Journal of Novel Applied Sciences
Available online at www.jnasci.org
©2015 JNAS Journal-2015-4-10/1085-1088
ISSN 2322-5149 ©2015 JNAS

The impact of Selection Intensity in level of Linkage Disequilibrium

Mohammad Reza Ghaemi[1,2]

1- Ghaemi International Academic Foundation (G.I.A.F)
2- Iranian Bioinfomatics Society (IBS), Karaj, Iran

***Corresponding author*:** Mohammad Reza Ghaemi

ABSTRACT: The extend and distribution of Linkage Disequilibrium is a topic of great current interest. In new breeding program such as Genome-Wide Prediction Linkage Disequilibrium (LD) between markers and Quantitative Trait Locus(QTL) play an important role. Closely related to the effect of extent of Linkage Disequilibrium on the accuracy of Genomic selection. The level of Linkage Disequilibrium is influenced by many factors. A genome consisiting5 chromosomes each with 100cM in length with 100 equally spaced markers (1cM) were simulated. After 50 generations of random mating in a finite population (Ne=100) in order to create sufficient Linkage Disequilibrium. Population was expended to size of 1000. This structure was conserved to generation 70.Linkage Disequilibrium was estimated from generation 51 to 70 with 5%, 10%, 20%, 40% and 50% selection intensity. The results of this study showed the level of linkage disequilibrium increased by increased of selection intensity.

***Keywords*:** Linkage Disequilibrium, Selection intensity, Simulation.

INTRODUCTION

Linkage Disequilibrium is occurrence of some combination s of alleles or Genetic markers in a population more often or less often than would be expected from random formation of haplotypes from alleles base on their frequency (Ghaemi et al, 2013) .

The classical definition of linkage disequilibrium refers to nonrandom association of alleles between two loci. Consider two markers, A and B that are on same chromosome .A has alleles A1and A2, B has alleles B1and B2.four haplotypes of markers are possible A1B1, A1B2, A2B1, A2B2if the frequencies of alleles A1, A2, B1 and B2 in the population are all 0.5 .Then we would expect the frequencies of each of four haplotypes in the population to be 0.25. Any deviation of haplotype frequencies from 0.25 is Linkage Disequilibrium (Hayes, 2007).

LD provides information about past event and constrains the potential response to both natural and artificial selection. Ld. throughout the genome reflects the population history, the breeding system and the pattern of geographic subdivision, whereas LD in each genomic region reflect the history of natural selection, gene conversation, mutation and other forces that cause gene frequency evolution (Slatkin, 2008).

The level of Linkage Disequilibrium is influenced by number factors , including genetic linkage , the rate of recombination , the rate of mutation, genetic drift , non randome mationg and population structure (Falconer, 1996). As such the objective of this paper was to investigate the impact of selection intensity in level of Linkage Disequilibrium.

MATERIALS AND METHODS

SIMULATION

A genome consisting 5 chromosome each with 100 CM (Centi Morgan) in length with 100 equally spaced SNPs 1 and total number of 50 QTLs 2 (that scattered on chromosome randomly) was generated for each individual . An effective population size of 100 individuals was simulated, of which 50 were male and 50 were female. This structure followed by 50 generations or random mating , implying that each individual had on average number of randome

offspring but only two of them selected in next generation .Each off spring selected by True Breeding Value(TBV). The paternal and maternal haplotypes for each individual were generated base on haldan mapping function (Haldan, 1919) to generate recombinant haplotypes. Fifty generations of random mating were practiced to generate sufficient Linkage Disequilibrium. After 50 generation population was expended to population size of 1000(500male and 500 female) this structure was conserved until generation 70.Two LD measurements r2 and D´ were used to calculated LD in generation (Hayes, 2007):

D´: freq(A1B1)*freq (A2B2) – freq(A1B2)*freq(A2B1)

Where freq(A1B1) is the frequency of A1B1 haplotype in population , and likewise for other haplotypes and also

r2: D2/freq(A1)*freq(A2)*freq(B1)*freq(B2)

