DETECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNP) IN MITOCHONDRIAL GENOME OF RED JUNGLE FOWL (RJF) AND ITS UTILIZATION IN ESTIMATING GENETIC DIVERSITY BETWEEN RJF AND DOMESTIC FOWL

Thesis

Submitted in partial fulfillment of the requirement for the degree
of
Doctor of Philosophy
in
POULTRY SCIENCE

By
Dr. Shahaj Uddin Ahmed
Roll No. 972

To
DEEMED UNIVERSITY
INDIAN VETERINARY RESEARCH INSTITUTE
Izatnagar – 243 122 (U.P.)
2009

DIVISION OF AVIAN GENETICS AND BREEDING
CENTRAL AVIAN RESEARCH INSTITUTE
IZATNAGAR - 243 122 (U.P)

Dr. Deepak Sharma,
Principal Scientist.

Tel : 0581-2302037
e-mail:dscari@rediffmail.com

Dated :.....

Certificate

Certified that the research work embodied in this thesis entitled **"Detection of Single Nucleotide Polymorphism (SNP) in Mitochondrial Genome of Red Jungle Fowl (RJF) and its Utilization in Estimating Genetic Diversity Between RJF and Domestic Fowl"** submitted by **Dr. Shahaj Uddin Ahmed, Roll No. 972,** for the award of **Doctor of Philosophy** degree in **Poultry Science** at Indian Veterinary Research Institute, Izatnagar, is the original work carried out by the candidate himself under my supervision and guidance.

It is further certified that **Dr. Shahaj Uddin Ahmed,** has worked for more than 30 months in this Institute and has put in more than 300 days attendance under me from the date of registration for the degree of **Doctor of Philosophy** of the Deemed University, as required under the relevant ordinance.

(Deepak Sharma)
Chairman
Advisory Committee

Certificate

Certified that the thesis entitled **"Detection of Single Nucleotide Polymorphism (SNP) in Mitochondrial Genome of Red Jungle Fowl (RJF) and its Utilization in Estimating Genetic Diversity Between RJF and Domestic Fowl"** submitted by **Dr. Shahaj Uddin Ahmed**, Roll No. **972**, in partial fulfillment of the requirement of **Doctor of Philosophy** degree in **Poultry Science,** Deemed University, **Indian Veterinary Research Institute,** Izatnagar, embodies the original work done by the candidate. The candidate has carried out his work sincerely and methodically.

We have carefully gone through the contents of the thesis and are fully satisfied with the work carried out by the candidate, which is being presented by him for the award of **Doctor of Philosophy** degree of this Institute.

It is further certified that the candidate has completed all the prescribed requirements governing the award of **Doctor of Philosophy** degree of **Indian Veterinary Research Institute**.

Signature of External Examiner

Dated : .....

(DEEPAK SHARMA)
Chairman
Advisory Committee

Date.....

MEMBERS OF STUDENT'S ADVISORY COMMITTEE

1. **Dr. B.P. Singh,** Pr. Scientist
   Division of AG&B, CARI .....
2. **Dr. D.P. Singh,** Pr. Scientist
   Division of AG&B, CARI .....
3. **Dr. (Mrs) Meena Kataria,** Pr. Scientist
   Division of Animal Biochemistry, IVRI .....
4. **Dr. Kuldeep Dhama,** Sr. Scientist
   Division of Avian Diseases, IVRI .....
5. **Dr. Pramod Kumar Tyagi,** Sr. Scientist
   Division of Avian Nutrition and Feed Technology, CARI .....

Acknowledgements

On the accomplishment of the present study, I would like to take the opportunity to extend my deepest sense of gratitude and words of appreciation towards those who helped me during the pursuit of this study. I deem it a proud privilege and feel immense pleasure to acknowledge all those who are directly or indirectly involved.

At the very outset, I express my profound sense of gratitude and indebtness to my guide Dr. Deepak Sharma, Principal Scientist, Division of Avian Genetics and Breeding, CARI, Izatnagar, for his excellent guidance, valuable advice whole-hearted encouragement and critical appreciation, kind affection and interaction with me during the entire period of study. He is the man who gave me the value of sincerity, perseverance, tolerane, patience to life and personality. Without him, this dissertation would not have been possible.

I express my heartiest gratefulness to the esteemed members of Advisory Committee, Dr. B. P. Singh, Director, CARI, Izatnagar, Dr. D. P. Singh, Principal Scientist, Division of Avian Genetics and Breeding, CARI, Izatnagar, Dr. (Mrs) Meena Kataria, Principal Scientist, Division of Animal Biochemistry, IVRI, Izatnagar, Dr.Kuldeep Dhama, Senior Scientist, Division of Avian Diseases, IVRI, Izatnagar, and Dr. Promod Kumar Tyagi, Senior scientist, Division of Avian Nutrition and Feed Technology, CARI, Izatnagar, for their innovative suggestions, critical evaluation and constant timely help and affectionate encouragement.

I also express most sincere gratitude to Dr. M.C. Kataria, Pr. Scientist and Head Division of AG&B, CARI, Izatnagar, Dr. R . P . Moudgal, Principal Scientist, Dr. Jog Mahan , Principal Scientist, Dr. A.B. Mandal, Principal Scientist, Dr. R . P. Singh, Principal Scientist, Dr. Sameer Majumder, Principal Scientist, Dr. V. K . Saxena, Senior Scientist, Dr. S . K. Bhanja, Senior Scientist, Dr. (Mrs) Simi Tomar, Senior Scientist, CARI, Izatnagar, Dr. P. Joshi , Principal Scientist, Division of Animal Biochemistry and Dr. S . D. Singh, Principal Scientist, and Dr. J . M. Kataria, Principal Scientist and Head , Division of Avian Diseases, I.V.R.I. Izatnagar, for their excellent teaching and kind advice.

I express my gratitude to Dr.Mukund Kadam, Dr. Amitav Bhattacharjya, Dr. Avishek Biswas, Dr. Ravi Sunderasan, Dr. Rakesh Thakur, Dr. Bal

Gangadhar Roy and Dr. Ajmat Alam Khan, Dr. Sadhagar, Dr. Nani Gopal Shit,Santosh Marandi, Rahul, Kirupasankar, Yagesh, Arun Chetri, Ram, Soib Akhtar and my dear Dr.Girraj Goyal for their multifarious help and affection during my study period.

It is great privilege to express my whole- hearted indebtedness and sincere regards to Dr.D.J.Kalita, Dr. Rituraj Konwar, Dr.Dhanjit Kr. Das, Dr. Prasanta Saikia, Dr. Sourabh Deuri, Dr.Purabi Barman, Dr.T.K. Rajkhowa, Dr. Gunjan Das, Dr. Farhat Pandit, Dr. N.C. Nath and Dr. Deepak Gogoi for their valuable suggestions, kind help and constant encouragement during my whole study period in I.V.R.I. Izatnagar.

The homely atmosphere and cheerful company of Dr. Dimpal Thakuria, Dr. Dr. Bikash Borthakur, Dr. Juwar Doley, Dr. Durlov Bora and Dr. Jayanta Kr. Chamuah and Dr. Raja Das which made my stay at I.V.R.I. most enjoyable and memorable and will always remain in my sweet memoirs.

Words are inadequate to express my greatefulness to my friends namely Dr. Bandu Mane, Durgadas Nandi, Ujjal Kumar De, Bablu Kumar, K.K. Mishra, Rahul , Vinod , Ravi for their cheerful company in IVRI.

I am thankful to my lab mates Satyendra, Somesh, Sanjeev Kumar Shukla, Ashutosh, Anjana, Vinay Tomar, Manish Mehra, Mr. P.D.Tiwari and Vijay for their pleasant cooperation and kind help during my research period.

I feel proud to express my gratefulness to my dear friends Dhiraj Kr. Ray, Atul Chandra Boro , Joynur Ali, Kandarpa Kalita, Nandita Das in this most coveted occasion fo r their unconditional help, constant inspiration, moral support, affection and whole hearted co operation, and evergreen company without which my study would not have been completed so smoothly in time.

I will always be remaining anew and vivid for the inestimable help , active support, constant inspiration, love and affection , rendered during this study period with zest and zeal by Naba and Dr. Pranab Jyoti Das.

It seems very difficult to express in words my sincere thanks to my achates Dr. Hemen Das, who has been a guiding star for my successful completion of entire study period.

I express my sincere gratitude to Director, Dr. S.P.S. Ahlawat , for providing the requisite facilities and financial assistance in the form pf I.V.R.I. Fellowship to pursue my study.

I have profound sense of respect to Dr. Dharmeswar Das, Joint Director (Academic), IVRI, for his precise and timely help during my whole study period in I.V.R.I. Izatnagar.

I owe a deep sense of gratitude to Dr. B. P. Singh, Director, CARI, Dr. H. P. Shrivastava, Head, PGE&T section, CARI, Izatnagar, for their valuable advice and timely help in extending all the required facilities during my study period..

The kind- hearted companionship, love and affection received from Dr. Anish D. and Dr. Suba is thankfully acknowledged.

I take this opportunity to pay tribute to my reverend teacher Dr. T. More, Pr. Ssientist, Div. of Biochemistry, IVRI, Izatnagar, Dr. Kushal konwar Sharma, Dr. H.F. Ahmed, Serifuddin Ahmed and Dhiren Medhi whose contribution has made me to achieve this goal.

I feel extremely delighted in extending my thanks to all my teachers, near and dear ones, friends, and all my well wishers for their blessings to achieve this goal.

With deep respect , I express my heart- felt gratefulness to my beloved parents, aunties, brothers, sisters, nephews and niece and Lony for their immortal love and blessings, moral support, constant encouragement and inspiration and unconditional help during the entire study period of my life.

I express my greatefulness to my elder sister, Sumi for her constant encouragement, moral support, immense love and prayer and also for standing by me during the whole study period without which the study would not have been completed.

Finally, I thank the Almighty for showering upon me His divine providence throughout the journey of my life.

Date : (**Shahaj UddinAhmed**)

Abbreviations

%	Per cent
µg	Microgram
µl	Microlitre
A	Adenine
ATP	Adenosine triphosphate
bp	Base pair
C	Cytosine
cDNA	Complementary Deoxyribonucleic acid
cm	Centimetre
COI	Cytochrome Oxidase I
CYTB	Cytochrome B
dATP	Deoxyadenosine triphosphate
ddNTP	Dideoxynucleoside triphosphate
DF	Domestic Fowl
D-loop	Displacement loop
DNA	Deoxy ribonucleic acid
dNTP	Deoxyribonucleoside triphosphate
e.g.	*exampli gratia* (for example)
EDTA	Ethylene diamine tetra acetic acid
et al	*et alii*
etc	*et cet·er·a*
G	Guanine
hr	Hour
IPTG	Isopropyl-beta-D-thiogalactopyranoside
min	Minute
ml	Millilitre
mm	Millimetre
mM	Millimolar
NADH	Nicotineamide Adinine Dehydrogenase
NCR	Non-coding region
nt	Nucleotide
^{o}C	Degree Centigrade
PCR	Polymerase Chain Reaction
pM	Pico mole
RE	Restrction Endonuclease

RJF	Red Jungle Fowl
RNA	Ribonucleic Acid
rpm	Rotation per minute
rRNA	Ribosomal Ribonucleic Acid
RT	Reverse transcription
RT-PCR	Reverse transcription - polymerase chain reaction
s	Second
SNP	Single Nucleotide Polymorphism
T	Thymine
TAE	Tris acetate ethylene diamine tetra acetate
TBE	Tris borate ethylene diamine tetra acetate
tRNA	Transfer Ribonucleic Acid
U	Unit
UV	Ultra Violet
w/v	Weight by volume
X-Gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
α	Alpha

List of Figures

Fig. No.	Title
2.1	Avian Mitochondrial DNA
3.1	The pTZ57R/T vector
4.1	CO1-114-C/T snp genotyping using allele specific PCR amplification in red jungle fowl (A) and domestic fowl (B). M : 100 bp ladder.
4.2	COI-255-G/A snp genotyping using allele specific PCR amplification in red jungle fowl (A) and domestic fowl (B). M : 100 bp ladder.
4.3	COII-279-C/T snp genotyping using allele specific PCR amplification in red jungle fowl (A) and domestic fowl (B). M : 100 bp ladder.
4.4	ATP6-294-G/A snp genotyping using allele specific PCR amplification in red jungle fowl (A) and domestic fowl (B). M : 100 bp ladder.
4.5	ND4-969-C/T snp genotyping using allele specific PCR amplification in red jungle fowl (A) and domestic fowl (B). M : 100 bp ladder.
4.6	CYTB-543-T/C snp genotyping using allele specific PCR amplification in red jungle fowl (A) and domestic fowl (B). M : 100 bp ladder.
4.7	Amplification of 1325 bp fragment in RJF (Lane 1-2). M 100 bp ladder plus.
4.8	Amplification of 1325 bp fragment in RJF from the colony PCR (Lane 1-2. M 100 bp ladder plus.
4.9	Resolution of 1325 bp fragment purified from gel. M : 100 bp ladder plus.
4.10	Resolution of 4241 bp fragment representing whole plasmid plus insert (lane 1) and *EcoR* I and *Pst* I digest 1401 bp fragment (insert) along with 2840 bp fragment (remaining plasmid DNA) (lane 2) and M : molecular weight marker.
4.11	Nucleotide Sequence (5'-3') of 1325 bp fragment containing 45 nucleotide from 3' end of tRNA-Glu, complete nucleotide of D loop region (1232) and 48 nucleotides from 5' end of tRNA-Phe.
4.12	Aligned sequences of the genus Gallus.

4.13 Plot of Gallus species control region variability in 50 base windows. Substitutions are indicated with black bars and indels with grey bars.

4.14 Phylogenetic tree constructed by NJ method based on the nucleotide sequences of D loop region in Indian RJF, other subspecies of RJF other jungle fowls an domesticated fowl.

4.15 Sequence alignment of first 400 bp of mitochondrial D loop region in different subspecies of Red Jungle Fowl.

4.16 Phylogenetic tree constructed by NJ method based on the sequence of first 400 nucleotides of D loop region in Indian RJF and other subspecies of RJF.

List of Tables

Table No.	Title
2.1	Avian Mitochondrial DNA.
2.2	The avian species having complete sequence of mt DNA.
3.1	Details of accession numbers used.
3.2	Details pf the primers used for SNP genotyping.
3.3	Details of accession numbers used.
3.4	Localities of the different samples from sub species of Red Jungle Fowl.
4.1	Size variation for different mitochondrial genes/r RNA /t RNA/ D loop region in domestic and red jungle fowl.
4.2	Polymorphic sites in different mitochondrial genes/r RNA /t RNA/ D loop region between as well as within in domestic and red jungle fowl.
4.3	Allelic frequency at different SNP position in RJF and Domestic fowl.
4.4	The size of different domain of D loop region in different Gallus species.
4.5	Base composition of different domains of D loop region in Indian Red Jungle Fowl.
4.6	Genetic distance (Jukes-Cantor) between different Gallus species based on nucleotide sequence variability in D loop region.
4.7	Genetic distance (Jukes-Cantor) between different *Gallus gallus* sub species based on first 400 nucleotide sequence variability in D loop region.

Contents

Sl. No.	CHAPTER	PAGE NO.
1.	INTRODUCTION	01-03
2.	REVIEW OF LITERATURE	04-20
3.	MATERIALS AND METHODS	21-29
4.	RESULTS	30-47
5.	DISCUSSION	48-62
6.	SUMMARY	63-67
7.	MINI ABSTRACT	68
8.	HINDI ABSTRACT	69
9.	BIBLIOGRAPHY	70-79

Introduction

Poultry industry in India has a strong standing globally, being third highest egg producer with more than 50000 million eggs and fifth highest broiler producer with 2.25 million tones of broiler meat. Largely, Indian poultry industry is chicken dominated and the rearing of high yielding varieties under intensive rearing system is the back bone of its present face. Presently, broiler industry is increasing at a growth rate of 10 %, while egg production has a growth rate of about 6 % (Singh and Sharma, 2007). This spectacular growth of Indian poultry industry has its root in the availability of high producing germplasm and the high yielding potential of these germplasm was well supported by advances in nutritional and health management and husbandry practices have also contributed.

