Anim. Breed. Genet. 20:364-367

THE EXTENT AND DISTRIBUTION OF LINKAGE DISEQUILIBRIUM IN EXTENSIVELY RAISED CHICKEN POPULATIONS OF SOUTHERN AFRICA

K.S. Khanyile[1,2], E. F. Dzomba[2] and F. C. Muchadeyi[1*]

[1]Biotechnology Platform, Agricultural Research Council, P Bag X5, Onderstepoort, 0110, South Africa
[2]University of KwaZulu-Natal, Discipline Genetics, School of Life Sciences, Private Bag X01, Scottsville 3209, Pietermaritzburg, South Africa
Correspondence: MuchadeyiF@arc.agric.za

SUMMARY

The amount of linkage disequilibrium (LD) is an important source of information about historical events of recombination and allows inferences about genetic diversity and genomic regions that have undergone selection. Linkage disequilibrium is equally important in studying effective population size and rate of inbreeding particularly in extensively raised and wild animal populations where pedigree records are scarce. The objective of this study was to investigate LD in village chicken populations of Southern Africa. These chickens are raised under scavenging systems of production characterized by uncontrolled breeding and frequent population bottlenecks due to disease outbreaks and fluctuations in feed supplies. DNA samples from 312 extensively raised chickens from South Africa, Malawi and Zimbabwe were genotyped using the Illumina iSelect chicken SNP60K BeadChip. A panel of 43,157 out of the total 57,636 (74.8%) SNPs was used in the final analysis after screening for those that had a minor allele frequency of less than 5%, were out of Hardy-Weinberg equilibrium ($P<0.01$) and had a call rate of less that 95%. Results indicated that LD averaged between 0.45 and 0.58 for SNPs that had a pairwise distance of less than 20 kb. LD dropped to 0.34 for SNPs between 20 and 100 kb after which it remained constant. LD was further analyzed for its decay over marker distance and differences between populations from different geographic locations. Results are discussed in terms of historical changes in effective population size and resultant recombination rates. The utility of the iSelect chicken SNP60K beadchip in investigating free-range chicken population genetics is demonstrated.

INTRODUCTION

Linkage disequilibrium (LD) is defined as a non-random association of alleles at two or more loci (Hendrik 2005; Qanbari *et al*. 2010). The importance of LD is in providing information about historical events of recombination thereby explaining genetic diversity in genomic regions undergoing selection. LD also allows estimation of effective population size and rate of inbreeding in extensively raised and wild animal populations without pedigree records (Wragg *et al*. 2012).

The village chicken production system in Africa is mainly based on scavenging village chickens (Kitalyi 1998), that are used to meet the multiple household social, economic and cultural needs and are crucial to biodiversity (Delany 2003). However, very little is known about the genetic composition of village chickens in developing regions like Southern Africa. Diversity studies using autosomal microsatellite (Muchadeyi *et al*. 2007) and mtDNA sequences (Mtileni *et al*. 2011) have not defined the genetic stability of these populations. Demographic population parameters such as effective population size and inbreeding levels, that influence the risk to extinction of these populations, remain uncharacterized due to the absence of pedigree and other population census records in these village chicken production systems. The availability of large-scale sequence data in chickens has resulted in an increase in the marker density and achieved a

comprehensive SNP coverage of the chicken genome. The chicken 60K SNP genotyping chip has the potential to unravel the genetic information in extensively raised chicken populations. Applying LD analysis will permit estimation of demographic and evolutionary parameters of these populations. The aim of this study was to investigate the extent and distribution of LD in extensively raised chicken populations of South Africa, Zimbabwe and Malawi using the Illumina iSelect chicken SNP60K BeadChip.

MATERIALS AND METHODS

Chicken populations, blood collection and DNA isolation. A total of 312 village chicken samples were collected from South Africa (n =147), Malawi (n = 30) and Zimbabwe (n = 135). In South Africa, village chickens representing Limpopo (n = 15), Eastern Cape (n = 26) and Northern Cape (n = 35) populations, and four conservation flocks of the Naked Neck (n = 20); Potchefstroom Koekoe (n = 20); Ovambo (n = 10) and Venda (n = 20) chickens kept at Agriculture Research Council Poultry Breeding Resource, were sampled as described in Mtileni et al. (2011). The sampling of the village chickens from Zimbabwe (n = 135) and Malawi (n = 30) populations is described in Muchadeyi et al. (2007). Blood was collected from the selected chickens onto FTA Micro Cards (Whatman Bio Science, UK) and DNA was isolated using a modified protocol of the Qiagen® DNA blood and tissue kit.

