Nig. J. Biotech. Vol. 35 No. 2 (2018): 184-191
ISSN: 0189 1731
Available online at
http://www.ajol.info/index.php/njb/index
and www.biotechsocietynigeria.org
DOI: https://dx.doi.org/10.4314/njb.v35i2.22

Molecular evolutionary genetics analysis of major histocompatibility complex class I, II and III genes in ruminants and monogastrics

Bolatito[1*], O. A., Ikeobi[2], C. O. N. and Onagbesan[2], O. M.

[1*]National Centre for Genetic Resources and Biotechnology, PMB 5382, Moor Plantation, Ibadan, Oyo State, Nigeria.

[2]Federal University of Agriculture, Abeokuta, Ogun State, Nigeria.

Abstract

The steady accumulation of mutations over long periods gives rise to MHC polymorphisms and this impacts how each species responds to the host organisms. The nucleotide sequences of MHC class I, II and III genes were retrieved from six selected mammalian livestock species from GenBank, NCBI. ClustalX 2.1 and MUSCLE in the MEGA 7 software were used for the multiple sequences alignment of the nucleotide and amino acids sequences. DnaSP 6.0 software was used to calculate the population genetics parameters. The evolutionary relationship of the MHC I to III proteins of the livestock species was determined by Phylogenetic trees which were constructed using the Neighbour-joining method in the MEGA 7. The DNA polymorphisms of MHC I to III genes among the ruminants and monogastrics revealed high and similar haplotypes diversity. However, the ruminants and monogastrics MHC I gene had low nucleotide diversity in contrary to the MHC II and III genes with high nucleotide diversity. MHC I to III genes had high guanine-cytosine, GC content except the MHC II monogastrics with low guanine-cytosine content. The evolutionary change observed on the livestock MHC I to III genes population deviated from Hardy-Weinberg Equilibrium. Phylogenetic and evolutionary analyses of mammalian livestock species MHC I to III molecules revealed they are slow-evolving at 5, 20 and 2% rates respectively. Similarly, the orthologous relationship existence among the ruminants and monogastrics MHC I to III molecules showed amino acids substitutions of 6 – 18, 5 - 29 and 1 – 12 % respectively since they diverged from the common ancestor. Likewise the optimal trees' sum of branch length (8×10^{-8}, 2×10^{8}, and 4×10^{-8}) indicated that the MHC protein I, II, III molecules of ruminants and monogastrics might have undergone birth-and-death evolution respectively. Since gene duplication and deletion seem to be prevalent among the MHC I, II and III genes and the primary function of MHC is to defend the host from various invaders and great diversity is required; therefore the evolutionary force is mandatory for the diversification. The above information can be utilized to assess the immunogenetic status of livestock population, which will help to increase knowledge on the importance of adaptive genetic polymorphisms in both free ranging and domesticated livestock populations.

Corresponding author: rontitodg@yahoo.com 08034671972

Introduction

Major Histocompatibility Complex (MHC) gene is a member of multigene families that constitute the most important genetic component of the mammalian immune system (Klein, 1986). The MHC molecules contain immune genes that are involved in the antigen presentation; the high level of polymorphism

of MHC genes is essential for generating host immune responses (Buitkamp et al., 1996; Konnai et al., 2003) in vertebrates and is usually arranged in haplotype blocks. A number of agriculturally important traits in animals, especially those related to disease resistance to various pathogenic viruses, bacteria and parasites, are closely linked to the MHC genes (Caron et al., 2013). Similarly, productivity in ruminants and monogastrics depends on the genetic factors governing the production of milk and meat. The general structure of the MHC class I to III genes with different functional roles is conserved among the mammalian species, (Amills et al., 1998), while some are varied when compared (Kelley et al., 2005).

The primary function of the MHC is to bind self or foreign peptides and present them to the T lymphocytes, thereby triggering immune response (Klein and Horejsi, 1997). There are two major types of MHC protein molecules: class I and class II. The former primarily responds to intracellular antigens while the latter binds to the extracellular antigens. The MHC also contains a variety of genes that code for other proteins: such as complement proteins, cytokines and enzymes, that is class III MHC molecules. MHC molecules can either be classical or non-classical; the former is highly polymorphic while the latter is less polymorphic. The high degree of polymorphism of the classical genes is essential for protecting the host from attack by various parasites which are always changing with time (Klein et al., 1993).