The divergence between Red Jungle Fowl (RJF) and Domestic Fowl (DF) is presumed to have originated some 8,000 years back. During all these years, the chicken genome has undergone tremendous changes due to intentional selection. Unlike other domestic species, where the ancestor from which the present day animals evolved are not existing. Chicken offers a unique system to compare the effect of natural selection verses man made artificial selection. Thus, it has immense value. Today, this bird is facing the danger of extinction in term of its purity. Peterson and Brisbin (1999) studied the 754 museum specimen of Red Jungle Fowl and using morphological characters, mainly presence of eclipse plumage and concluded that almost 99% of the captive population and potentially all wild population are contaminated by feral and

domestic chicken. This aspect needs meticulous appraisal on scientific lines. However, in reality the position may not be so worse. It is expected that Red Jungle Fowl population of Himachal Pradesh, Uttar Pradesh and Uttaranchal may have negligible or no contamination. The scientific studies in this regard are required.

In the recent past, molecular genetics have provided several powerful tools such as DNA based genetic markers that can exploit the great wealth of polymorphism at DNA level. DNA-based markers can be grouped into clone/ sequence based (CSB) markers and fingerprint markers (Dodgson *et al.,* 1997). The CSB markers include microsatellite (Weber and May, 1989), RFLP, Sequence Tagged Site (STS) and Expressed Sequence Tags (EST). On the other hand, the fingerprint (FP) markers include variable number of tandem repeats (VNTR). These VNTR are often called minisatellites for large loci (Jeffreys *et al.,* 1985) and microsatellites for smaller VNTR loci (Ali *et al.*, 1986). Single nucleotide polymorphism (SNP), a new type of molecular marker and its more recent concept has basically arisen from the recent need for the densities of genetic markers for the studies of multifactorial diseases and the recent progress in polymorphism detection and genotyping techniques. Detection of SNPs in the genome is important to determine whether they cause silent changes, contribute to the diversity of the population, or cause mutations leading to diseases. SNPs are widely distributed throughout the genomes (Halushka *et al.*, 1999). The occurrence and distribution of SNPs varies among species. In chicken, 1 SNP per 225 base pair (bp) was observed in a survey of 31,000 bases analysed from broiler and layer lines (Schmid *et al.*, 2000) and 1 SNP per 2,119 bp in chicken expressed sequence tags (EST) (Kim *et al.*, 2002). Several other SNPs were also reported in chicken. Amills *et al.* (2003) reported that the genotypes of the 4 SNPs in the chicken IGF-2 gene are associated with growth and feeding traits in chicken. The SNPs of the chicken insulin like factor-binding protein-2 gene associated with chicken growth and carcass traits was reported by Lie *et al.* (2005). The identification of SNPs of adipocyte fatty acid-binding protein gene and its association with fatness traits in the chicken was reported by Wang *et*

al. (2006). Qui *et al.* (2006) studied the association of single nucleotide polymorphisms in the insulin gene of chicken for early growth and fat deposition.

The primary advantage of mtDNA is that it is present in higher copy numbers within the cells and therefore more likely to be recovered even from highly degraded specimen (Kline *et al.*, 2005). Ingman *et al.* (2001) described mtDNA as a potent tool for understanding human evolution, owing to its high copy number, apparent lack of recombination, high substitution rate and maternal mode of inheritance. The SNPs that occur in the mitochondrial genome are very important tool in the species identification, forensic studies as well as anthropological and evolutionary research (Vigilant *et al.*, 1991 and Boles *et al.*, 1995). Harumi *et al.* (2004) observed the single nucleotide polymorphism in the D-loop region in chicken mitochondrial DNA and reported 11 SNPs. For six of the SNP sites, polymerase chain reaction (PCR) primers that had each base (A, C, G and T) as the penultimate base at the 3' end were produced to type the polymorphisms. Twenty-one out of 96 primers succeeded in distinguishing the SNPs by the presence or absence of PCR product. This method provides an easy way to discriminate SNP in chicken mitochondrial DNA. A rapid identification of SNPs and estimation of allele frequencies using sequences from DNA pools was reported by Ye *et al.*(2006).

Hence, the present study has been mainly aimed for the detection of SNPs in the mitochondrial DNA of Red Jungle Fowl and its utilization in estimating genetic diversity between RJF and Domestic Fowl with the following objectives.

1. ***Genotyping the potential SNPs by allele specific PCR and/or PCR RFLP in red jungle fowl and domestic fowl population.***
2. ***Estimating genetic diversity between RJF and DF using SNPs.***

Review of Literature

Mitochondrial DNA (mtDNA)

In 1949, Boris Ephrusi suggested that mitochondria have their own genome and several years later, in the mid 1970, mitochondria were found to contain DNA. The DNA found in the mitochondrion is called mitochondrial DNA. The first complete sequence to be resolved was that of the human mitochondrial genome (Anderson *et al.*, 1981). Since then several vertebrate mitochondrial genomes have been sequenced, including that of the domestic fowl.

Avian mitochondrial DNA

Desjardins and Morais (1990) studied the complete mitochondrial (mt) genome sequences of WLH chicken (*Gallus gallus domesticus*). Sequence analysis revealed that the WLH chicken mt genome is 16,775 bp and the genome encodes the same set of genes (13 proteins, 2 rRNAs and 22 tRNAs) as do other vertebrate mt DNAs and is organized in a very similar economical fashion (Fig.2.1). The protein genes are highly similar to their mammalian and amphibian counterparts and are translated according to the same variant genetic code. Despite these highly conserved features, the chicken mitochondrial genome displays two distinctive characteristics. First, it exhibits a novel gene order, the contiguous $tRNA^{Glu}$ and ND6 genes are located immediately adjacent to the displacement loop region of the molecule, just ahead of the contiguous $tRNA^{Pro}$, $tRNA^{Thr}$ and Cyt b gene which border the D-loop region in other vertebrate mt genomes. This unusual gene order is conserved among the galliform birds. Second, a

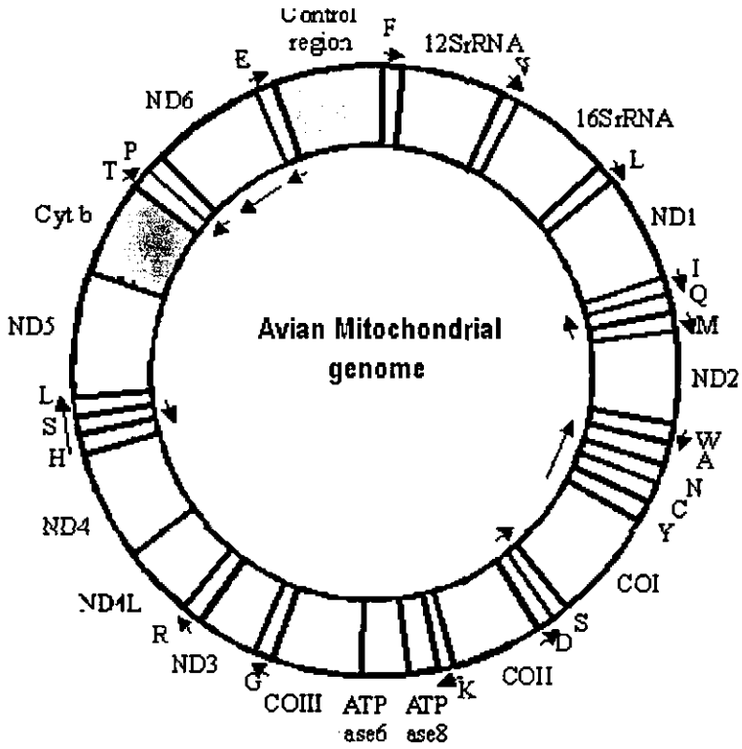

Fig 2.1. Avian Mitochondrial DNA

light stand replication origin, equivalent to the conserved sequence found between mt tRNA Cys and tRNA-Asn genes in all vertebrate mt genome sequences thus far, is absent in the chicken genome. These observations indicate that galliform mt genomes departed from their mammalian and amphibian counterparts during the course of evolution of vertebrate species. The complete mitochondrial sequence of Red Jungle Fowl along with their gene position and size is presented (Table 2.1). The mitochondrial genome of other chicken breeds, avian species and jungle fowl has also been completely sequenced (Table 2.2).

Desjardins *et al.* (1990) deduced the gene organization of the Peking duck mitochondrial (mt) DNA through heterologous hybridization using different cloned fragments of the chicken or Japanese quail mitochondrial genome as probe. As in the chicken and other gallinaceous birds, the Peking duck mtDNA displays a novel gene order which differs from that of other vertebrates by the unusual localizationn of the tRNA Glu and ND6 genes next to the displacement (D) loop region of the molecule. The position of these genes with respect to the mitochondrial D-loop region, the cytochome oxidase subunits I, II and III, the NADH ehydrogenase subunit I and the ribosomal rRNAs was confirmed by the partial nucleotide sequence of cloned mt DNA fragments.

Desjardins and Morais (1991) cloned and sequenced the segments of the Japanese quail mitochondrial genome encompassing many tRNA and protein genes, the small and part of the large rRNA genes, and the control region. Analysis of the relative position of these genes confirmed that the tRNA Glu and ND6 genes in galliform mitochondrial DNA are located immediately adjacent to the control region of the molecule instead of between the cytochrome b and ND5 genes as in other vertebrates. Japanese quail and chicken display another distinctive characteristic, that is, they both lack an equivalent to the light strand replication origin found between the tRNA Cys and tRNA Asn genes in all vertebrate mitochondrial genomes sequenced thus far. Comparison of the protein-encoding genes revealed that a great proportion of the substitutions are silent and involve many transitions. Sequence alignment indicated that the two avian control regions evolve mainly through base substitutions but are also characterized by the occurrence of a 57-bp deletion/addition event at their 5' end.

Table 2.1. Avian Mitochondrial DNA

Sl No	Name of genes	Code
1	D loop control region	D-loop
2	tRNA phenylalanine	tRNA-Phe
3	12S ribosomal RNA	12S rRNA
4	tRNA valine	tRNA-Val
5	16S ribosomal RNA	16S rRNA
6	tRNA leucine	tRNA-Leu1
7	NADH dehydrogenase 1	ND1
8	tRNA isoleucine	tRNA-Ile
9	tRNA glutamine	tRNA-Gln
10	tRNA methionine	tRNA-Met
11	NADH dehydrogenase 2	ND2
12	tRNA tryptophan	tRNA-Trp
13	tRNA alanine	tRNA-Ala
14	tRNA asparagine	tRNA-Asn
15	tRNA cysteine	tRNA-Cys
16	tRNA tyrosine	tRNA-Tyr
17	Cytochrome Oxidase I	COI
18	tRNA serine 1	tRNA-Ser
19	tRNA aspartic acid	tRNA-Asp
20	Cytochrome Oxidase II	COII
21	tRNA lysine	tRNA-Lys
22	ATP Synthase 8	ATPase8
23	ATP Synthase 6	ATPase 6
24	Cytochrome Oxidase III	COIII
25	tRNA glycine	tRNA-Gly
26	NADH dehydrogenase 3	ND3
27	tRNA arginine	tRNA-Arg
28	NADH dehydrogenase 4L	ND4L
29	NADH dehydrogenase 4	ND4
30	tRNA histidine	tRNA-His
31	tRNA serine 2	tRNA-Ser
32	tRNA leucine	tRNA-Leu2
33	NADH dehydrogenase 5	ND5
34	Cytochrome B	Cytb
35	tRNA threonine	tRNA-Thr
36	tRNA proline	tRNA-Pro
37	NADH dehydrogenase 6	ND6
38	tRNA glutamic acid	tRNA-Glu

Table 2.2. The avian species having complete sequence of mt DNA

Sl No	Species / Breed	Common name / Breed	Total bases	References
1	*G. g. domesticus*	White Leghorn	16,775	Derjardius & Morais, 1990
2	*C. japonica*	Japanese quail	16,697	Nishibori *et al.* (2001)
3	*Numida meleagris*	Guineafowl	16,726	Nishibori *et al.* (2004)
4	*G. g. var. domesticus*	Silkie fowl	16,784	Wada *et al.* (2004)
5	*G. g. murghi*	Red jungle fowl	16,785	Nishibori *et al.* (2005)
6	*G. g. Spadiceus*	Burmese Red Jungle fowl	16,785	Nishibori *et al.* (2005)
7	*G. g. gallus*	Cochin chinese RJF	16,785	Nishibori *et al.* (2005)
8	*G. g. bankiva*	Javan RJF	16,785	Nishibori *et al.* (2005)

L'Abbe *et al.* (1991) mapped the transcription start sites of chicken mitochondrial DNA in the control region by direct sequencing of *in vitro* capped mitochondrial RNA species, by primer extension and by S 1 nuclease protection analysis. Transcription of the heavy strand initiates predominantly at a site 156 nucleotides upstream of the tRNAPhe gene, *i.e.* about 135 nucleotides further upstream than the corresponding sites in amphibia and mammals. On the opposite strand, transcription starts predominantly one nucleotide removed from the site in the heavy strand. The L-strand position start site is similar to that found in other vertebrates. The chicken mitochondrial DNA control region thus contains one major transcriptional promoter, whose bidirectional capacity is similar to the situation in amphibia but which contrasts to the mainly unidirectional capacity of mammalian promoters. In chicken mitochondria, the sequence comprising the start sites is A+T rich and contains an almost perfect inverted repeat which can be folded into a cruciform structure. The heavy and light strand initiation sites are flanked on their respective 3' ends by an octanucleotide sequence matching those surrounding the start sites in *Xenopus laevis* (5'- ACPUTTATA-3'). This motif is found associated with the H-strand start sites in mouse but is not present in human or bovine mitochondrial DNA promoters.

Liu *et al.* (1996) sequenced the displacement loop (D-loop) and the ribosomal RNA genes of mtDNA for *Cairina moschata* and *Anas platyrhynchos*. Duck mtDNA has a specific gene order 5-upstream of the D-loop that is identical to fowl mtDNA.The nucleotide diversity is greatest in the D-loop while being most conserved in the 12S r RNA gene, as indicated by a comparisons of sequences of duck and fowl mtDNA.

Marshal and Baker (1997) sequenced the entire control region and portions of flanking genes (tRNAPhe, tRNAGiu, and ND6) in the common chaffinch *(Fringilla coelebs),* blue chaffinch *(F. teydea),* brambling *(F. montifringiZZa),* and greenfinch *(Carduelis chloris).*In these finches the control region is similar in length (1,223-1,237 bp) and has the same flanking gene order as in other birds, and contains a putative TAS element and the highly conserved CSB-1 and F, D, and C boxes recognizable in most vertebrates. Cloverleaf-like structures associated with the TAS element at the 5' end and

CSB-1 at the 3' end of the control region may be involved with the stop and start of D-loop synthesis, respectively. The pattern of nucleotide and substitution bias is similar to that in other vertebrates, and consequently the finch control region can be subdivided into a central, conserved G-rich domain (domain II) flanked by hypervariable 5'-C-rich (domain I) and 3'-AT-rich (domain III) segments. In pairwise comparisons among finch species, the central domain has unusually low transition / transversion ratios, which suggests that increased G+T content is a functional constraint, possibly for DNA primase efficiency. In finches the relative rates of evolution vary among domains according to a ratio of 4.2 (domain III) to 2.2 (domain I) to 1 (domain II), and extensively among sites within domains I and II. Domain I and III sequences are extremely useful in recovering intraspecific phylogeographic splits between populations in Africa and Europe, Madeira, and a basal lineage in Nefza, Tunisia. Domain II sequences are highly conserved, and are therefore only useful in conjunction with sequences from domains I and III in phylogenetic studies of closely related species.

Harlid *et al.* (1997) reported the complete mtDNA sequence of ostrich. The size of the molecule was found to be 16,591 bp. Since the ostrich is a representative of the palaeognathous birds, comparison with the mt DNA of the fowl (neognathous) the only bird species reported so far in databases. Relative to other vertebrates, the NADH-Glu genes of birds are positioned upstream of the control region rather than the cytochrome b gene.The NADH 3 gene of the ostrich was found to be terminated by a stop codon at the position 207. Thus, the gene was about 140 nucleotides shorter than in other vertebrates. The sequence for the L – strand origin of replication was missing in both species, and 4 transfer RNA genes of the 2 tRNA of vertebrate mtDNA by having an adenine at position 8. Phylgenetic analysis of complete cytochrome b gene of 6 orders of bird showed that the Passeriformes represent the earliest divergence among recent birds, contradicting the commonly accepted notion of a basal position of the palaeognathae among recent birds.