Where frequency of the A1 allele in population and likewise for other alleles in population. To creat phenotypic values for each training population, 50 QTL were randomly picked from the set of segregating QTL in that population and TBV was estimated by QTLs frequency (Falconer, 1996):

$$\mathbf{QTL\text{—}AA=2q\alpha,\quad Aa=(q-p)\alpha\ \ and\ \ aa=-2p\alpha}$$

Whereα Estimated By;

$$x=\frac{\sqrt{\delta^2 \mathbf{QTL}}}{2pq}$$

Where;

$$\delta^2\,\mathbf{QTL}=h^2\delta^2\mathbf{P}$$

Where h2 is heritability and δ2P is phenotypic variance and were used 100. Phenotypic records for each animal was estimated by P=TBV+ei where TBV is True Breeding Value and ei is envoirmental effects estimated by ei=N*δ2e where δ2e estimated by δ2e= (1-h2) (δ2P) where N is randome effect. R2 was estimated in population with different selection intensity (i) including 5,10,20,40 and 50 percent. The details of simulation are shown in table1:

Table 1. The parameters used for simulation program

Item	
Genome size	100(cM)
Number of chromosome	5
Number of marker per chromosome	100
Marker density per cM	1
Number of segregation QTL	50
QTL effect	Normal distribution
Dominance of QTL effects	0
Recombination and Crossingover	Haldan Mappng Function
Population size	
Generation 0 to 50	Ne=100
Generation 51 to 70	Ne=100
Heritability	0.5

RESULTS AND DISCUSSION

After fifty generation of random mating in finite population (Ne=100, 50male and 50 female) , considered Linkage Disequilibrium between markers was created two measurement ,D' and r²(in this study r² were estimated for each

generation with 20 repeat) were used to measure the amount of LD between pairs markers in the individuals of generation of 50 was D´±SE=0.6±0.002 and r^2 ±SE=0.18±0.021 .

In this study to create Linkage Disequilibrium was using a willumsen *et al*. (2009) method. Our results show in generation of 50 linkage disequilibrium was around 0.2 and created reference population .similar results have been reported Calus (2009), Saatchi *et al. (*2010) and Ghaemi *et al*. (2013).

Generated the paternal and maternal haplotypes for each individual based on haldan mapping function (at this study crossing over and recombination fraction was also estimated around 0.001) .The population was expanded to the size of 1000(validation set).Phenotypic records and True Breeding Value were estimated as QLTs frequency by the Falconer (1996) method. Linkage Disequilibrium was estimated with different selection intensity including 5%, 10%, 20%, 40%and50% respectively from generation of 51 to generation of 70 and our results shows in figure 1:

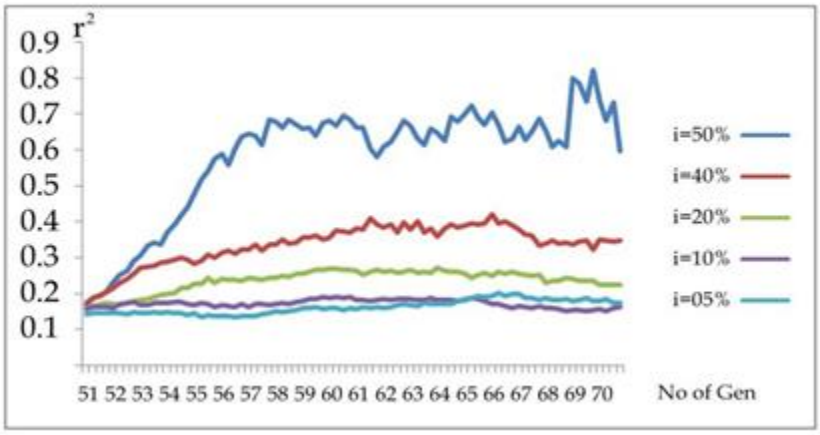

Figure 1. Changes of Linkage disequilibrium in 5%, 10%, 20%, 40% and 50% Selection intensity respectively

The extend and distribution of linkage disequilibrium in animal is a topic of great current interest .LD plays a fundamental role in gene mapping both as a tool for fine mapping of complex disease genes and in proposed genome wide association studies .LD is also of interest for what it can reveal about evolution of population. One of most important factor in selection and allele frequency is a selection intensity. Our results shows the linkage disequilibrium is depend on selection intensity directly on the other hand with increased selection intensity level of linkage disequilibrium was increased (Figure 1).