Evolution of MHC genes polymorphisms is explained by another model, birth-and-death model of evolution proposed by Nei and Hughes (1992). The individuals are usually species (lineage); new genes are created by repeated gene duplication (speciation) and some duplicated genes are maintained in the genome for a long time while other are deleted (extinction) or become nonfunctional by deleterious mutations. Takahashi et al. (2000) and Nei & Rooney, (2005) corroborated that the MHC class I and II are subject to evolution by birth-and-death process. However, several authors have associated the MHC polymorphisms with concerted evolution (Hood et al., 1975; Ohta, 1980), cross over or gene conversion (Lopez de Castrol et al., 1982; Ohta, 1983), speciation (Figueroa at al., 1988), resistance and susceptibility of variety of infectious and parasitic diseases (Zidi et al., 2008) and so on. With the public availability of molecular data, the phylogenetic analysis of MHC and other immune genes actually revealed different patterns of evolution. The phylogenetic tree do infers the evolutionary relationship of species and patterns of gene duplications and their common ancestors (Nei and Kumar, 2000). To understand the evolutionary pattern of MHC genes, this research work computationally characterized the DNA polymorphisms and phylogeny of major histocompatibility complex class I, II and III genes variability among the six mammalian livestock species.

Materials and Methods

Retrieval of MHC I, II and III nucleotide and protein sequences

The resource population comprised ruminant and monogastrics animals. Four nucleotide sequences of the classical leucocytes antigens (LA) were mined from the National Center for Biotechnology Information (NCBI, USA) database for each of the livestock species which comprised of cattle, sheep, goat, horse, pig and rabbits. In the same manner were the protein sequences of the MHC I, II, and III (TNF) genes retrieved from the same database. The FASTA format of the nucleotide and protein sequences of MHC genes at the NCBI (NCBI, USA), saved in notepad was used for the study.

Multiple Sequence alignment of MHC I, II and III nucleotide and protein sequences and population genetics analysis

ClustalW and MUSCLE in MEGA 7 (Kumar et al., 2016) were used for the multiple sequence alignment of the nucleotide and protein sequences respectively. Following the retrieval of these genes, the nucleotide sequences were edited manually to remove gaps and ClustalW in MEGA 7 (Kumar et al., 2016) was used for the multiple sequence alignment. Thereafter, DnaSP 6.0 (Julio et al., 2009) software was used to calculate the haplotype (gene) diversity (Hd), number of haplotype (h), nucleotide diversity (Pi), nucleotide diversity (Jukes and Cantor) Pi(JC), average number of nucleotide differences

(k) (Tajima, 1983), parsimony informative sites (PIP), singleton variable sites (SP), number of polymorphic/segregating sites (S) and GC content (G+C) contents for each of the livestock species respectively.

Phylogeny construction of MHC I, II and III Proteins

The MHC I, II and III protein sequences alignment was also executed with MUSCLE in the MEGA 7 (Kumar et al., 2016) for the animal species respectively. The unconserved sites of the multiple alignments were edited manually for each of the MHC protein classes. Thereafter, the Phylogenetic trees were constructed from the protein sequences of MHC class I, II and III genes derived from six livestock species using the Neighbour-joining method (Saitou and Nei, 1987). The trees were drawn to scale for each of the MHC genes classes, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of the number of amino acid substitutions per site. The analysis involved 6 amino acid sequences each. All positions containing gaps and missing data were eliminated. Bootstrapping with NJ (1000 replicates) was used in assessing the reliability of individual branches (Felsenstein 1985) that clustered together.

Results

Tables 1 - 3 show the nucleotide polymorphisms of major histocompatibity complex (MHC) I, II and III genes in ruminants and monogastrics. In like manner, Table 4 and Figures 1 - 3 show the evolutionary relationships of livestock major histocompatibity complex (MHC) class I, II and III proteins.