Mindell *et al.* (1998) reported that mitochondrial genomes of all vertebrate animals analyzed to date have the same 37 genes, whose arrangement in the circular DNA molecule varies only in the relative position of a few genes. This

relative conservation suggests that mitochondrial gene order characters have potential utility as phylogenetic markers for higher-level vertebrate taxa. We report discovery of a mitochondrial gene order that has had multiple independent originations within birds, based on sampling of 137 species representing 13 traditionally recognized orders. This provides evidence of parallel evolution in mitochondrial gene order for animals. Our results indicate operation of physical constraints on mitochondrial gene order changes and support models for gene order change based on replication error. Bird mitochondria have a displaced O_L (origin of light-strand replication site) as do various other Reptilia taxa prone to gene order changes. Our findings point to the need for broad taxonomic sampling in using mitochondrial gene order for phylogenetic analyses. We found, however, that the alternative mitochondrial gene orders distinguish the two primary groups of songbirds (order Passeriformes), oscines and suboscines, in agreement with other molecular as well as morphological data sets. Thus, although mitochondrial gene order characters appear susceptible to some parallel evolution because of mechanistic constraints, they do hold promise for phylogenetic studies.

Randi and Lucchini (1998) sequenced the entire mitochondrial DNA control region (mtDNA D loop} in the seven extant species *of Alectoris* partridges. The D-loop length is much conserved (1155 ± 2 nucleotides), and substitution rate are lower than for the mitochondrial cytochrome b geneof the same species, on average. Comparative analysis suggest that these D-loops can be divided into three domains corresponding to the highly variable peripheral domains I and III and to the central conserved domain II of vertebrates. Nevertheless. first 161 nucleotides of domain I of the *Alectoris,* immediately flanking the tRNA Glu, evolve at an unusual ly slow rate and show motifs similar to the mammalian extended termination-associated sequences [ETAS1 and ETAS2, which can form stable secondary structures. The second part of domain I contains a hypervariable region with two divergent copies of a tandemly repeated sequence described previously in other species of anseriforms and galliforms. Some of the conserved sequence blocks of mammals can be mapped in the central domain of *Alectoris*. Domain III is highly variable and has sequences similar to mammalian CSB1. The bidirectional transcription promoter HSP/LSP box of the chicken is partially conserved among the *Alectoris*.

Quinn and Wilson (1993) cloned and sequenced 3.4 kb of snow goose mtDNA and found that the ND5 gene is followed by the genes for cytochrome b, tRNA Thr, tRNA Pro, ND6, tRNA Glu, the control region, tRNA Phe and srRNA. This order is identical to that of chicken, quail, and duck mtDNA but differs from that of mammals and *a frog (Xenopus).* The mean extent of difference due to base substitution between goose and chicken is generally closer to the same comparison between rat and mouse but less than that between human and cow. For one of the nine regions compared (tRNA Glu), the bird differences appear to be anomalous, possibly implicating altered functional constraints. Within the control region, several short sequences common to mammals are also conserved in the birds. Comparison of the goose control region with that of quail and chicken suggests that a sequence element with similarity to CSB-1 duplicated once prior to the divergence of goose and chicken and again on the lineage leading to chicken. Between goose (or duck) and chicken there are four times more transversions at the third positions of fourfold-degenerate codons in mitochondrial than in nuclear genes.

Crochet and Desmarais (2000) sequenced part of the mitochondrial control region and the cytochrome *b* gene in 72 specimens from 32 gull species (Laridae, Larini) and 2 out-group representatives (terns: Laridae, Sternini). Our control region segment spanned the conserved central domain II and the usually hyper-variable 3' domain III. Apart from some heteroplasmy at the 3' end of the control region, domain III was not more variable than domain II or the cytochrome *b* gene. Furthermore, variation in the tempo of evolution of domain III was apparent between phyletic species groups. The lack of variation of the gull control region could not be explained by an increase in the proportion of conserved sequences in these birds, and the gull control region showed an organization similar to those of other avian control regions studied to date. A novel invariant direct repeat was identified in domain II of gulls, and in domain III, two to three inverted, sometimes imperfect, repeats are able to form a significantly stable stem-and-loop structure. These putative secondary structures have not been reported before, and a comparison between species groups showed that they are more stable in the group with the more conserved control region. The unusually slow rate of evolution of control region part III of the gulls could thus

be partly explained by the existence of secondary structures in domain III of these species.

Roquis *et al.* (2004) reported the first entire mitochondrial DNA (mtDNA) control region sequences in two endangered vulture species, the bearded vulture *(Gypaetus barbatus)* and the Egyptian vulture (*Neophron percnopterus*). Results showed that the general organization of vulture control regions was very similar to other birds, with three distinct domains: a left variable domain (DI), a central conserved one (DII) including the F, E, D, and C boxes, and a right domain (DIII) containing the CSB1 sequence. However, due to the presence of long tandem repeats, vulture control regions differed from other avian control regions both in size and nucleotide composition. The Egyptian vulture control region was found to be the largest sequenced so far (2031 bp), due to the simultaneous presence of repeats in both DI (80 bp) and DIII (77 bp). Low variation was found in vulture control regions, particularly in *G. barbatus*, as the probable result of populations declines in the last few centuries.

Applications of mt DNA polymorphism

A. Species identification

Highly conserved region of mitochondrial DNA that includes Cytochrome b gene, 12S rRNA gene, 16S rRNA gene and hypervariable region like D-loop region have been used by various workers for species differentiation. Cytochrome b gene is the most commonly used target for PCR-RFLP. In one of the first report, Meyer *et al.* (1995) PCR amplified 359 bp Cytochrome b fragment and cut with *Alu*I, *Rsa*I, *Taq*I and *Hinf*I to identify pig, cattle, wild boar, buffalo, sheep, goat, horse, chicken and turkey meat. Zimmermann *et al.* (1998) developed a technique for identification of cattle, pigs, sheep, chickens, turkeys, rabbits, European hares, dogs, cats, fallow deer, red deer, roe deer and bison by RFLP using *Alu*I and *Nco*I and base sequence analysis of 981 bp DNA fragment from cytochrome b gene.

Matsunaga *et al.*(1999) developed multiplex PCR using primers targeted against Cytochrome b gene. They were able to detect adulteration upto 1% of chevon, chicken, beef and upto 2.2 and 3.6% of mutton and horse meat, respectively added to pork. Chevon, chicken and beef were detected at 1% level

with mutton or horse as base meat. Amplification of a fragment of the mitochondrial encoded Cytochrome b gene for discrimination of raw meat from chickens, turkeys, swine, cows and sheep was standardized by Herman (2001). Gene amplification of Cytochrome b1 and b2 can be applied for identification of chicken, turkey meat and derived products (Ferrari *et al.*, 2001). To differentiate emu (*Dromaius novaehollandiae*) and ostrich (*Struthio camelus*), a zone of gene coding for Cytochrome b can be amplified (Colombo *et al.*, 2000). This gives fragments of 543 bp in ostrich vs. 229 bp in emu, distinguishable by horizontal PAGE.

Other mitochondrial genes were also used for meat speciation. Colgan *et al.* (2001) designed PCR assay targeting mitochondrial ATPase 6/8 gene in bovine, ovine, porcine and poultry species. A detection limit of 0.3% was determined for both the bovine and ovine-specific primers, while a limit of 1% was obtained with porcine-specific primers. Verkaar *et al.* (2002) described bovine species identification method based on species-specific mutation on mt DNA of Cytochrome b and Cytochrome oxidase II.

Rodriguez *et al.* (2003) clearly identified the goose (*Anser anser*), mule duck (*Anas platyrhynchos* x *Cairina moschata*), chicken (*Gallus gallus*), turkey (*Meleagris gallopavo*) and swine (*Sus scrofa domesticus*) in foie gras by using the common forward primer designed on a conserved DNA sequence in the mitochondrial 12S rRNA gene and reverse primers were designed to hybridize on species-specific DNA sequences of each species considered by a specific polymerase chain reaction and detects approximately 1% of adulteration.

Bellagamba *et al.* (2003) described a DNA monitoring method to examine fishmeal for contamination with mammalian and poultry products to avoid the risk of the diffusion of mad cow disease. A PCR method based on the nucleotide sequence variation in the 12S r RNA gene of mitochondrial DNA was developed and evaluated. The PCR specifically detected mammalian and poultry adulteration in fishmeal containing 0.125% beef, 0.125% sheep, 0.125% pig, 0.125% chicken and 0.5% goat. They also developed a multiplex PCR assay for detection of adulteration of fishmeal by ruminant and pig meat.

Martin *et al.* (2007) applied the PCR technique for the qualitative identification of chicken (*Gallus gallus*), turkey (*Meleagris gallipavo*), duck (*Anas platyrhynchos* x *Cairina muschata*), and goose (*Anser anser*) tissues in feed-stuffs, on an individual basis using the oligonucleotide primers that are specific for each avian species, targeting the 12S rRNA mitochondrial gene which generated amplicons of 95, 122, 64 and 98 bp length for chicken, turkey, duck and goose, respectively and also reported that the performance of this method was not affected by prolonged heat-treatment (upto 133°C for 20 min at 300 kPa), and consequently, it could be very useful for the accurate identification of tissues from these 4 avian species in products submitted to denaturing technologies, for which other methods cannot be applied.

Identification of different poultry meat species by using the mitochondrial 12S rRNA gene was also reported by Girish *et al*. (2007). Saini *et al.* (2007) investigated the use of polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) of mitochondrial 12S rRNA gene for identification of peacocks (*Pava cristatus*) using universal primers designed on mitochondrial 12S rRNA gene and amplified the 446 bp amplicon followed by digestion with commonly available restriction enzymes *Alu*I and Sau3AI produced a differential pattern for identifying these poultry species unambiguously.

B. Estimation of genetic diversity among species / breeds

Another potential application of sequence variation in mitochondrial genes is the estimation of genetic diversity among species / breeds and establishing a phylogenetic relationship among them. There are ample reports on this aspect in poultry itself.

Lee *et al.* (1989) examined the nucleotide sequence of the structural genes for the mt asp, lys, ser-tRNAs from chicken including the CoxII gene. The sequence arrangement was virtually identical to that found in other vertebrate mtDNA. The tRNA (asp) and tRNA (lys) were positioned upstream and downstream of the CoxII gene respectively, while tRNA (ser) was transcribed from the opposite strand.

Pan *et al.* (1993) studied the Cytochrome c oxidase subunit II of domesticated ducks (*Anas platyrhynchos*) and their hybrid and found 91%

sequence homology. The nucleotide substitution pattern in the gene revealed a high transition/ transvertion ratio (8.8: 1.0).

Fumihito *et al.* (1995) studied the mt DNA control region for establishing the genetic link between the Chinese bamboo partridge (*Bambusicola thoracica*) and the chicken and jungle fowls of the genus Gallus. A comparison of mt control region, DNA base sequences of 16 species belonging to the subfamily phasianidae revealed that generalized perdicine birds (quails and partridges) are of ancient lineages. Even the closest pair, the common quail (*Coturnix coturnix* japonica) and the Chinese bamboo partridge maintained only 85.71% identify.

Fumihito *et al.* (1996) determined the D-loop regions of the mt DNA sequence of birds. They determined a total of 21 birds, which 12 samples belonged to Red Jungle Fowl (*Gallus gallus*), comprising 3 subspeceis, 6 *G. g. gallus*, 3 *G. g. spadiceus* and 3 *G. g. bankiva* and 9 represented diverse domestic breeds (*G.g. domesticus*). Four green jungle fowls (*G. lafayettei*) and one grey jungle fowl (*G. Sonneratii*) were also sequenced. A continental population of *G. g. gallus* was found to be the real matriarchy origin of all types of domestic poultry.

Munechika *et al.* (1997) analysed the restriction endonuclease cleavage patterns of mtDNA from the sequence divergence between *G. gallus* and *G. sonnerati*, *G. varius* and *G.lafatattei* was 0.9. 10.5 and 12.6 % respectively. Assuming that the rate of evolution of mt DNA is 3 % per million years. It was estimated that the divergence of mtDNA of *G. gallus* from that of *G. sonnerati*, *G. varius* and *G. lafayattei* occurred 0.3, 3.5 and 4.5 million years ago, respectively.

Zhon *et al.* (1997) studied the variation of restriction cleavage pattern of mtDNA in chahua, Banna Fighter, wuding and Nyxi breeds of Yunnan.Samples were digested with 11 restriction enzymes and analyses by agarose gel electrophoresis. All birds produced the same 41 cleavage fragments. The size of mtDNA and 4 breeds. The study suggested that these fowl were derived from a common wild predecessor.

Nishibori *et al.* (2001) studied the complete mt sequence of the Japanese quail (*Cotrunix japonica*) and its genetic relationship with related species

sequence analysis reveal that the Japanese quail mt genome is a circular DNA of 16.697 bp, which is smaller than chicken mt DNA of 16.785 bp but the genomic structure of Japanese quail mt DNA is same as that of chicken. The sequence homology of an mt genes including those for 12s rRNA and 16s rRNA, between Japanese quail and chicken ranges from 78 to 89.9%

Nishibori *et al.* (2002) studied the complete nucleotide sequences of the *Coturnix chinensis* mt genome and a phylogenetic analysis with related species. They reported that the mt genome of blue-breasted quail is 16,687 bp and it is smaller than mt Japanese quail and chicken by 10 and 88 bp respectively. They also reported that the sequence identity of all mt genes, including those for 12s and 16s rRNA between blue breasted quail and Japanese quail ranged from 84.5% to 93.5% between blue breasted quail and chicken sequence identity ranged from 78.0 to 89.6%.

Dong *et al.* (2002) compared the mt DNA D-loop region of 5 chinese native chicken breeds (*G. g. domesticus*), a foreign breed (white leghorn) and four jungle fowl i.e. *G. gallus*, *G. Sonneratii*, *G. various* and *G. lafayettei.* The study revealed that *G. g. domesticus* was closest the Red Jungle Fowl (*G. gallus*).

Yan *et al.* (2002) sequenced a part of the mt D-loop region of silky chicken and other 3 native breeds and compared 24 variation sites were detected among 20 DNA sequences, with a sequence divergence of 4.45%, which suggested that there was a less differences between silky chicken and other native breeds, In genetic relationship, silky chicken is closest to Baiyiner chicken, followed by Xiaoshan chicken and the farthest from Xianju chicken.

Wada *et al.* (2004) sequenced the whole mt DNA of the silkie fowl (*G. g.* var. *domesticus*) and found 99.77 % similarity between silkie and white leghorn (39 nucleotide differences in 16,784 bp). Yan *et al.* (2002) sequenced a part of the mt D-loop region of silky chicken and other 3 native breeds and reported a sequence divergence of 4.45%, which suggested fewer differences between silky chicken and other native breeds.

Nishihori *et al.* (2004) studied the complete nucleotide sequence of *Numida meleagris* mt genome and it was revealed that the mt DNA is a circular DNA of 16,726 bp with a genomic structure the same as that of *G. g.*

var. domesticus and *Coturnix japonica* mt DNAs, though it is 62 bp smaller than *G. g. domesticus* mt DNA and 29 bp larger than C. *japonica* mt DNA. Similarities of the 13 genes and two ribosomal RNAs except D-loop and tRNAs between *N. meleagris* and *C. japonica* ranged from 77.0 to 88.8% and from 76.2 to 88.4% respectively.

Liu *et al.* (2004) investigation the genetic variability of mt DNA hypervariable HVSI) in a total of 48 birds belonging to 12 chinese native chicken breeds. Sixteen haplotypes were identified from 35 polymorphic nucleotide sites, which accounted for 6.4% of a sequenced 544 bp fragment. Diversity analysis of the haplotypes showed that Tibetan, Langshan and Henan cockfight chicken had only one haplotypes, while ancient haplotypes existed in Taihe silkie and Chahua chicken. Phylogenetic analysis of the haplotypes suggested that Chinese native chicken breeds shared the same maternal lineage, regardless of their external features and ecological types. Both divergent and phylogenetic analysis of the haplotypes indicated the close genetic relationship between the Chinese native chicken breeds and *G. gallus* and G.*g. spadiceus* from different areas, which implied that G. *g. gallus* and G. *g. spadiceus* were the original ancestors of Chinese native chicken breeds.

Komiyama *et al.* (2004) studied the evolutionary origin of long- crowning chicken with fighting cocks disclosed by the mt DNA sequence analysis. This study revealed that the Nanganaki dori varieties were morophyletic and originated from a fighting cook.