SNP genotypes and quality control. The chicken DNA samples were genotyped using the iSelect chicken SNP60K bead chip produced by Illumina Inc. SNP quality control was done using Plink (1.07) software to remove SNPs that were either out of Hardy-Weinberg equilibrium (HWE) *(P< 0.01)*, showing a minor allele frequency (MAF) of at least 5%, had low call rate (< 95%) and with missing genotypes (> 5%). SNPs that were on unknown chromosomes, mtDNA, linkage groups and/or sex chromosomes were excluded from further analyses. After filtering, 45676, 44667,46905 and 43157 SNPs on 28 autosomal chromosomes were used for each of the Malawi, South Africa, Zimbabwe and combined populations, respectively.

Linkage Disequilibrium analysis. A pair-wise LD (r^2) was estimated using PLINK (1.07) software for SNPs on chromosome 1 to 28 for the individuals belonging to the three populations using the following formula:

$$r^2 = \frac{(f11f22 - f12f21)^2}{fA1fA2fB1fB2}$$

A Generalized Linear Model procedure (Proc GLM) in the Statistical Analysis System (SAS) was used to determine the effects of SNP marker interval (bp), chromosome, and population group and interaction of chromosome-by- population on the decay of LD using the following model:

$$r^2_{ij} = \mu + Pop_i + Gga_j + (Pop \times Gga)_{ij} + b\text{SNPint} + e_{ik},$$

Where: Pop_i was the effect of *i*th chicken population of either, Malawi, Zimbabwe or South Africa; Gga_j was the effect of the *j*th chromosome 1-28; and SNPint the effect of SNP interval fit as a covariate with *b* the regression coefficient.

RESULTS AND DISCUSSION

Effects of chicken population, chromosome and distance between SNPs on LD. LD was calculated on 28 of the 38 chicken autosomes. The chromosome size, SNP interval distance and number of SNPs per chromosome support the differences between macrochromosome 1-5 that had high number of SNPs and large intervals between SNPs and micro-chromosomes 16-28, which are smaller and had less SNPs that were relatively close together (Megens *et al*. 2009). Linkage disequilibrium (r^2±SD) averaged 0.38 ± 0.20 and ranged from 0.34 ± 0.14 -0.45 ± 0.24 in Malawi,

0.34 ± 0.15 - 0.52 ± 0.27 in Zimbabwe and 0.34 ± 0.14 - 0.50 ± 0.27 in South African chicken populations. Overall, there was no significant difference in r^2 values ($P<0.05$) between populations indicating similarities between the Malawian, Zimbabwean and South African village chicken populations. However, LD varied significantly between chromosomes ($P<0.001$) with chromosome 8 having the highest LD of 0.52 ± 0.26 followed by chromosome 22 with an r^2 ± SD value of 0.49 ± 0.28. The high LD might be an indication of selection at genes on these chromosomes (Hendrick 2005) particularly natural selection pressures as these chicken populations are raised under extensive systems of production where human selection pressures are minimal (Mtileni *et al*. 2010). Although population did not influence genome-wide LD, a population by chromosome interaction was observed whereby the Zimbabwean chicken population had the highest LD on chromosome 8 (0.52 ± 0.267) and the South African chicken population was highest on chromosome 22 (0.49 ± 0.29). Such interactions need to be further investigated as they might indicate different selection pressures in different populations (Wragg *et al*. 2012).

Another factor that influenced LD was the SNP interval. To further understand this, LD was computed at different distance interval of 0-1 kb, 1-10 kb, 10-20 kb, 20-40 kb, 40-60 kb, 60-100 and 100kb plus using SNP data from chromosomes 1-28 (Fig 1a) and from chromosomes 8; 22 and 13 as indicated in Figures 2b, c and d respectively.