Table 1: Nucleotide polymorphisms of Major histocompatibity complex (MHC) I genes in ruminants and monogastrics

Livestock species	Number of Sequence	h	Hd	K	Pi	ᵖi(JC)	PIP	Sp	S	G+C content
Cattle	4	3	0.83	701	0.22	0.29	858	23	1081	0.53
Sheep	4	4	1.00	96	0.09	1.00	46	121	167	0.63
Goat	2	2	1.00	102	0.17	0.19	0.0	102	102	0.64
Horse	4	4	1.00	140	0.15	0.17	38	214	252	0.56
Swine	4	4	1.00	366	0.28	0.41	43	617	660	0.57

Number of Haplotype **h**, Haplotype (gene) diversity **Hd**, Average number of nucleotide differences **K,** Nucleotide diversity **Pi**, Nucleotide diversity (Jukes and Cantor) **Pi (JC),** Parsimony informative sites **PIP**, singleton variable sites **Sp**, Number of polymorphic (segregating) sites **S** (PIP+Sp), GC content (**G+C).**

Table 2: Nucleotide polymorphisms of Major histocompatibity complex (MHC) II genes in ruminants and monogastrics.

Livestock species	Number of Sequence	h	Hd	K	Pi	Pi(JC)	PIP	Sp	S	G+C content
Cattle	4	3	1	2355	0.62	1.33	658	2784	3442	0.43
Sheep	4	4	1	603	0.55	1.08	221	683	904	0.52
Goat	4	4	1	133	0.44	0.77	46	176	222	0.52
Horse	4	4	1	3607	0.61	1.23	4336	935	5271	0.41
Swine	4	4	1	3187	0.61	1.28	865	3791	4656	0.44

Number of Haplotype **h**, Haplotype (gene) diversity **Hd**, Average number of nucleotide differences **K,** Nucleotide diversity **Pi**, Nucleotide diversity (Jukes and Cantor) **Pi (JC),** Parsimony informative sites **PIP**, singleton variable sites **Sp**, Number of polymorphic (segregating) sites **S** (PIP+Sp), GC content (**G+C).**

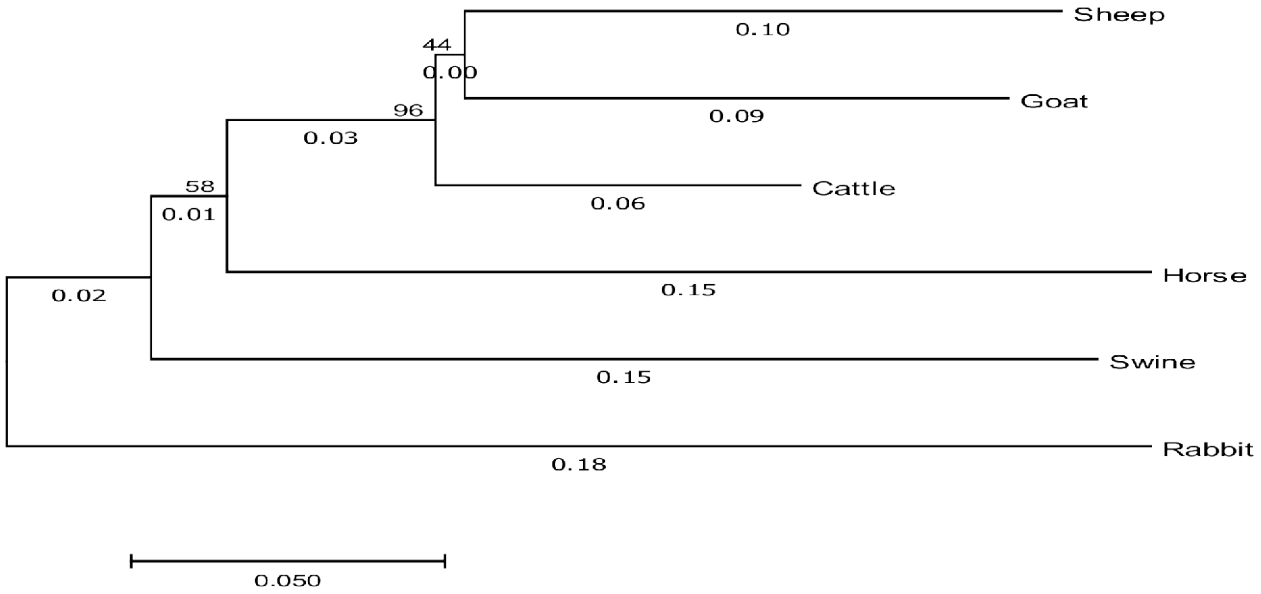
Table 3: Nucleotide polymorphisms of Major histocompatibity complex (MHC) III genes in ruminants and monogastrics.