Nishibori *et al.* (2005) studied the phylogenetic analysis based on mt DNA sequences and revealed that two grey jungle fowls (GJF) were clustered in a clade with RJFs and chicken, and that one GJF was located in a remote position close to Ceylon jungle fowl (CJF). These findings taken together strongly indicate that interspecies hybridizations have occurred between GJF and RJF/ chicken and between GJF and CJF.

Liu *et al.* (2006) analysed the mt DNA hypervariable segment I (HVS-I) for 834 domestic chickens (*G. g. domesticus)* and 66 wild Red Jungle Fowl (*G. gallus*) from Southeast Asia and China. There distinct distribution patterns and expansion signatures suggest that different clades may originate form

different regions, such as Yunnam, South and South West China and or surrounding areas i.e. Vietnam, Burma, Thiland and the Indian subcontinent, respectively, which support the theory of multiple origin in South and Southeast Asia.

Lee *et al*. (2007) studied mitochondrial DNA (mtDNA) sequence variations in Korean Ogol chicken. A total of 31 Korean Ogol chicken was grouped into four haplotypes and the large haplotype was represented in 12 individuals. The un-rooted neighbor-joining tree indicates that the Korean Ogol chicken shared three (A to C) major chicken lineages representing the high genetic variability of this breed. These results can be used for making the breeding and conservation strategies for the Korean Ogol chicken.

Kanginakudru *et al*. (2008) analyzed 76 Indian birds that included 56 *G. g. murghi* (RJF), 16 *G. g. domesticus* (domestic chicken) and 4 *G. sonneratii* (Grey JF) using both microsatellite markers and mitochondrial D-loop sequences. Microsatellite marker analyses of Indian birds indicated an average FST of 0.126 within *G. g. murghi*, and 0.154 within *G. g. domesticus* while it was more than 0.2 between the two groups. The microsatellite-based phylogenetic trees showed a clear separation of *G. g. domesticus* from *G. g. murghi*, and *G. sonneratii*. Mitochondrial DNA based mismatch distribution analyses showed a lower Harpending's raggedness index in both *G. g. murghi* (0.001515) and in Indian *G. g. domesticus* (0.0149) birds indicating population expansion. When meta analysis of global populations of 855 birds was carried out using median joining haplotype network, 43 Indian birds of *G. g. domesticus* (19 haplotypes) were distributed throughout the network sharing haplotypes with the RJFs of different origins. These results suggested that the domestication of chicken has occurred independently in different locations of Asia including India. We found evidence for domestication of Indian birds from *G. g. spadiceus* and *G. g. gallus* as well as from *G. g. murghi*, corroborating multiple domestication of Indian and other domestic chicken. In contrast to the commonly held view that RJF and domestic birds hybridize in nature, the present study shows that *G. g. murghi* is relatively pure. Further, the study also suggested that the chicken populations have undergone population expansion, especially in the Indus valley.

Single nucleotide polymorphism (SNP)

As suggested by the Acronym, an SNP marker is just a single base change in a DNA sequence, with a usual alternative of two possible nucleotides at a given position. SNPs are the nucleotide variations in the DNA sequence of individuals in a population and constitute the most abundant molecular markers in the genome. They can be used for various purposes including the construction of the high- density genetic maps and association studies related to genetic disorders (Douabin-Gicquel *et al.*, 2001).

SNPs are widely distributed throughout the genomes (Halushka *et al.*, 1999). The occurrence and distribution of SNPs varies among species. In human, an estimated 1 SNP per 1,000 base pair (Sachidanandan *et al.*, 2001). In chicken, 1 SNP per 225 base pair (bp) was observed in a survey of 31,000 bases analysed from broiler and layer lines (Schmid *el al.*, 2000) and 1 SNP per 2,119 bp in chicken expressed sequence tags (EST) (Kim *et al.*, 2002).

Detection of SNPs in the genome is important to determine whether they cause silent changes, contribute to the diversity of the population, or cause mutations leading to diseases. The SNPs that occur in the mitochondrial genome are very important tool in the species identification, forensic studies as well as anthropological and evolutionary research (Vigilant *et al.*, 1991; Boles *et al.*, 1995).

Soleimani *et al*. (2003) reported the Allele Specific PCR with 3 primers as an efficient, cost effective, and reliable way for SNP validation in barley cultivars. Bundock *et al*. (2005) reported an alternative cheap, simple, transferable and robust method of SNP genotyping, based on a three primer nested allele specific PCR. They developed markers for 40 SNP sites in 28 barley genes based on the nested allele specific PCR method. We obtained a 95% success rate for the development of allele specific markers (for one or both alleles) with this approach. Markers for both alleles at a SNP site were developed in 84% of cases. Gaudet *et al*. (2007) reported a new method for polymerase chain reaction (PCR) amplification of multiple specific alleles (PAMSA), which allows efficient discrimination of SNP polymorphisms in one reaction tube with standard PCR conditions. This improved PAMSA requires only three unlabeled primers: a

common reverse primer and two allele-specific primers having a tail of different length to differentiate the two SNP alleles by the size of amplification products on agarose gel. A destabilizing mismatch within the five bases of the 3' end is also added to improve the allele specificity. To validate the accuracy of this method, 94 full-sib individuals were genotyped with three SNPs and compared to the genotypes obtained by cleaved amplified polymorphic sequence (CAPS) or derived CAPS. This method is flexible, inexpensive, and well suited for high throughput and automated genotyping.

Li *et al*. (1998) described a single-nucleotide polymorphism (SNP) in the NADH dehydrogenase subunit IV gene that appeared to be associated with resistance to Marek's disease in White Leghorns. In chicken, Zhiliang *et al.* (2002) studied myostatin gene in different chicken popuations and identified seven SNPs in promoter and untranslated regions. Zhiliang *et al.* (2004) observed significant association of different genotypes with skeletal muscle growth (fat, body weight and muscle parameters).

Amills *et al.* (2003) reported that the genotypes of the 4 SNPs in the chicken IGF-2 gene are associated with growth and feeding traits in chicken. The SNPs of the chicken insulin like factor-binding protein 2 gene associated with chicken growth and carcass traits was reported by Lie *et al*. (2005).

Harumi *et al*. (2004) identified 6 SNPs in mt D loop regions from sequence alignment of D loop region sequences from six chickens. For six of these, they designed the PCR primers that had each base (A,C G and T) as the penultimate base at the 3' end (N2 base) and bases specific to both alleles at the 3' end. These primers were effective in typing the SNP. Twenty one out of 96 primers succeeded in distinguishing the SNP by the presence or absence f the PCR product.

Sano *et al*. (2005) developed a new screening test for male putative germline chimeric chickens by polymerase chain reaction (PCR) that uses single nucleotide polymorphism (SNP) detection primers. Putative germline chimeric chickens were produced by the transfer of stage X blastodermal cells or circulating primordial germ cells from Barred Plymouth Rocks (BPR) to White Leghorns (WL). For screening prior to the progeny test, DNA was extracted

from the semen and used for PCR analysis, using SNP detection primers at position 020 in the chicken mitochondrial DNA D-loop region (DNA database, AB*3+**2). In this study, the type of SNP in all BPR at position 020 was fixed as base G, and as base A in all WL. When the PCR product was typed as both donor (base G)- and recipient (base A)-derived, the male was determined as the germline chimera. The male and female putative germline chimeric chickens produced in this study were given the progeny test. The results of the male screening test showed good agreement with the progeny test for detecting germline chimeric chickens.

Qui *et al*. (2006) studied the association of single nucleotide polymorphisms in the insulin gene of chicken for early growth and fat deposition. Rapid identification of SNPs and estimation of allele frequencies using sequences from DNA pools was reported by Ye *et al.* (2006).

Sironi *et al.* (2006) reported the estimation of the SNP frequency within the avian Tapasin gene. The identification of SNPs of adipocyte fatty acid- binding protein gene and its association with fatness traits in the chicken was reported by Wang *et al.* (2006).

Bharani Kumar *et al.* (2007) generated the data on ten SNP markers of Myostatin gene (GDF-8) was generated on nine breeds/populations of indigenous poultry and Red Jungle Fowl (RJF). The SNPs were five in promoter region, one each in intron 1 and 2 and three in exon 1. The data was analyzed to find out the genetic relationship among the indigenous chicken populations. PCR-RFLP was carried out to genotype the populations at seven SNPs while three were genotyped by SNaPshot method using automated DNA sequencer as no restriction sites were found at the SNP sites. ANOVA revealed 18% of the variation among the populations. The correspondence analysis separated out Punjab Brown, Red jungle fowl and Nicobari populations from the rest of the indigenous chicken populations. It is postulated that the populations of Punjab Brown and Nicobari showed the recent common ancestry with Red jungle fowl.

Guan *et al.* (2007) used experimental and in silico tools to identify nucleotide variants in the mt Genome, including the coding and non-coding (D-loop) regions. The distribution of the experimentally identified mitochondnal

DNA variants in meat-(broilers) and egg-type (White Leghorn) chickens was also assessed. A total of 113 single-nucleotide polymorphisms (SNPs) were identified. The in silico analysis revealed a total of 91 SNPs, with 70 in the coding region and 21 in the non-coding region. Of the 41 experimentally identified SNPs, 27 were in the D loop. Together, the experimentally identified SNPs in the non-coding region formed I I haplotypes, whereas the 14 SNPs in the coding region formed 6. Though, 9 of the D loop region haplotypes were observed only in broilers, 3 of the 6 haplotypes from the coding region occurred at a significantly higher frequency in broilers. To our knowledge, this investigation represents the first whole-mt Genome scan for variation and an evaluation, though limited in sample size, of the haplotype distribution in meat- and egg-type populations, using the SNPs and haplotypes identified.

Materials and Methods

3.1. Identifying the probable SNPs in mitochondrial genome

Complete mitochondrial DNA sequences from different chicken breeds and *G. gallus* subspecies (Table 3.1) were retrieved from database (http://www.ncbi.nlm.nih.gov.in). Genes identified for study were the two rRNA i.e. 12S ribosomal RNA (12S-rRNA) and 16S ribosomal RNA (16S-rRNA), 13 mRNA i.e. ATPase unit 6 and 8 (ATPase 6 and ATPase 8, respectively), Cytochrome Oxidase subunit I, II and III (COI, COII & COIII, respectively), Cytochrome B (Cytb), NADP dehrdrogenase subunit 1, 2, 3, 4, 4L, 5 and 6 (ND1, ND2, ND3, ND4, ND4L, ND5 and ND6, respectively) and 22 tRNAs i.e. tRNA Phe, tRNA Val, tRNA Leu, tRNA Ile, tRNA Gln, tRNA Met, tRNA Trp, tRNA Ala, tRNA Asn , tRNA Cys , tRNA Tyr , tRNA Ser , tRNA Asp , tRNA Lys tRNA Gly, tRNA Arg, tRNA His, tRNA Ser, tRNA Leu , tRNA Thr , tRNA Pro, and tRNA Glu. Apart from these control region or D loop region was also included. Complete mitochondrial sequences were inspected manually and edited using sequence editor option of GENETOOL and the sequences corresponding to the different genes were isolated and saved. Subsequently, the sequences were aligned using CLUSTALX (Thompson *et al.*, 1994), website (http://www.cbi.ac.uk/clustalw/). The nucleotide sequence variability was studied using the computer program Molecular Evolutionary Genetic Analysis (MEGA Version 4.0). Probable SNP sites were identified through manual inspection of the aligned sequences.

Table 3.1. Details of accession numbers used

Species	Subspecies	Breeds	Accession number
Gallus gallus	*G. g. gallus*	Red jungle fowl	AP003322
	G. g. gallus	Cochin Chinese red jungle fowl	NC_007236
	G. g. bankiva	Javan red jungle fowl	AP003323
	G.g. spadiceus	Burmese Red jungle fowl	AP003321
	G.g. domesticus	White Leghorn	AP003317
	G.g. domesticus	White Plymouth Rock	AP003318
	G.g. domesticus	White Leghorn	AP003580
	G.g. domesticus	New Hampshire	AY235570

3.2. SNP genotyping

3.2.1. Using PCR-RFLP

The restriction enzyme maps were developed for the sequences where the probable SNPs were identified by using GENETOOL software. These maps were studied for identifying the restriction enzymes, whose sites were polymorphic at the particular SNP site. No such restriction enzyme could be identified, hence the PCR-RFLP was not used for SNP genotyping.

3.2.2. Using Allele Specific PCR

3.2.2.1. Resource population

A total of 18 individuals from Indian red jungle fowl and 18 individuals from White Leghorn population, representing domestic fowl were used for SNP genotyping. The DNA was obtained from the Genome Mapping Laboratory, CARI, Izatnagar.

3.2.2.2. Designing of the primers

The primers were designed for Allele Specific PCR using GENETOOL computer software. For six SNPs, the primers with good score could be designed. The details of these primers have been presented in Table 3.2.

3.2.2.3. PCR amplification

The PCR amplification carried out for SNP genotyping was a multiplex PCR using the general primers as well as the allele specific primers for each SNP and for each population.

The PCR amplification was carried out in a final volume of 25 µl reaction mixtures in 0.2 ml thin wall PCR tubes. Each PCR tube containing 25-50 ng genomic DNA, 1.5 mM $MgCl_2$. 50 mM Kcl, 10 mM Tris-Hcl, pH=8.8, 0.1% Triton X-100, 0.01% gelatin, 200 µM of each dNTP (dATP, dGTP, dCTP and dTTP), 1.0 unit of Taq DNA polymerase enzyme and 10 pico mole of each forward and reverse primers from each set.

The amplification was carried out in an i-cycler (Biorad). Protocol for PCR reaction consisted of an initial denaturation at 94°C for 5 min. Followed

Table 3.2. Details of the primers used for SNP genotyping

SNP	Type of PCR	Primer Type	Primer Code	Primer Sequence	Size of amplicon
COI-114-C/T	General PCR	Forward	COI-F	ggcttcgatcccatacaattt	603
		Reverse	COI-R	cggctaaagggggtaaactgt	
	Allele Specific PCR	Forward	COI-F	ggcttcgatcccatacaattt	334
		Reverse	COI-114-R-RJF	ggctgtcctagttctgcgcgg	
		Reverse	COI-114-R-RJF	ggctgtcctagttctgcgcga	
COI-255-G/A	General PCR	Forward	COI-F	ggcttcgatcccatacaattt	603
		Reverse	COI-R	cggctaaagggggtaaactgt	
	Allele Specific PCR	Forward	COI-F	ggcttcgatcccatacaattt	475
		Reverse	COI-255-R-RJF	tctggggcaccgattataagc	
		Reverse	COI-255-R-RJF	tctggggcaccgattataagt	
COII-279-C/T	General PCR	Forward	COII-F	ggccaaccactcccaactag	523
		Reverse	COII-R	agggattgcgtctgtttttacc	
	Allele Specific PCR	Forward	COII-279-F-RJF	agacgaaatcgacgaacctgat	268
		Forward	COII-279-F-DF	agacgaaatcgacgaacctgac	
		Reverse	COII-R	agggattgcgtctgtttttacc	

ATP6-294-G/A	General PCR	Forward	ATP6-F	atccctctaatcctcccatcact	496
		Reverse	ATP6-R	gggcgattgtggctgtagaga	
	Allele Specific PCR	Forward	ATP6-294-F-RJF	ccttcaccccaactacccaact	270
		Forward	ATP6-294-F-DF	ccttcaccccaactacccagct	
		Reverse	ATP6-R	gggcgattgtggctgtagaga	
ND4-969-C/T	General PCR	Forward	ND4-F	gcccatgtagaagcaccaatc	521
		Reverse	ND4-R	gtgggggaggatcagttgaagag	
	Allele Specific PCR	Forward	ND4-F	gcccatgtagaagcaccaatc	335
		Reverse	ND4-969-R-RJF	gctaggcagaataaaagggaaga	
		Reverse	ND4-969-R-DF	gctaggcagaataaaagggagga	
CYTB-543-T/C	General PCR	Forward	CYTB-F	cccacacttgccggaacgtac	514
		Reverse	CYTB-R	ggggtgagtatgagagttaagcc	
	Allele Specific PCR	Forward	CYTB-F	cccacacttgccggaacgtac	363
		Reverse	CYTB-543-R-RJF	caaaggggaggaggaagtgtaag	
		Reverse	CYTB-543-R-DF	caaaggggaggaggaagtgtaaa	

by 35 cycles of PCR, each cycle consisting of 1 min at 94°C, 1 min at 55 °C and 2 min at 72°C, and followed by a final extension step of 10 min at 72 °C.