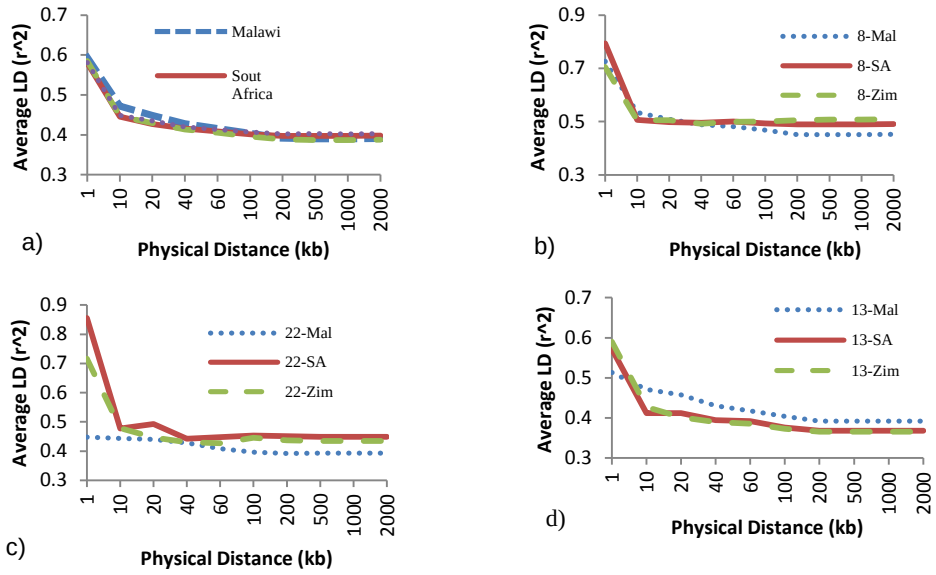

Figure 1. Average LD decay with an increase in physical distance between SNPs for a) chromosomes 1-28, b) chromosome 8; c) chromosome 22; and d) chromosome 18.

The LD averaged 0.58 for SNPs within a 10 kb interval and decayed to 0.45 -0.47 for SNPs between 10-30 kb after which they remained constant. The LD decay at chromosome 8 of the Malawi chickens continued to decline after 40kb. In Zimbabwe and South African chickens, LD at chromosome 22 made a sharp decay from 0.7 (Zimbabwe) and 0.85 (South Africa) to an r^2 below 0.5 at 10kb after which it stayed constant. On the same chromosome LD was maintained around 0.45 over all sliding windows in the Malawi chicken population.

Overall, a higher LD was observed in the Southern African chicken populations compared to

other chicken populations observed in other studies (Qanbari *et al*. 2010; Wragg *et al*. 2012). For example, in a commercial egg laying flock, r^2 averaged 0.32 ± 0.33 with a minimum 0.21 ± 0.26 (Qanbari *et al*. 2010) whereas it was maintained around 0.38 in this study.

CONCLUSION

A relatively high LD that persisted over long SNP intervals was observed in the South African, Zimbabwean and Malawian chicken populations. This LD pattern seems to be consistent with low and steady effective population sizes. The study recommends for a further investigation on the role of selection and population bottlenecks on chromosomes 8 and 22 that had significantly high LD.

ACKNOWLEDGEMENTS

Genotyping of samples was funded by the Agricultural Research Council-Biotechnology Platform (ARC-BTP) while Mr KS Kanyile held fellowships from the ARC-Professional Development Program and the University of KwaZulu-Natal postgraduate program.

REFERENCES

Delany M.E., (2004) *Mec. of Dev*. **121**: 1169.

Hendrick P. W. (2005) Genetics of populations 3rd. Jones and Bartlett publishers, Canada

Illumina Technical notes: DNA analysis, Analyzing GoldenGate Genotyping Data

Kitalyi A. J. (1998). Village chicken production systems in rural Africa: Household food security and gender issues. FAO Animal Production and Health Paper 142, Rome, Italy.

Megens H., Crooijmans R.P.M.A., Bastiaansen J. V. M., Kersten H. H. D., Coster A., Jalving R., Vereijken A., Silva P., Muir W. M., Chen H. H., Hanotte O. and Groene M. A. M. (2009) *BMC Genetics*. **10**: 86.

Mtileni B. J., Muchadeyi F. C., Maiwashe A., Groeneveld E., Groeneveld L. F., Dzama K. and Weigend S. (2011) *Journal of Animal Breeding and Genetics*. **128**: 209.

Mtileni B. J., Muchadeyi F. C., Maiwashe A., Phitsane P. M., Halimani T. E., Chimonyo M. and Dzama, K. (2010) *Applied Animal Husbandry and Rural Develop*. **2**:18.

Muchadeyi F. C., Eding H., Wollny C. B. A., Groeneveld E., Makuza S. M., Shamseldin R., Simianer H. and Weigend S. (2007) *Animal Genetics*. **38**: 332.

Qanbari S., Hansen M., Weigend S., Rudolf P. and Simianer H. (2010) *BMC Genetics*. **11**: 471.

Wragg D., Mwacharo J., Alcalde J., Hocking P. and Hannotte O. (2012) *Heredity*. **109**: 6.