Livestock Species	Number of Sequence	h	Hd	K	Pi	Pi(JC)	PIP	Sp	S	G+C content
Cattle	4	4	1	1900	0.63	1.38	658	2784	3442	0.52
Sheep	4	4	1	1211	0.61	1.31	0.00	2598	2598	0.52
Horse	4	4	1	1469	0.63	1.43	1771	365	2136	0.49
Swine	3	3	1	2109	0.60	1.24	0.00	2825	2825	0.49
Rabbit	4	4	1	1984	0.62	1.33	4336	935	5271	0.49

Number of Haplotype **h**, Haplotype (gene) diversity **Hd**, Average number of nucleotide differences **K,** Nucleotide diversity **Pi**, Nucleotide diversity (Jukes and Cantor) **Pi (JC),** Parsimony informative sites **PIP**, Singleton Variable sites **Sp**, Number of polymorphic (segregating) sites **S** (PIP+Sp), GC content (**G+C).**

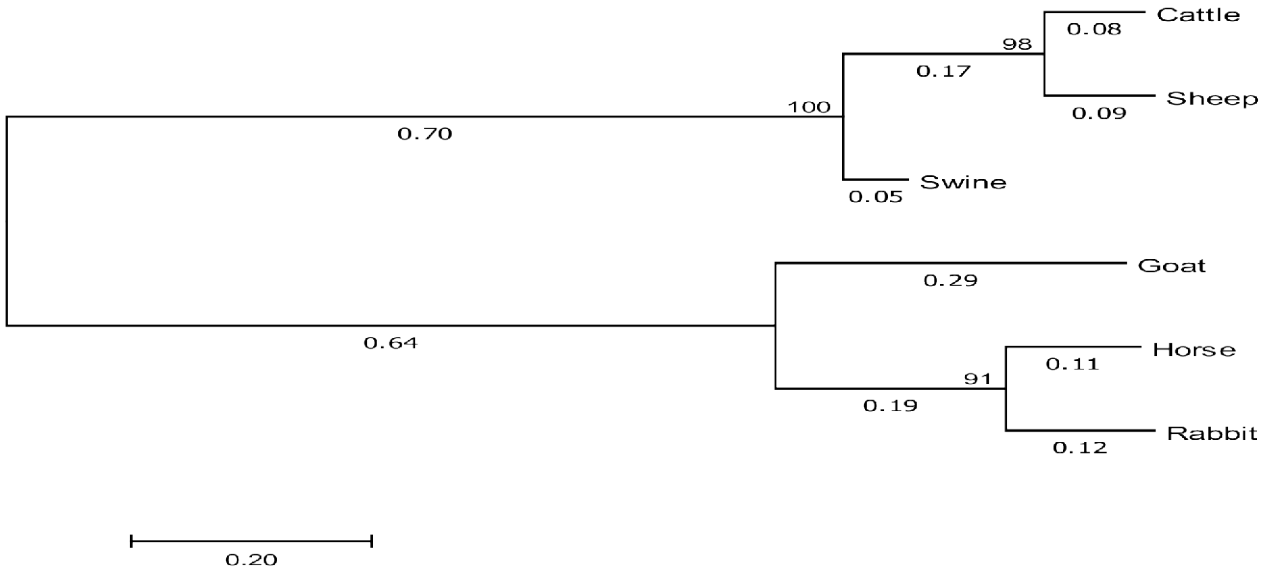
Table 4: Amino acid sequence of six livestock MHC class I to III genes with their ascension numbers and sequence lengths.