3.2.2.4. Resolution of the amplified products and documentation

The PCR products were resolved on 1.6% agarose gel in 1x TBE. Electrophoresis was done at 90 volts for 10 minutes, then at 50 volts for 2 hour. Gel was viewed under a UV Trans-illuminator and photographed with gel documentation system (Minibis) for future analysis. Molecular size of the amplified product was estimated by using 100 bp DNA ladder (MBI Fermentas).

3.2.2.5. Statistical analysis

The allelic frequency for population specific allele was estimated as the proportion of that allele in the sample population.

3.3. Polymorphism in D loop region

3.3.1. Amplification of complete D loop region in Indian Red Jungle Fowl

The primers were designed to amplify the complete D loop region. The forward primer was taken from tRNA-Glu (5'-AGG ACT ACG GCT TGA AAA GC-3'), while the reverse primer was taken from tRNA-Phe (5'-CAT CTT GGC ATC TTC AGT GCC A-3').

PCR reactions were set up in 25 µl reaction volume containing 2.5 µl of 10 X Assay buffer (100 mM Tris- HC1, pH 9.0, 15 mM $MgCl_2$, 500mM KC1 and 0.1% gelatin), 200 µM of dNTP mix, 10 pM of forward and reverse primer, 1U Taq DNA polymerase, 50 ng of genomic DNA and autoclaved milliQ water to make up the volume. The amplification was carried out in an i-cycler (Biorad)). Protocol for PCR reaction consisted of an initial denaturation at 94°C for 5 min. followed by 35 cycles of PCR, each cycle consisting of 1 min at 94°C, 1 min at 55 °C and 2 min at 72°C, and followed by a final extension step of 10 min at 72 °C.

The PCR products were resolved on 1.6% agarose gel in 1x TBE. Electrophoresis was done at 90 volts for 10 minutes, then at 50 volts for 2 hour. Gel was viewed under a UV Trans-illuminator and photographed with gel

documentation system (Minibis) for future analysis. Molecular size of the amplified product was estimated by using 100 bp DNA ladder plus (MBI Fermentas).

3.3.2. Gel purification of PCR products

The 1325 bp PCR products were purified from gel using QIAquick Gel Extraction Kit (QIAGEN Inc. Valencia, CA, USA) as follows:

- A preparative gel was prepared in TAE buffer and electrophoresis was done at 30 V for 6 hours. The gel was briefly visualized in low range UV light and the desired band was cut using a sterile paragon blade. The gel slice was collected in a pre-weighed sterile 1.5 ml micro centrifuge tube.
- The gel slice was weighed and 3 volumes of buffer QG was added to 1 volume of gel slice and incubated at 50°C for 10 min. with intermittent vortexing every 2-3 min.
- After the gel slice was completely dissolved, isopropanol equivalent to gel volume was added and mixed.
- Then, the sample was applied to the QIAquick spin column, placed in 2 ml collection tube provided in the kit.
- QIAquick spin column was centrifuged for 1 min at 12000 rpm.
- The flow-through was discarded and the QIAquick spin column was placed back in the same collection tube.
- 500 µl of buffer QG was added to QIAquick spin column and centrifuged for another 1 min at 12000 rpm.
- For washing, 750 µl of Buffer PE was added to column and then centrifuged at 12000 rpm for 1min.
- The flow-through was discarded and the column was placed back in the same collection tube and centrifuged at 12000 rpm for additional 2 min.
- The column was then transferred to a clean 1.5 ml micro centrifuge tube.
- For elution of the Plasmid DNA, 30 µl of buffer EB was added to the center of the column, kept as such for 5 min and then centrifuged at 12000 rpm for 1 min.

- 5 µl of the purified PCR product was run in 1.6% agarose gel in 1x TBE along with 100 bp molecular weight marker (Gene ruler 100 bp ladder, MBI Fermentas).

3.3.3. Cloning of the purified PCR products in pTZ57R/T vector

3.3.3.1. Preparation of ligation reaction

The gel purified PCR products were cloned into pTZ57R/T vector system (MBI Fermentas), according to the manufacturer's protocol. The ligation reaction was prepared in 10 µl reaction in a 0.5ml microcentrifuge tube:

5X Rapid ligation buffer	2.0 µl
pTZ57R/T Vector (50 ng)	1.0 µl
Purified PCR product (100ng)	5.0 µl
T4 DNA ligase (5 U)	1.0 µl
Nuclease free water	2.0 µl

The reaction mixture was incubated at 4°C overnight. The ligated DNA was diluted to 200 µl in 1.1x TCM solutions and kept at 4°C until used for transformation.

3.3.3.2. Preparation of competent cells and transformation

In order to transform DH5α cells (E. coli cells) with the ligated DNA, fresh competent cells were prepared. The steps as follows:

- Fresh cultures of DH5α cells were grown in 25ml LB medium in shaking incubator at 37°C overnight.
- The cells were diluted 200 times in LB medium and incubated in shaking incubator at 37°C for about 2-3 hrs until culture Attained an O.D. of 0.3-0.4 at A600 nm.
- The cells were kept on ice for 20 min.
- The cells were pelleted at 6000 rpm for 10 min at 4°C and the pellet was re-suspended in 1/10 volume of chilled TSS solution and incubated on ice for 1 hour.
- Competent cells were then ready for transformation and kept at –20°C till further use.

- 200 µl of competent cells were gently added to the 200µl ligated DNA (diluted in TCM), mixed and incubated on ice for 20 min.
- The DNA-competent cells mixture were subjected to heat shock at 42°C for 45-50 sec. and were immediately snap-chilled on ice for 2 min.
- 950 µl SOC medium was added to the transformants and incubated at 37°C for 1.5 hrs with shaking.
- Transformants were then plated on LB/Amp/X-gal/IPTG plates and incubated overnight at 37°C.
- The white colonies were picked on the next day and grown in 10 ml LB/ Amp broth for overnight at 37°C with shaking for plasmid extraction.

3.3.3.3. Screening of recombinant clones

In order minimize the number of clones to be handled, clones were initially checked by performing the colony.

3.3.3.3.1. Preparation of master plate

To ensure in situ maintenance of the clones until the analysis is over a master plate was prepared with LB agar with ampicillin and marked appropriately for identification of the colonies. The single white colony was picked using a sterile toothpick. For each colony, after picking the colony, the respective toothpick was first touched onto the marked place of the master plate and then dipped into pre-numbered PCR tube containing MiIIiQ water. Finally, the toothpick was immersed into the test tube containing LB broth plus ampicillin. The master plate was incubated at 37°C for overnight and the LB broth was cultured overnight at 37°C with shaking. The PCR tube was subjected to colony PCR.

3.3.3.3.2. Colony PCR

The toothpick, after touching the master plate was transferred to pre-numbered PCR tube containing 10 µI of autoclaved MilliQ water. Using pipette the colonies were dissolved properly. Colony PCR was performed using the same PCR protocol as used for the amplification of the D loop region. PCR was performed in 25 µl reaction volume. The amplified PCR product (5 µI) was run

in 1.6% agarose gel in 1x TBE along with 100 bp molecular weight marker (Gene ruler 100 bp ladder, MBI Fermentas). The amplified product was visualized, documented and its size was estimated. All the clones showed the amplification and were propagated for further analysis.

3.3.3.4. Isolation of plasmid DNA from PCR-positive clones

As mentioned earlier, following the master plate preparation, the tooth pick was immersed in the 50 ml test tubes containing 10 ml LB broth with the antibiotic ampicillin (100 mg/ml) (@ 1 µl/ml of broth). The tubes were incubated at 37°C in the shaker incubator at 150 rpm for overnight. A glycerol stock of the positive clones (bacterial cultures) was prepared and stored at -70°C.

From these overnight grown cultures the plasmid DNA was extracted following the small scale alkali lysis method as described by Sambrook et al. (1989) as follows:

- Bacterial culture (1.5 ml twice in the same tube) was centrifuged at 13000 rpm (12000 g) at 4°C for 1 min and supernatant was discarded thoroughly.
- The pellet was re-suspended in 100 µl of ice-cold Solution I by vigorous vortexing. Then 200 µl of freshly prepared Solution II (RT) was added to the suspension and mixed gently by inversion. Next, 150 µl of ice-cold Solution III was added and inverted gently for mixing. The viscous lysate was kept on ice for 3-5 min.
- The bacterial lysate was centrifuged at 13000 rpm (12000 g) at 4°C for 5 min. The supernatant was transferred to a new microcentrifuge tube.
- Equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) was added and mixed well by vortexing. The emulsion was centrifuged at 13000 rpm at 4°C for 2 min and the upper aqueous layer was transferred to a new microfuge tube.
- Equal volume of chloroform: isoamyl alcohol (24:1) was added and centrifuged at 13000 rpm at 4°C for 2 min and the upper aqueous layer was transferred to a new microfuge tube.

- The nucleic acid was precipitated by adding two volumes of ethanol at RT and incubated for 2-5 min at RT after brief vortexing. Then it was centrifuged at 13000 rpm at 4°C for 5 min.
- The supernatant was discarded thoroughly by inverting the tube onto a paper towel and 1 ml of 70% ethanol was added to the pellet. The tube was inverted several times for proper mixing and then centrifuged at 13000 rpm (12000 g) at 4°C for 2 min.
- The supernatant was removed completely and the residual ethanol was evaporated by keeping the tube open at RT for few min.
- The plasmid was dissolved in 25 µl TE buffer (pH 8.0) containing 20 µg/ml DNase-free RNase A and briefly vortexed and incubated at RT for 30 min.
- The quality of the plasmid DNA was checked in 1% agarose gel. The plasmid was then stored at -20°C until further use.

3.3.3.5. Restriction enzyme analysis of plasmid DNA

The plasmid DNA was double-digested with Restriction enzymes, EcoR I (G?AA TIC) and Pst I (CTGCA?G) having cutting sites on either side of the MCS in the cloning vector pTZ57R/T (as depicted in the Fig 3.1. Accordingly, RE digestion reaction was set up in 1.5 ml microcentrifuge tube using EcoR I and Pst I (MBI Fermentas) as follows:

Plasmid DNA	3.0 µl
Tango Buffer (2X)	4.0µl
EcoRI (10 U/µl)	0 .5 µl
PstI (10 U/µl)	0.5 µl
Nuclease free water	12 µl

The contents were gently mixed by pipetting and incubated in a waterbath at 37°C for 5-6 hrs. The RE digestion was studied by running the 5 µl aliquot of the reaction on 1% agarose gel along with the 100 bp molecular weight marker (100 bp ladder, MBI Fermentas).

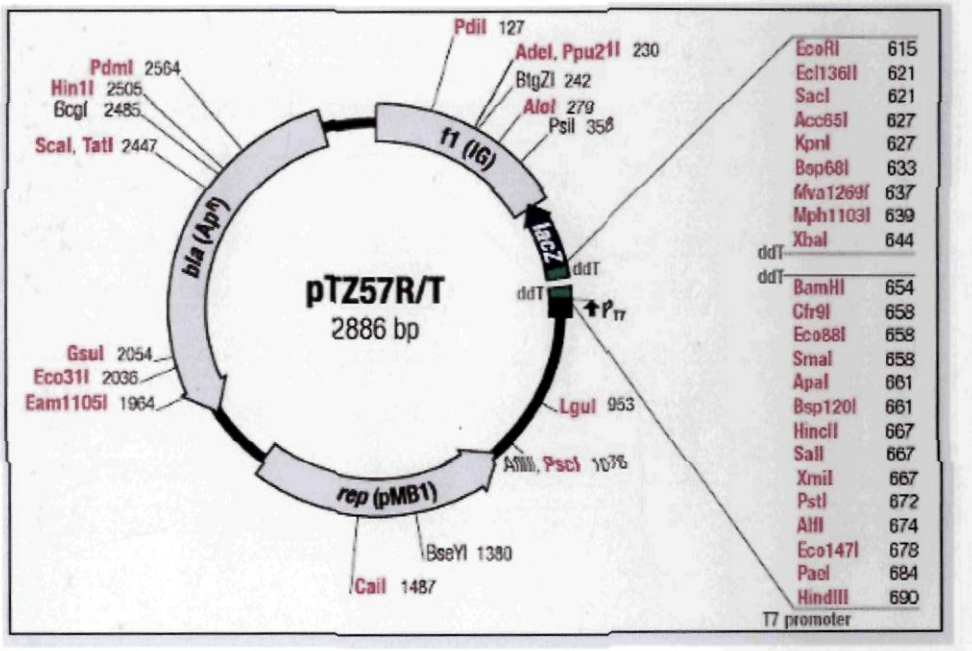

Fig 3.1. The pTZ57R/T vector

3.3.4. Sequencing of complete D loop in Indian RJF

Sequencing was done at DNA sequencing facility, Department of Biochemistry, University of Delhi South campus, New Delhi (UDSC) by automated sequencer using Sanger's dideoxy chain termination method. The positive clones were submitted in the form of the stab culture. Sequencing was done 5'-3' as well as 3'-5' using the M13 Forward and M13 reverse primers, respectively.

3.3.5. Sequence homology comparisons

The related sequences (Table 3.3) were obtained from Genebank (www.ncbi.nlm.nih.gov). The sequences were edited by using GENETOOL software to get comparable sequences. Subsequently, the sequences were aligned using CLUSTALW (Thompson *et al.*, 1994), website (http://www.cbi.ac.uk/clustalw/). Jukes-Cantor genetic distances were estimated using the computer program Molecular Evolutionary Genetic Analysis (MEGA Version 4.0). Jukes-Cantor estimates were used because all mitochondrial sequences were not much divergent (< 6 % divergence from raw counts) and no strong transition bias was evident. Phylogenetic trees were constructed with neighbour joining (NJ) procedure using MEGA Version 4.0.

3.4 Polymorphism between first 400 nucleotides of D loop region in *Gallus gallus* species

The related sequences (Table 3.4) were obtained from Genebank (www.ncbi.nlm.nih.gov) and from published literature. The sequences were edited by using GENETOOL software to get comparable sequences. Subsequently, the sequences were aligned using CLUSTALW (Thompson et al., 1994), website (http://www.cbi.ac.uk/clustalw/). Jukes-Cantor genetic distances were estimated using the computer program Molecular Evolutionary Genetic Analysis (MEGA Version 4.0). Jukes-Cantor estimates were used because all mitochondrial sequences were not much divergent (< 6 % divergence from raw counts) and no strong transition bias was evident. Phylogenetic trees were constructed with neighbour joining (NJ) procedure using MEGA Version 4.0.

Table 3.3. Details of accession numbers used

Species/ Subspecies	Breeds	Common name	Accession number	Reference
G. g. murghi	-	Indian Red Jungle Fowl	-	Present work
G. varius	-	Green jungle fowl	AP003324	Nishibori *et al*. (2005)
G. lafayettei	-	Ceylon jungle fowl	AP003325	-do-
G. sonneratii	-	Grey jungle fowl	AP006741	-do-
G. g. gallus	-	Cochin Chinese red jungle fowl	AP003322	-do-
G. g. bankiva	-	Javan red jungle fowl	AP003323	-do-
G. g.spadiceus	-	Burmese Red jungle fowl	AP003321	-do-
G.g. domesticus	White Leghorn (WLH)	Domestic fowl	AP003317	Nishibori *et al.* (2003)
G.g. domesticus	White Plymouth Rock (WPR)	Domestic fowl	AP003318	Nishibori *et al*. (2003)
G.g. domesticus	Native Laos (NL)	Domestic fowl	AP00319	Nishibori *et al*.(2005)

Table 3.4. Localities of the different samples from subspecies of Red Jungle Fowl

Subspecies / Code	Place of Collection	Reference
Gallus gallus murghi		
Our-RJF	Pilibhit, UP, India	Present work
Gallus gallus bankiva		
*GG_bankiva*15	Singaraja, Bali, Indonesia	Fumihito *et al.* (1996)
*GG_bankiva*18	West Java, Indonesia	Fumihito *et al.* (1996)
*GG_bankiva*15	Lanpung, East Sumatra, Indonesia	Fumihito *et al.* (1996)
Gallus gallus gallus		
*GG_gallus*11	Tama Zoological Garden, Tokyo, Japan	Fumihito *et al.* (1996)
*GG_gallus*39	Palembang, East Sumatra, Indonesia	Fumihito *et al.* (1996)
*GG_gallus*58	Palembang, East Sumatra, Indonesia	Fumihito *et al.* (1996)
*GG_gallus*8	Deptt of Forestry, Thailand	Fumihito *et al.* (1996)
*GG_gallus*10	Deptt of Forestry, Thailand	Fumihito *et al.* (1996)
*GG_gallus*3322	Manila, Philippines	Nishibori *et al.*, 2005
Gallus gallus spadicus		
*GG-spadiceus*3	Deptt of Forestry, Thailand	Fumihito *et al.* (1996)
*GG-spadiceus*4	Deptt of Forestry, Thailand	Fumihito *et al.* (1996)
*GG-spadiceus*5	Deptt of Forestry, Thailand	Fumihito *et al.* (1996)

Results

4.1. Identifying the probable SNPs in mitochondrial genome

A total of eight complete mitochondrial genome sequences (Table 3.1) were used. Among these, four sequences were from *G. g. domesticus* (two from White Leghorn, one from New Hampshire and one from White Plymouth Rock). Among the other four sequences, two were from *G. g. gallus*, one from *G. g. spadiceus* and one from *G. g. bankiva*. These sequences were compared using MEGA 4. The size of the different genes and nucleotide sequence variability has been presented in Table 4.1 and 4.2.