For genomic selection to be successful as in the simulations of Meuwissen *et al* .(2001),the level of LD between adjacent markers should be r2≥0.2.our simulation results shown minimum level of selection intensity for successful genome-wide prediction is 20 percent (figure1).

Closely related to the effect of the extent of linkage disequilibrium on the accuracy of genomic selection. Calus *et al*. (2009) compared the accuracy of GEBVs with different level of linkage disequilibrium .They shown accuracy of genomic selection increased when the level of the r2 was increased. The results of this study shown the level of Linkage Disequilibrium increased by increased of selection intensity so for successful genomic selection need high selection intensity.

CONCLUSION

The results demonstrate the level of Linkage Disequilibrium influenced by selection intensity. For genomic selection to be successful as in the simulation, the selection intensity should be upper than 20 percent in breeding program.

ACKNOWLEDGEMENT

I acknowledge with warmest thanks to dr.Mehdi Saatchi (Department of Animal Science, Iowa State University), dr. Abbas Jahanbakhshi, dr Mehdi Aminafshar and ms. Farideh Farzin for supported and publish this research article.

REFERENCES

CALUS MPL and VEERKAMP RF. 2007. Accuracy of breeding value when using and ignoring the polygenic effect in Genomic breeding value estimation with a marker density of one SNP per CM .J Anim breed Genet .124:362-368.

CALUS MPL, MEUWISSEN THE, DEROOSE APW and VEERKAMP RF. 2008. Accuracy of Genomic selection using different methods to define haplotype .Genetics.178:553-561.

CALUS MPL. 2009. Genomicbreeding value prediction: method and procedures animal. 4:2,157-164.

FALCONER DS and MACKAY TFC. 1996. Introduction to quantitative Genetics 4th edition .Addison wesley Longman, Harlow, Essex, UK.

GHAEMI MR, AMINAFSHAR M, ASGARI-JAFARABADI G and HAJIBANDEH N. 2013. Effect of heritability coefficient of inbreeding coefficient in genomic evaluation. Animal Science and Research journal 23(4); 99-121.

GODDARD ME. 2008. Genomic selection: prediction of accuracy and maximisation of Long term response. Genetica. Doi: 10.1007.

GODDARD ME and HAYES BJ. 2009. Mapping genes for complex traits in domestic animals and their use in breeding programmes. Nat Rev Genet. 10:381-391.

HABIER D, FERNANDO RL and DEKKERS JCM. 2007. The impact of Genetic Relationship in formation on Genome assisted breeding values. Dairy Sci. 92:433-443.

HALDANE JBS. 1919. The combination of Linkage values and calculation of distance between the loci of linked factors.Journal of Genetics.2:3-19.

HAYES BJ .2007. QTL Mapping, Mas, and Genomic selection. Animal breeding and Genetics Depatment of Animal sceince Iowa state university.

HILL WG and ROBERTSON A. 1968. Linkage Disequilibrium finite populations. Theor.Appl. Genet.38:226-231. HILL, W.G .1981. Estimation of Linkage disquilibrium in randomly matting populations. Heridity.33:229-239.

MEUWISSEN THE, HAYES BJ and GODDARD ME. 2001. predistion of total genetic value using genome – wide dense marker maps. Genetics. 157:1819-1829.

MEUWISSEN THE and GODDARD ME. 2001. Prediction of identity by Descent probabilities from marker haplotypes. Genet.Sel.Evol.33:605-634.

SAATCHI M, MIRAIE-ASHTIANI SR, NEJATY-JAVAREMI A, MORADI SHAHREBABAK M and MEHRBANI YEGANE H. 2010. the impact of information quantity and strenghth of relationship between training set and validation set on accuracy of genomic estimated breeding values. Afr J Biotechnol. 9(4): 438-442.

SLAKTIN M.1972. On treating the chromosome the unit of selection .Genetics.11:147-164.

TOOSI A, FERNANDO RL and DEKKERS JCM. 2010. Genomic selection in admix and crossbred population J Anim Sci. 88:32-46.