Species	MHC Genes Name	MHC class I to III Genes	Amino acids sequence length	GeneBank accession number/ NCBI Ref. Number for protein
Cattle	BoLA	I	171	AAX21792.1
		II	261	BAA07171.1
		TNF	234	AAI34756.1
Sheep	OLA	I	365	AGN98174.1
		II	261	SMA53616.1
		TNF	234	NP_001020031.1
Goat	Cahi-LA	I	358	ABQ14768.1
		II	253	BAA23386.1
		TNF	234	BAA13130.1
Swine	SLA	I	361	AAN35113.1
		II	261	BAA33887.1
		TNF	232	NP_999187.1
Rabbit	RLA	I	361	AAA98730.1
		II	82	NP_001122065.1
		TNF	235	AAA31484.1
Horse	Eqca-LA	I	303	ABV44773.1
		II	255	ACD65012.1
		TNF	234	NP_001075288.2



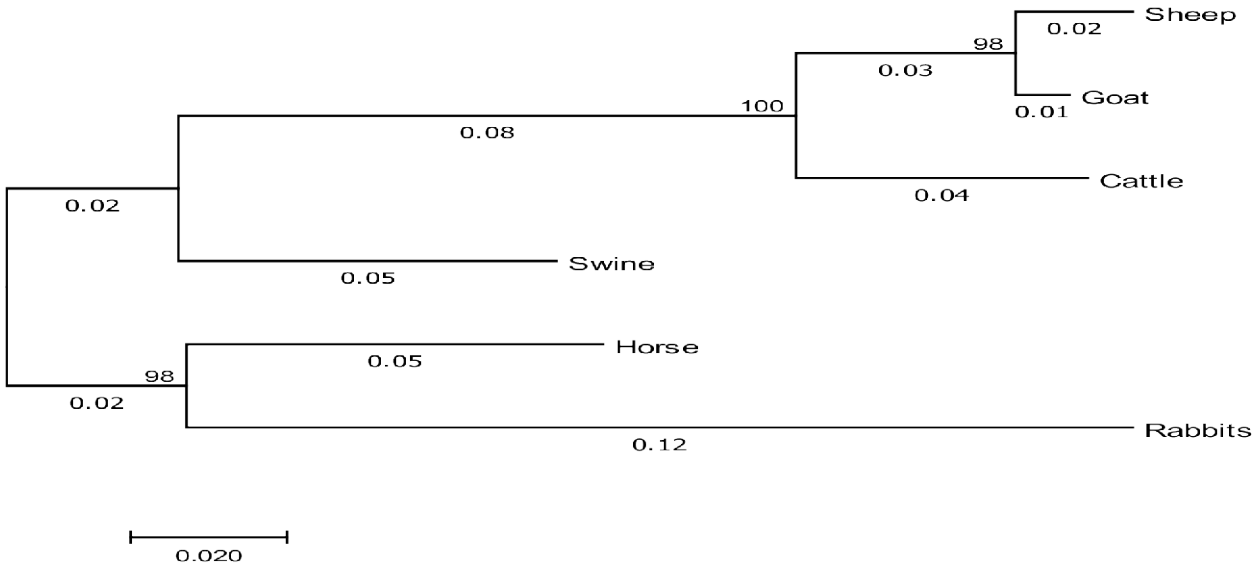
The optimal tree's sum of branch length was 0.79337230 (8×10^{-8}).

Figure 1: Neighbour-joining phylogenetic tree showing the evolutionary relationships among the ruminants and monogastrics MHC class I proteins.



The optimal tree's sum of branch length was 2.44885709 (2×10^{8}).

Figure 2: Neighbour-joining phylogenetic tree showing the evolutionary relationships among the ruminants and monogastrics MHC class II proteins.



The optimal tree's sum of branch length was 0.43970902 (4×10^{-8}).

Figure 3: Neighbour-joining phylogenetic tree showing the evolutionary relationships among the ruminants and monogastrics MHC class III proteins.