4.1.1. D-Loop

For D loop, size variation was observed within *G. gallus* subspecies. While in *G. g. domesticus*, it ranged from 1231 to 1232 bp, but in other *G. gallus* subspecies, it was 1232. This size variation was due to deletion of one nucleotide in WLH(1) and WPR. Apart from, this deletion, 16 nucleotide substitutions were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these substitutions were transitions. Three of the nucleotide substitutions i.e. T by C at 217 nt, C by T at 446 nt as well as 1214 nt positions were specific to *G. g. domesticus*.

4.1.2. tRNA-Phenylalanine

For tRNA Phe, no size variation (69 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

Table 4.1. Size variation for different mitochondrial genes/ rRNA/ tRNA / D loop region in domestic and red jungle fowl.

Acc. No	AP00 3317	AP00 3318	AP00 3580	AY23 5570	AP00 3322	NC_00 7236	AP00 3323	AP00 3321
Breed/ name	WLH1	WPR	WLH2	NH	RJF	CoCh-RJF	Jav-RJF	Brm-RJF
D-loop	1231	1231	1232	1232	1232	1232	1232	1232
tRNA-Phe	69	69	69	69	69	69	69	69
12S rRNA	976	976	976	975	976	976	976	976
tRNA-Val	73	73	73	73	73	73	73	73
16S rRNA	1626	1623	1625	1621	1622	1622	1622	1622
tRNA-Leu1	74	74	74	74	74	74	74	74
ND1	975	975	975	975	975	975	975	975
tRNA-Ile	72	72	72	72	72	72	72	72
tRNA-Gln	71	71	71	71	71	71	71	71
tRNA-Met	69	69	69	68	69	69	69	69
ND2	1041	1041	1041	1041	1041	1041	1041	1041
tRNA-Trp	76	76	76	76	76	76	76	76
tRNA-Ala	69	69	69	69	69	69	69	69
tRNA-Asn	73	73	73	73	73	73	73	73
tRNA-Cys	66	66	66	66	66	66	66	66
tRNA-Tyr	71	71	71	71	71	71	71	71
COI	1551	1551	1551	1551	1551	1551	1551	1551
tRNA-Ser	75	75	75	75	75	75	75	75

tRNA-Asp	69	69	69	69	69	69	69	69
COII	684	684	684	684	684	684	684	684
tRNA-Lys	68	68	68	68	68	68	68	68
ATPase8	165	165	165	165	165	165	165	165
ATPase 6	684	684	684	684	684	684	684	684
COIII	784	784	784	787	786	786	786	786
tRNA-Gly	68	68	68	68	68	68	68	68
ND3	352	352	352	352	352	352	352	352
tRNA-Arg	68	68	68	68	68	68	68	68
ND4L	297	297	297	297	297	297	297	297
ND4	1378	1378	1378	1378	1380	1380	1380	1380
tRNA-His	69	69	69	69	69	69	69	69
tRNA-Ser	65	65	65	65	65	65	65	65
tRNA-Leu2	71	71	71	71	71	71	71	71
ND5	1818	1818	1818	1818	1818	1818	1818	1818
Cytb	1143	1143	1143	1143	1143	1143	1143	1143
tRNA-Thr	69	69	69	69	69	69	69	69
tRNA-Pro	70	70	70	70	70	70	70	70
ND6	522	522	522	522	522	522	522	522
tRNA-Glu	68	68	68	68	68	68	68	68

Table 4.2. Polymorphic sites in different mitochondrial genes/ rRNA/ tRNA / D loop region between as well as within in domestic and red jungle fowl.

	AP00 3317 WLH1	AP00 3318 WPR	AP00 3580 WLH2	AY23 5570 NH	AP00 3322 RHF	NC_00 7236 CoCh-RJF	AP00 3323 Ja-RJF	AP00 3321 Brm-RJF
D-loop								
212	G	G	G	G	G	G	G	A
217	C	C	C	C	T	T	T	T
243	C	C	C	C	C	C	C	T
246	C	C	C	C	C	C	C	T
256	C	C	C	C	C	C	C	T
261	T	T	T	T	T	T	T	C
281	A	A	A	A	A	G	G	G
296	C	C	C	C	T	T	C	C
307	T	T	T	T	C	C	C	T
310	T	T	T	T	T	T	T	C
315	C	C	C	C	C	C	C	T
342	A	A	A	A	G	G	G	A
446	T	T	T	T	C	C	C	C
686	G	G	G	G	A	A	G	G
859	-	C	-	C	C	C	C	C
1214	T	T	T	T	C	C	C	C
1215	G	G	G	G	G	G	G	A

tRNA-Phe								
	-	-	-	-	-	-	-	-
12S rRNA								
1	A	A	-	A	A	A	A	A
398	C	C	C	C	C	C	C	A
721	G	G	G	A	G	G	G	G
726	A	G	G	G	G	G	G	G
908	T	T	T	T	T	T	T	C
tRNA-Val								
	-	-	-	-	-	-	-	-
16S rRNA								
12	C	-	-	-	-	-	-	-
246	G	G	G	G	A	A	A	G
1488	A	A	A	A	A	A	A	G
1592	C	C	C	C	T	T	T	T
1597	C	C	C	-	C	C	C	C
1598	C	C	C	-	-	-	-	-
1599	C	-	C	-	-	-	-	-
1600	C	-	C	-	-	-	-	-
tRNA-Leu1								
	-	-	-	-	-	-	-	-
ND1								
236	T	T	T	T	C	C	C	T
531	A	A	A	A	G	G	G	G
717	A	A	A	A	G	G	G	A

tRNA-Ile								
	-	-	-	-	-	-	-	-
tRNA-Gln								
	-	-	-	-	-	-	-	-
tRNA-Met								
	-	-	-	-	-	-	-	-
ND2								
478	T	T	T	T	G	G	G	T
586	T	T	T	T	T	T	T	C
tRNA-Trp								
	-	-	-	-	-	-	-	-
tRNA-Ala								
	-	-	-	-	-	-	-	-
tRNA-Asn								
	-	-	-	-	-	-	-	-
tRNA-Cys								
	-	-	-	-	-	-	-	-
tRNA-Tyr								
	-	-	-	-	-	-	-	-
COI								
114	C	C	C	C	T	T	T	T
156	C	C	C	C	T	T	T	T
255	G	G	G	G	A	A	A	A
1429	C	C	C	C	T	T	T	T
1542	G	G	G	G	G	G	A	G

tRNA-Ser								
	-	-	-	-	-	-	-	-
tRNA-Asp								
	-	-	-	-	-	-	-	-
COII								
134	C	C	C	C	T	T	T	T
279	C	C	C	C	T	T	T	T
303	C	C	C	C	T	T	T	C
tRNA-Lys								
	-	-	-	-	-	-	-	-
ATPase8								
	-	-	-	-	-	-	-	-
ATPase6								
294	G	G	G	G	A	A	A	A
354	A	A	A	A	A	A	A	G
COIII								
150	G	G	G	G	A	A	A	A
176	G	G	G	G	G	G	A	G
339	G	G	G	G	G	G	G	A
785	G	-	-	-	G	-	G	G
786	C	-	-	-	C	-	C	C
787	T	-	-	-	-	-	-	-
tRNA-Gly								
	-	-	-	-	-	-	-	-

ND3								
193	C	C	C	C	T	T	T	T
222	C	C	C	C	C	C	C	T
tRNA-Arg								
50	A	A	A	A	A	G	A	A
ND4L								
176	T	T	T	T	G	G	G	T
183	T	T	T	T	C	C	C	C
242	C	C	C	C	T	T	C	C
260	C	C	C	C	C	C	T	C
ND4								
141	C	C	C	C	C	C	T	C
198	A	A	A	A	G	G	G	A
478	T	T	T	T	C	C	C	C
720	A	A	A	A	G	G	A	A
969	C	C	C	C	T	T	T	T
1195	G	A	G	G	G	G	G	G
1379	-	-	-	-	G	G	G	G
1380	-	-	-	-	C	C	C	C
tRNA-His								
	-	-	-	-	-	-	-	-
tRNA-Ser								
	-	-	-	-	-	-	-	-
tRNA-Leu2								
	-	-	-	-	-	-	-	-

ND5								
177	G	G	A	G	G	G	G	G
489	T	C	T	T	T	T	T	T
804	A	A	A	A	A	A	A	G
1435	T	T	T	T	T	T	C	T
CYTB								
460	A	A	A	G	A	A	A	A
507	T	T	T	T	C	C	C	T
543	T	T	T	T	C	C	C	C
713	C	C	C	C	C	C	T	C
791	T	T	T	T	T	T	T	A
973	C	C	C	C	T	T	C	T
tRNA-Thr								
	-	-	-	-	-	-	-	-
tRNA-Pro								
14	G	G	G	G	A	A	A	A
ND6								
273	C	C	C	C	C	C	T	C
403	G	G	G	G	G	G	G	A
tRNA-Glu								
-	-	-	-	-	-	-	-	

4.1.3. 12S- rRNA gene

The 12SrRNA showed size variation from 975 to 976 bp within *G. gallus* subspecies. This size variation was due to deletion of one nucleotide at 1 nt position in NH breed. Over 976 nucleotides, only four polymorphic nucleotide sites were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these polymorphism was due to nucleotide substitution. Out of these 4 nucleotide substitution, 3 were transitions, while 1 was transversion. None of the nucleotide substitution was specific to *G. g. domesticus*.

4.1.4. tRNA-Valine

For tRNA Val, no size variation (73 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.5. 16S- rRNA gene

For 16S rRNA, size variation was observed within *G. gallus* subspecies. While in *G. g. domesticus*, it ranged from 1621 to 1626 bp, but in other *G. gallus* subspecies, it was 1622. The size variation within *G. g. domesticus* was due to deletion of five nucleotide in NH, deletion of three nucleotides in WPR and deletion of one nucleotide in WLH(2) against the 1626 bp in WLH(1). Apart from, these deletions/ additions, only three nucleotide substitutions were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these substitutions were transitions. One of the nucleotide substitution i.e. T by C at 1592 nt position was specific to *G. g. domesticus*.

4.1.6. tRNA-Leucine (1)

For tRNA Leu, no size variation (74 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.7. NADH dehydrogenase 1

The NDI showed no size variation within *G. gallus* subspecies. Over 975 nucleotides, only 3 polymorphic nucleotide sites were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these polymorphism was

due to nucleotide substitution. All the three nucleotide substitutions were transitions. One nucleotide substitution i.e. G by A at 531 nt position was specific to *G. g. domesticus*.

4.1.8. tRNA-Isoleucine

For tRNA Ile, no size variation (72 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.9. tRNA-Glutamine

For tRNA Gln, no size variation (71 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.10. tRNA-Methionine

For tRNA Met, no size variation (69 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.11. NADH dehydrogenase 2

The ND2 showed no size variation within *G. gallus* subspecies. Over 1041 nucleotides, only 2 polymorphic nucleotide sites were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these polymorphism was due to nucleotide substitution. Out of these 2 nucleotide substitution, 1 was transitions, while 1 was transversion. None of the nucleotide substitution was specific to *G. g. domesticus*.

4.1.12. tRNA-Tryptophan

For tRNA Trp, no size variation (76 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.13. tRNA-Alanine

For tRNA Ala, no size variation (69 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.14. tRNA-Asparagine

For tRNA Asn, no size variation (73 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.15. tRNA-Cysteine

For tRNA Cys, no size variation (66 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.16. tRNA-Tyrosine

For tRNA Tyr, no size variation (71 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.17. Cytochrome Oxidase I

For COI, no size variation was observed within *G. gallus* subspecies. Over 1551 nucleotides, only five polymorphic nucleotide sites were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these polymorphism was due to nucleotide substitution. All the five substitutions were transitions and four of these i.e. T by C at 114, 156 as well as at 1429 nt positions and A by G at 255 nt position were specific to *G. g. domesticus*.

4.1.18. tRNA-Serine(1)

For tRNA Ser, no size variation (75 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.19. tRNA-Aspertic acid

For tRNA Asp, no size variation (69 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.20. Cytochrome Oxidase II

For COII, no size variation was observed within *G. gallus* subspecies. Over 684 nucleotides, only three polymorphic nucleotide sites were observed

between *G. g. domesticus* and other *G. gallus* subspecies and all these polymorphism was due to nucleotide substitution. All the three substitutions were transitions and two of them i.e. T by C at 134 nt as well as 279 nt positions were specific to *G. g. domesticus*.

4.1.21. tRNA-Lysine

For tRNA Lys, no size variation (68 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.22. ATPase8

For ATPase8, no size variation (165 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.23. ATPase6

For ATPase6, no size variation was observed within *G. gallus* subspecies. Over 684 nucleotides, only two polymorphic nucleotide sites were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these polymorphism was due to nucleotide substitution. Both of the substitutions were transitions and one of them i.e. A by G at 294 nt position was specific to *G. g. domesticus*.

4.1.24. Cytochrome Oxidase III

For COIII, size variation was observed within *G. gallus* subspecies. While in *G. g. domesticus*, it ranged from 784 to 787 bp, but in other *G. gallus* subspecies, it ranged from 784 to 786. This size variation was due to addition of 3 nucleotides in NH against the 784 bp in other breeds. Similarly, addition of 2 nucleotides in *G. g. spadiceus*, *G. g. gallus* and *G. g. bankiva* resulted in 786 nt COIII. Apart from these deletions / additions, three nucleotide substitutions were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these substitutions were transitions. The nucleotide substitution A by G at 150 nt position was specific to *G. g. domesticus*.

4.1.25. tRNA-Glycine

For tRNA Gly, no size variation (68 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.26. NADH dehydrogenase 3

The ND3 showed no size variation within *G. gallus* subspecies. Over 352 nucleotides, only 2 polymorphic nucleotide sites were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these polymorphism was due to nucleotide substitution. Both the nucleotide substitutions were transitions and one of them i.e. T by C at 193 nt position was specific to *G. g. domesticus*.

4.1.27. tRNA-Arginine

For tRNA Arg, no size variation (68 bp) was observed within *G. gallus* subspecies, however one nucleotide substitution (Transition) was found between *G. g. domesticus* and other *G. gallus* subspecies. This nucleotide substitution was not specific to *G. g. domesticus*.

4.1.28. NADH dehydrogenase 4L

The ND4L showed no size variation within *G. gallus* subspecies. Over 297 nucleotides, only 4 polymorphic nucleotide sites were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these polymorphism was due to nucleotide substitution. Out of these, 3 were transitions and one was tranversion. One nucleotide substitution i.e. C by T at 183 nt position was specific to *G. g. domesticus*.

4.1.29. NADH dehydrogenase 4

For ND4, size variation was observed within *G. gallus* subspecies. While in *G. g. domesticus*, it was 1378, in other *G. gallus* subspecies, it was 1380. This size variation was due to addition of 2 nucleotides in other *G. gallus* subspecies. Apart from these deletions/ additions, 6 nucleotide substitutions were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these substitutions were transitions. The nucleotide substitutions i.e. C by T at 478 position and T by C at 969 position were specific to *G. g. domesticus*.

4.1.30. tRNA-Histidine

For tRNA His, no size variation (69 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.31. tRNA-Serine(2)

For tRNA Ser, no size variation (65 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.32. tRNA-Leucine(2)

For tRNA Leu, no size variation (71 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.33. NADH dehydrogenase 5

The ND5 showed no size variation within *G. gallus* subspecies. Over 1818 nucleotides, only 4 polymorphic nucleotide sites were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these polymorphism was due to nucleotide substitution. All the nucleotide substitutions were transitions and none of them was specific to *G. g. domesticus*.

4.1.34. Cytochrome B

The CYTB showed no size variation within *G. gallus* subspecies. Over 1143 nucleotides, only six polymorphic nucleotide sites were observed between *G. g. domesticus* and other *G. gallus* subspecies and all these polymorphism was due to nucleotide substitution. Out of these 6 nucleotide substitution, 5 were transitions, while 1 was transversion. One nucleotide substitution i.e. C by T at 543 nt position was specific to *G. gallus domesticus*.

4.1.35. tRNA-Threonine

For tRNA Thr, no size variation (69 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.1.36. tRNA-Proline

For tRNA Pro , no size variation (70 bp) was observed within *G. gallus* subspecies, however one nucleotide substitution i.e. A by G (Transition) was found between *G. g. domesticus* and other *G. gallus* subspecies and it was specific to *G. g. domesticus*.

4.1.37. NADH dehydrogenase 6

The ND6 showed no size variation within *G. gallus* subspecies. Over 522 nucleotides, only 2 polymorphic nucleotide sites were observed between *G. g. domesticus* and other *G. gallus* subspecies and were due to nucleotide substitution. Both the nucleotide substitutions were transitions and not specific to *G. g. domesticus*.

4.1.38. tRNA-Glutamic acid

For tRNA Glu, no size variation (68 bp) was observed within *G. gallus* subspecies and it was found to be completely conserved in all the *G. gallus* subspecies.

4.2. SNP genotyping

A total of 19 population specific nucleotide substitutions were identified. Attempts were made to use these SNPs in developing the allele specific PCR and for PCR-RFLP. None of these substitutions created or abolished the restriction site of the commonly available restriction enzymes. However, for 6 SNPs, primers with good score could be designed (Table 3.2). The primers were designed using GENETOOL software. In general two set of primers were designed for each SNP. The first set i.e. general set, was designed for amplification of the common fragment containing the SNP. Other set of primers was allele specific, having one common primer, which used to be one of the primer of general set of primers, while the second primer of this set used to be allele specific, having the SNP as either the most terminal nucleotide at 3' end or one of the 3 terminal nucleotides at 3' end. Hence, for each SNP, from all the individuals, general set of primers was expected to amplify a common band in all the populations, while the allele specific set of primer was expected to amplify the band from the individuals having population specific nucleotide only.

4.2.1. COI-114-C/T SNP

The forward general primer i.e. COI-F was designed from the 5' end (12 nt to 32 nt) of tRNA asparagines, while the reverse general primer i.e. COI-R was taken from the initial part of COI gene (382 nt-404 nt). This set of primer was expected to amplify a 603 bp fragment from either populations. For allele specific set of primers, the forward primer was the same i.e. COI-F, while the reverse primer specific to DF population i.e. COI-114-R-DF as well as reverse primer specific to RJF population i.e. COI-114-R-RJF were taken from 114 nt to 134 nt. These primers were expected to amplify a 334 bp fragment in either population.

Using COI-F and COI-R primers, a 603 bp fragment was amplified in all the individuals from both the populations. Using COI-F and COI-114-R-DF primers, 17 out of 18 individuals showed amplification of 334 bp bands, while using COI-F and COI-114-R-RJF primers, 16 out of 18 individuals showed amplification of 334 bp bands (Fig 4.1). The allelic frequency of C allele in DF population and T allele in RJF population was found to be 0.94 and 0.89, respectively (Table 4.3).

4.2.2. COI-255-G/A SNP

The forward and reverse primers i.e. COI-F and COI-R were the same as used for COI-114-C/T snp and were expected to amplify a 603 bp fragment from either populations. For allele specific set of primers, the forward primer was the same i.e. COI-F, while the reverse primer specific to DF population i.e. COI-255-R-DF as well as reverse primer specific to RJF population i.e. COI-255-R-RJF were taken from 255 nt to 275 nt. These primers were expected to amplify a 475 bp fragment in either population.

Using COI-F and COI-R primers, a 603 bp fragment was amplified in all the individuals from both the populations. Using COI-F and COI-255-R-DF primers, 15 out of 18 individuals showed amplification of 475 bp bands, while using COI-F and COI-255-R-RJF primers, 16 out of 18 individuals showed amplification of 475 bp bands (Fig 4.2). The allelic frequency of G allele in DF population and A allele in RJF population was found to be 0.83 and 0.89, respectively (Table 4.3).

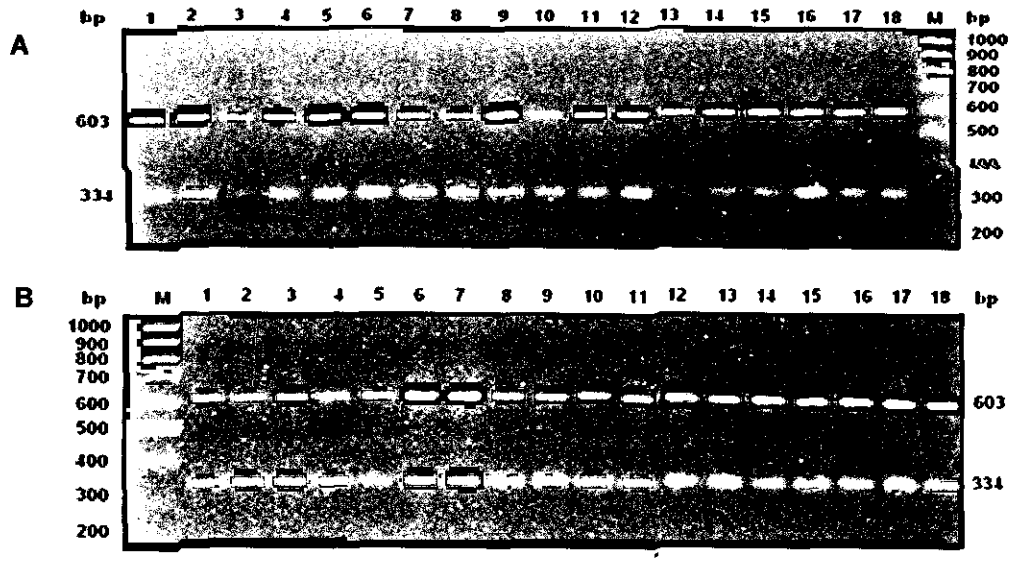

Fig 4.1. COI-114-C/T SNP genotyping using allele specific PCR amplification in red jungle fowl (A) and domestic fowl (B). M : 100 bp ladder.

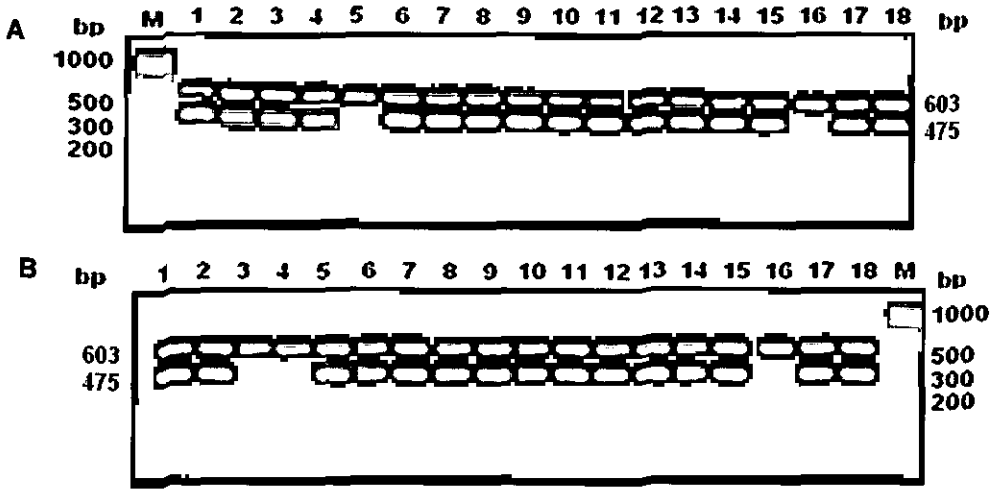

Fig 4.2. COI-255-G/A SNP genotyping using allele specific PCR amplification in red jungle fowl (A) and domestic fowl (B). M : 100 bp ladder.

Table 4.3. Allelic frequency at different SNP positions in RJF and domestic fowl

SNP	Population	Presumed allele	Allelic frequency
COI-114-C/T	Domestic Fowl	C	0.94
	Red Jungle Fowl	T	0.89
COI-255-G/A	Domestic Fowl	G	0.83
	Red Jungle Fowl	A	0.89
COII-279-C/T	Domestic Fowl	C	0.78
	Red Jungle Fowl	T	0.94
ATP6-294-G/A	Domestic Fowl	G	0.89
	Red Jungle Fowl	A	0.83
ND4-969-C/T	Domestic Fowl	C	0.83
	Red Jungle Fowl	T	0.94
CYTB-543-T/C	Domestic Fowl	T	0.94
	Red Jungle Fowl	C	0.78

4.2.3. COII-279-C/T SNP

The forward general primer i.e. COII-F was designed from the 5‘ end (3 nt to 22 nt) of COII gene, while the reverse general primer was taken from the 3‘ part (504 nt- 525 nt). This set of primer was expected to amplify a 523 bp fragment from either populations. For allele specific set of primers, the reverse primer was the same i.e. COI-R, while the forward primer specific to DF population i.e. COII-279-F-DF as well as forward primer specific to RJF population i.e. COII-279-F-RJF were taken from 258 nt to 279 nt. These primers were expected to amplify a 268 bp fragment in either population.

Using COII-F and COII-R primers, a 523 bp fragment was amplified in all the individuals from both the populations. Using COII-279-F-DF and COII-R primers, 14 out of 18 individuals showed amplification of 268 bands, while using COII-279-F-RJF and COII-R primers, 17 out of 18 individuals showed amplification of 268 bp bands (Fig 4.3). The allelic frequency of C allele in DF population and T allele in RJF population was found to be 0.78 and 0.94, respectively (Table 4.3).

4.2.4. ATP6-294-G/A SNP

The forward general primer i.e. ATP6-F was designed from the 5‘ end (49 nt to 71 nt) of ATP6 gene, while the reverse general primer was taken from the 3‘ part (524 nt- 544 nt). This set of primer was expected to amplify a 496 bp fragment from either populations. For allele specific set of primers, the reverse primer was the same i.e. ATP6-R, while the forward primer specific to DF population i.e. ATP6-294-F-DF as well as forward primer specific to RJF population i.e. ATP6-294-F-RJF were taken from 275 nt to 296 nt. These primers were expected to amplify a 270 bp fragment in either population.

Using ATP6-F and ATP6-R primers, a 496 bp fragment was amplified in all the individuals from both the populations. Using ATP6-294-F-DF and ATP6-R primers, 16 out of 18 individuals showed amplification of 270 bp bands, while using ATP6-294-F-RJF and ATP6-R primers, 15 out of 18 individuals showed amplification of 270 bp bands (Fig 4.4). The allelic frequency of G allele in DF population and A allele in RJF population was found to be 0.94 and 0.83, respectively (Table 4.3).

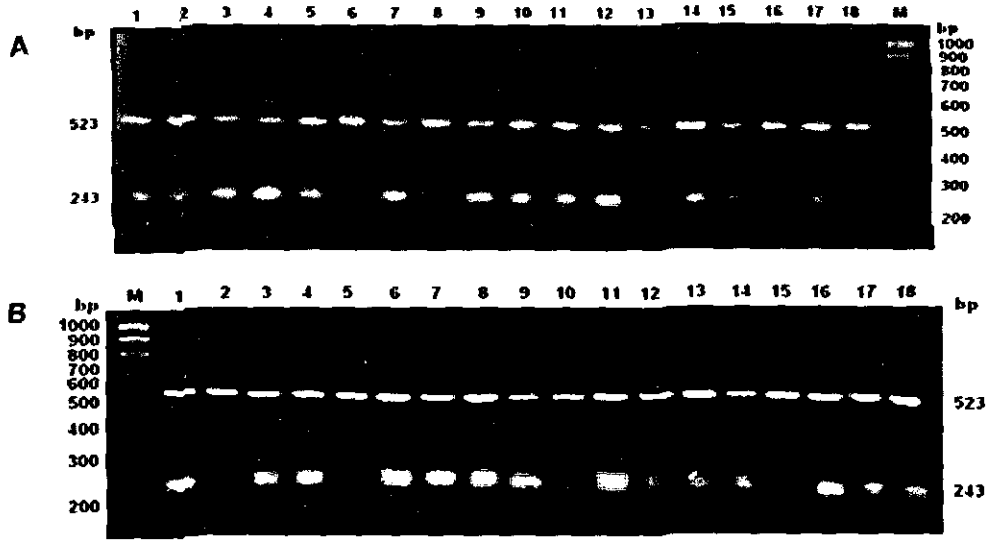

Fig 4.3. COII-279-C/T SNP genotyping using allele specific PCR amplification in red jungle fowl (A) and domestic fowl (B). M : 100 bp ladder.

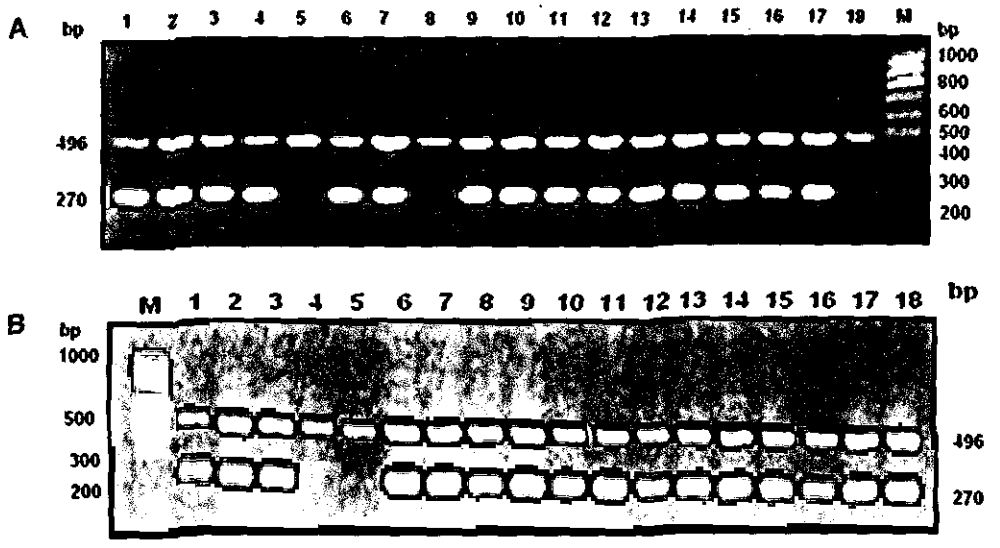

Fig 4.4. ATP6-294-G/A SNP genotyping using allele specific PCR amplification in red jungle fowl (A) and domestic fowl (B). M : 100 bp ladder.

4.2.5. ND4-969-C/T SNP

The forward general primer i.e. ND4-F was designed from the 5‘ end (655 nt to 675 nt) while, the reverse general primer was taken from the 3’ end (1153 nt-1175 nt) of ND4 gene. This set of primer was expected to amplify a 521 bp fragment from either populations. For allele specific set of primers, The forward primer was the same i.e. ND4-F, while the reverse primer specific to DF population i.e. ND4-969-R-DF as well as reverse primer specific to RJF population i.e. ND4-969-R-RJF were taken from 967 nt to 989 nt. These primers were expected to amplify a 335 bp fragment in either population.

Using ND4-F and ND4-R primers, a 521 bp fragment was amplified in all the individuals from both the populations. Using ND4-F and ND4-969-R-DF primers, 15 out of 18 individuals showed amplification of 335 bp bands, while using ND4-F and ND4-969-R-RJF primers, 17 out of 18 individuals showed amplification of 335 bp bands (Fig 4.5). The allelic frequency of C allele in DF population and T allele in RJF population was found to be 0.83 and 0.94, respectively (Table 4.3).