Discussion and Conclusion

Major Histocompatibity Complex genes polymorphisms

The DNA polymorphisms of the MHC class I to III genes among the ruminants and monogastrics revealed high and similar haplotypes diversity. Behl et al. (2012) inferred the high degree of observed polymorphism may help in the identification of superior haplotypes for disease resistance in an individual. However, both the ruminants and monogastrics MHC class I gene had low nucleotide diversity and this depicted that the gene is conserved in them, contrary to the MHC II and III genes with high nucleotide diversity in the ruminants and monogastrics. The high intraspecific diversity of the MHC II gene could be attributed to its monitoring activity against the extracellular and foreign peptides parasites (Dengiel et al., 2005; Sommer, 2005). When MHC genes of different mammals are compared, some regions appear to be well conserved and while others vary widely (Kelley et al., 2005). High guanine-cytosine (GC) content obtained in the MHC class I, II and III genes among the livestock species studied except in the monogastrics MHC II genes with low GC content. GC content is responsible for the polymorphic stability of the MHC genes. The evolutionary diversity observed on the livestock MHC I to III genes population deviated from Hardy-Weinberg Equilibrium. This evolutionary pattern of the MHC I to III genes could be a consequence of high level of exposure to extracellular pathogens and the hypothesis by Hedrick and Kim (2000) confirmed that selection acting on certain population is expected to have effect on genotypic frequencies within populations and thus affecting the level of heterozygosity.

Phylogeny analysis of Major histocompatibility complex protein molecules

Phylogenetic and evolutionary analyses of MHC protein I, II and III molecules revealed they are slow-evolving at 5, 20 and 2% rates respectively among the six livestock species studied. This could be the maintenance mechanism for the MHC genes to remove the non-synonymous nucleotide substitutions that are deleterious to immune protein functions (Hughes and Nei, 1988; 1989). The dendogram obtained showed that MHC proteins of different livestock species belong to different

taxonomic classes which were clearly separated from one another. Ruminants MHC I protein clustered together, likewise the monogastrics having 6 – 18 % amino acid substitutions since they diverged from the common ancestor. Similarly were the swine MHC II proteins clustered with cattle and sheep; while goat clustered with horse and rabbits with 5 - 29 % accumulated changes. MHC III (TNF) proteins had two clades of ruminants and monogastrics clustering except swine which clustered with the ruminants' clade with 1 - 12 % amino acid substitutions since they diverged from the common ancestor. The orthologous relationships of MHC class I, II and III protein molecules in the studied animals is an evidence that these genes (proteins) evolved independently of one another during the evolution of eutherian mammals. Yue and Zhiwen (2011) confirmed this, stating that strains of similar animals always clustered together and this could be pointing out to the evidence of the need for a specific immune response to a common pathogen**.** Likewise, the optimal trees' sum of branch length indicated that the MHC protein I, II, III molecules of ruminants and monogastrics might have undergone birth-and-death evolution respectively. The reports by Hughes and Nei (1990), Klein et al. (1993), Takahashi et al. (2000) and Nei and Rooney (2005) substantiated this evolutionary pattern result of MHC protein molecules to have evolved by birth-and-death process.

References

Amills, M., Ramiya, V., Norimine, J. and Lewin, H. A. (1998). The major histocompatibility complex of ruminants. Rev. Sci. Tech. - O.I.E. (Off. Int. Epizoot.) 17: 108-120.

Behl, J. D., Verma, N. K., Tyagi, N., Mishra, P., Behl, R. and B. K. Joshi. (2012). The Major Histocompatibility Complex in Bovines: A Review. Vet. Sci. Vol. 2012:12p.

Buitkamp, J., Filmether, P., Stear, M. J. and Epplen, J. T. (1996). Class I and class II major histocompatiblity complex alleles are associated with faecal egg counts following natural, predominantly *Ostertagia circumcincta* infection. Parasit. Res. 82:693–696.

Caron, J., Malo, D., Schutta, C., J. W. Templeton, and L. G. Adams. 2013. Genetic Susceptibility to Infectious Diseases Linked to *NRAMP1* Gene in Farm Animals. Madame Curie Bioscience Database. Copyright © 2000-2013, Landes Bioscience. NCBI Bookshelf ID: NBK6283.

Dengjel, J., Schoor, O., Fischer, R., Reich, M. and Kraus, M. (2005). Autophagy promotes MHC class II presentation of peptides from intracellular source proteins. Proc Nat Acad Sci USA. 102: 7922-7927.

Felsenstein J. (1985). Confidence limits on phylogenies: An approach using the bootstrap. Evolution, 39:783-791.

Figueroa, F., Gunther, E. and Klein, J. (1988). MHC polymorphism pre-dating speciation. Nature, 335:265–67.

Hedrick, P. and Kim, T. J. (2000). Genetics of complex polymorphisms: parasites and maintenance of the major histocompatibility complex variation. In R. S. Singh and C. B. Krimbas (eds). Evol. Genet. 1:204-234. From molecules to morphology. Cambridge University Press, Cambridge, UK.

Hughes, A. L. and Nei, M. (1988). Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature, 335:167

Hughes, A. L. and Nei, M. (1989). Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proceedings of the National Academy of Sciences USA, 86:958.

Hughes, A. L. and Nei, M. (1990). Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. Mol. Biol. Evol. 7:491–514.

Julio, R., Librado, P., Sánchez-DelBarrio, J. C., Messeguer, X. and Rozas, R. (2009). DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics, 25(11):1451-1452.

Kelley, J., Walter, L. and Trowsdale, J. (2005). Comparative genomics of major histocompatibility complexes. Immunogenetics, 56: 683-695.

Klein, J. (1986). *Natural History of the Major Histocompatibility Complex*. New York, Wiley & Son.

Klein, J., Ono, H., Klein. D. and O'hUigin, C. (1993). The accordion model of MHC evolution. Prog. Immunol. 8:137–43.

Klein, J. and Horejsi, V. (1997). Immunology. 2nd ed. Blackwell Sci.; London.

Konnai, S., Takeshima, S. N., Tajima, S., Yin, S. A., Okada, K., Onuma, M. and Aida, Y. (2003). The influence of ovine MHC class II DRB1 alleles on immune response in bovine leukemia virus infection. Microbiology and Immunology; 47(3):223-232.

Kumar S., Stecher G., and Tamura K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Bio. and Evol. 33:1870-1874.

Lopez de Castro, J. A., Strominger, J. L., Strong, D. M. and Orr, H. T. (1982). Structure of cross-reactive human histocompatibility antigens HLA-A28 and HLA-A2: possible implications for the generation of HLA polymorphism. Proc. Natl. Acad. Sci. USA. 79:3813–17.

Nei, M. and Hughes, A. L. (1992). Balanced polymorphism and evolution by the birth-and-death process in the MHC loci. In: Tsuji K, Aizawa M, Sasazuki T, editors. 11th Histocompatibility Workshop and Conference. Oxford Univ. Press; Oxford, UK.

Nei, M. and Kumar, S. (2000). Molecular evolution and phylogenetics. Oxford University Press New York, pp 333.

Nei, M. and Rooney, A. P. (2005). Concerted and Birth-and-death Evolution of Multigene Families. Annu Rev Genet. 39:121-152.

Ohta, T. (1980). Evolution and Variation of Multigene Families. Springer-Verlag; Berlin.

Ohta, T. (1983). On the evolution of multigene families. Theor. Popul. Biol. 23:216– 40.

Saitou N. and Nei M. (1987). The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Bio. and Evol. 4:406-425.

Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Front. Zool. 2: 16.

Tajima, F., 1983. Evolutionary relationship of DNA sequences in finite populations. Genetics, 105(2):437-460.

Takahashi, K., Rooney, A. P. and Nei, M. (2000). Origins and divergence times of mammalian class II MHC gene clusters. J. Hered. 19:198–204.

Yue, Yi. And Zhiwen, Xu. (2011). Bioinformatics analysis and characteristics of the giant Panda Interferon-alpha. MECS Press 1:45-54.

Zidi, A., Sanchez, A., Obexer-Ruff, G. and Amills, M. (2008). Sequence analysis of goat major histocompatibility complex class I genes. J. Dairy Sci. 91(2):814-7.

Zuckerkandl E. and Pauling L. (1965). Evolutionary divergence and convergence in proteins. Edited in Evolving Genes and Proteins by V. Bryson and H. J. Vogel, pp. 97-166. Academic Press, New York.