4.2.6. CYTB-543-T/C SNP

The forward general primer i.e. CYTB-F was designed from the 5‘ end (203 nt to 223 nt), while the reverse general primer was taken from the middle part (694 nt-716 nt) of CYTB gene. This set of primer was expected to amplify a 514 bp fragment from either populations. For allele specific set of primers, The forward primer was the same i.e. CYTB-F, while the reverse primer specific to DF population i.e. CYTB-543-R-DF as well as reverse primer specific to RJF population i.e. CYTB-543-R-RJF were taken from 543 nt to 565 nt. These primers were expected to amplify a 363 bp fragment in either population.

Using CYTB-F and CYTB-R primers, a 514 bp fragment was amplified in all the individuals from both the populations. Using CYTB-F and CYTB-543-R-DF primers, 17 out of 18 individuals showed amplification of 363 bp bands, while using CYTB-F and CYTB-543-R-RJF primers, 14 out of 18 individuals showed amplification of 363 bp bands (Fig 4.6). The allelic frequency of T allele in DF population and C allele in RJF population was found to be 0.94 and 0.78, respectively (Table 4.3).

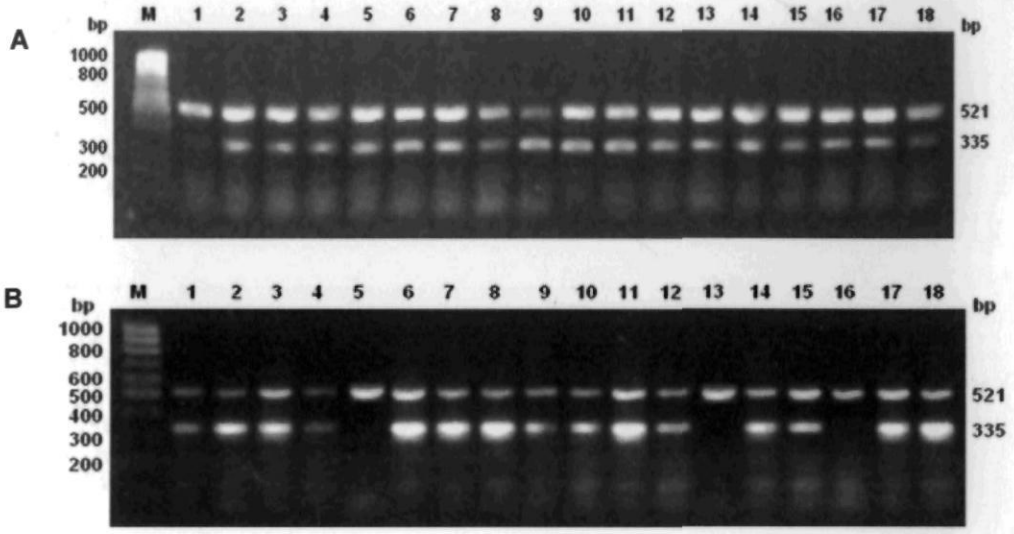

Fig 4.5. ND4-969-C/T SNP genotyping using allele specific PCR amplification in red jungle fowl (A) and domestic fowl (B). M : 100 bp ladder.

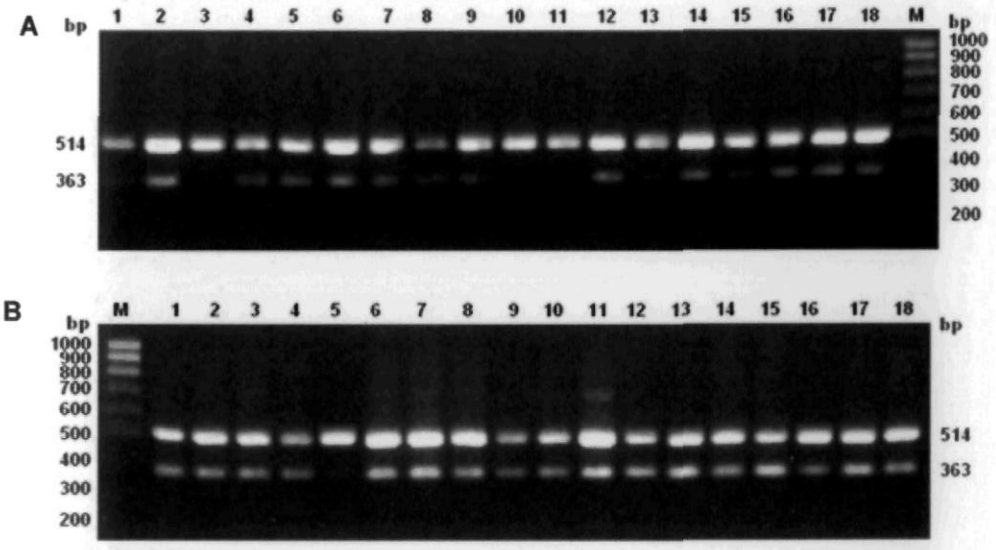

Fig 4.6. CYTB-543-T/C SNP genotyping using allele specific PCR amplification in red jungle fowl (A) and domestic fowl (B). M : 100 bp ladder.

4.3. Polymorphism in D loop region

4.3.1. Amplification of complete D loop region in Indian Red Jungle Fowl

The primers were designed to amplify the complete D loop region. The forward primer was taken from tRNA-Glu (5'-AGG ACT ACG GCT TGA AAA GC-3'), while the reverse primer was taken from tRNA-Phe (5'-CAT CTT GGC ATC TTC AGT GCC A-3'). These primers were expected to amplify a 1325 bp fragment. This 1325 bp fragment is comprised of 45 nucleotide from 3' end of tRNA-Glu, complete nucleotide of D loop region (1232) and 48 nucleotides from 5' end of tRNA-Phe. On resolution of the PCR amplification product on 1.6 % gel showed a single distinct band (Fig 4.7). The size of this fragment was estimated to be 1325 bp by using a molecular size marker. No non-specific amplification showed the specificity of the primers. Absence of any primers or dimmers suggested the good quality of primer designing.

4.3.2. Cloning of 1325 bp fragment

This 1325 bp amplicon was purified from agarose gel using QIAquick Gel Extraction Kit (QIAGEN Inc. Valencia, CA, USA) as per the manufacturer's protocol (Fig 4.8). The purified PCR product was seen as a single and specific band of 1325 bp. This purified product was directly cloned into the TA-cloning vector (e.g., pTZ57R/T vector, MBI Fermentas)) using TA-cloning strategy. Three µl of the ligated DNA was transformed into competent DH5 cells and plated on LB/Amp/X-gal/IPTG plates and incubated over night at 37°C. For primary screening of the positive transformants carrying the insert, blue-white screening was used. White clones were screened for the presence of insert. The clones were initially checked by performing colony PCR and later confirmed by RE digestion of the isolated plasmid.

4.3.2.1. Colony PCR

Four individual white colonies were picked up and subjected to colony PCR using the same set of primers applied for amplification of D loop region. As evident in Fig 4.9, amplification of a single product of 1325 bp was seen with all four colonies Thus the colonies may contain the plasmid with expected insert.

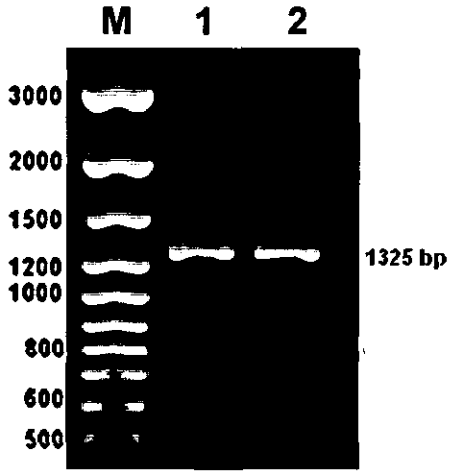

Fig 4.7. Amplification of 1325 bp fragment in RJF (Lane 1-2). M 100 bp ladder plus

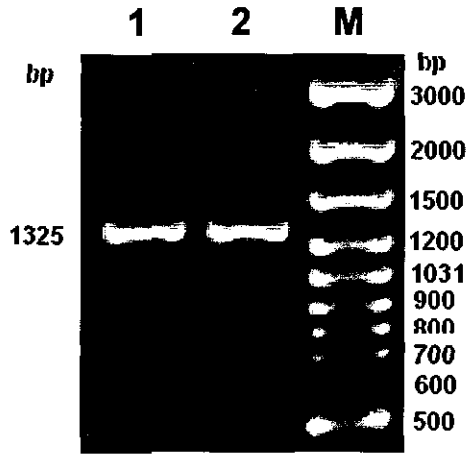

Fig 4.8. Amplification of 1325 bp fragment in RJF from the colony PCR (Lane 1-2. M *100 bp ladder plus*

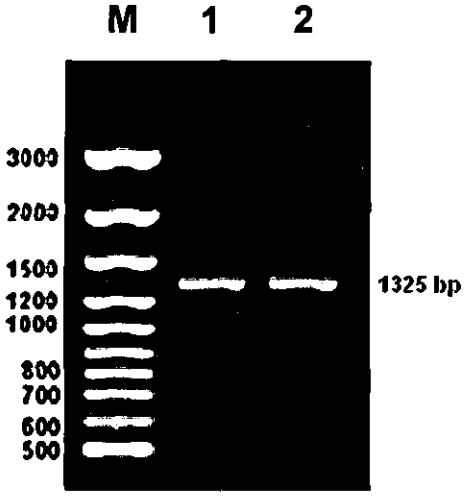

Fig 4.9. Resolution of 1325 bp fragment purified from gel. M 100 bp ladder plus

Fig 4.10. Resolution of 4241 bp fragment representing whole plasmid plus insert (lane 1) and *EcoR* I and *Pst* I digest 1401 bp fragment (insert) along with 2840 bp fragment (*remaining plasmid* DNA) (lane 2) and M : molecular weight marker

4.3.2.2. Restriction enzyme digestion

Consequent upon the fact that the cloning vector pTZ57R/T has one restriction site for *EcoR* I *(G*?*AA* TIC) and one site for *Pst* I (CTGCA?G) on either side of the MCS (as depicted in the Fig 3.1. Accordingly, *EcoR* I and *Pst* I double-digestion of plasmid DNA was done in 10 µl reaction mixture volume. Presence of two bands i.e. 1401 bp and 2840 bp on resolution of restriction digestion mixture on 0.8 % agarose gel confirmed the release of an insert of expected size of 1325 bp in positive clones (Fig 4.10).

4.3.2.3. Recovery of 1325 bp fragment

After confirming the successful cloning of the 1325 bp fragment, attempts were made to retrieve the 1325 bp fragment from plasmid. Restriction enzyme was set up in two reaction mixture volume of 100 µl each. The digestion products were resolved on 1.4 % agarose gel after loading them in larger wells. The 1325 bp fragments were cut from the gel and eluted. A recovery of ~ 85 % was observed.

4.3.3. Sequence homology comparison with other related sequences

The 1325 bp fragment was sequenced on ABI DNA sequencer and the 5'-3' sequence is presented in Fig 4.11. The sequencing revealed the D loop region is 1232 bp in *G. g. murghi.* The related sequence from other Gallus species, *G. gallus* subspecies including *G. g. domesticus* (Table 3.3) were retrieved from the data bank and aligned using ClustalX programme (Fig 4.12).

The length of D loop region showed ranged from 1227 in WPR to 1292 bp in CJF. The size variation was observed within the *G. g. domesticus* subspecies. While in WPR, it was 1227 bp, in WL, NLAO, Silky and Tibetan it was 1231 bp and in NH, it was 1232 bp. In WPR, the reduction in size of D loop is due to deletion of 4 nucleotides at 712 nt, 713 nt, 714 nt and 715 nt positions in comparison the 1231 nt in WL, NLAO, silky and Tibetan. Similarly, addition of 1 nucleotide at 859 nt position in NH resulted in 1232 nt in D loop region. In all the other red jungle fowl subspecies, the size of D loop region was 1232 bp and was due to addition of 1 nucleotide at 859 nt position in comparison to other most of the chicken breeds. Among, other Gallus species, it was 1288 nt in *G. sonnerattii*, 1228 nt in *G. varius* and 1292 nt in *G. lafayettei*. All the three

```
<----------------------------------- Glu ------------------------------------------------------->
aggactacgg cttgaaaagc cattgttgtt ctcaactacg ggaacaattt tatttttaa    60
cctaactccc ctactaagtg tacccccccct ttcccccccca gggggggtat actatgcata  120
atcgtgcata catttatata ccacatatat tatggtaccg gtaatatata ctatatatgt  180
actaaaccca ttatatgtat acgggcatta atctatattc cacatttctc ccaatgtcca  240
ttccatgcat gatccaggac atactcattc accctcccca tagacagctc caaaccacta  300
ccaagtcacc taactatgaa tggttgcagg acataaatct cactctcatg ctcttccccc  360
aacaagtcac ctaactatga atggttgcag gacatacatt taactactat gctctaaccc  420
atttggttat gctcgccgta tcagatggat ttattgatcg tccacctcac gagagatcag  480
caacccctgc ctgtaatgta cttcatgacc agtctcaggc ccattctttc cccctacacc  540
cctcgcccta cttgccttcc accgtgcctc tggttcctcg gtcaggcaca tcccatgcat  600
aactcctgaa ctttctcact tttcacgaag tcatctgtgg attatcttcc cctctttagt  660
ccgtgatcgc ggcatcttct ctcttctatt gctgttggtt ccttctcttt ttggggcttc  720
ttcacaggtt gcccttcaca gtgcgggtgc ggagtgctat tcaagtgaag cctggactac  780
acctgcgttg cgtcctatcc tagtcctctc gtgtccctcg atgagacggt ttgcgtgtat  840
ggggaatcat cttgacactg atgcactttg gatcgcattt ggttatggtt cttccacccc  900
ccccggtaaa tggtgctatt tagtgaatgc ttgtcggaca tatttttatc aattttcact  960
tcctctattt tcttcacaaa actaggaaat tcaccacaat tttttctttg ttatttttta 1020
attttttttt tatttttaa aaacattttt taaaaaacta aattacatac aaactaccgc 1080
ataaaatccc tcaaactata caaacgttta tcgtataata tatatacatt attgtttatt 1140
ctatcattat tagagaaact ccactaccaa aaccatcatt aaaacaaaaa tttacatgcc 1200
           <-------------------------------------------------------
acttaactcc cctcacaaac aatcgttatt tatattgtta attagcaaac acaaaacccg 1260
-------------------------------------Phe---------------------------
ccttctacca ctataaagcc cccatagctt aacccacaaa gcatggcact gaagatgcca 1320
---->
agatg                                                             1325
```

Fig. 4.11. Nucleotide Sequence (5'-3') of 1325 bp fragment containing 45 nucleotide from 3' end of tRNA-Glu, complete nucleotide of D loop region (1232) and 48 nucleotides from 5' end of tRNA-Phe.

Domain I

Hairpin

stop

C- Strech

```
our-rjf  -AATTTTATT TTTTAACCTA ACTCCCCTAC TAAGTGTACC CCCCCTTTCC CCCCCAGGGG GGGTATACTA
AP3317   -......... .......... .......... .......... .......... .......... ..........
AP3318   -......... .......... .......... .......... .......... .......... ..........
AP3319   -......... .......... .......... .......... .......... .......... ..........
AP3321   -......... .......... .......... .......... .......... .......... ..........
AP3322   -......... .......... .......... .......... .......... .......... ..........
AP3323   -......... .......... .......... .......... .......... .......... ..........
AP6741   A......... .......... .......... .......... .......... .......... ..........
AP3324   -......... ........C. .......... .......... .......... .......... ..........
AP3325   -......... .......... .......... .......... .......... .......... ..........
```

ETAS 2

ETAS1

TAS TAS

```
our-rjf  TGCATAATCG TGCATACATT TATATACCAC ATATATTATG GTACCGGTAA TATATACTAT ATATGTACTA
AP3317   .......... .......... .......... .......... .......... .......... ..........
AP3318   .......... .......... .......... .......... .......... .......... ..........
AP3319   .......... .......... .......... .......... .......... .......... ..........
AP3321   .......... .......... .......... .......... .......... .......... ..........
AP3322   .......... .......... .......... .......... .......... .......... ..........
AP3323   .......... .......... .......... .......... .......... .......... ..........
AP6741   ..T....... .......... .......... .......... .......... .......... ..........
AP3324   .......... .......... .......... .......... .......... .......... ..........
AP3325   .......... .......... .......... .......... .......... .......... ..........
```

ETAS 2

```
our-rjf  AACCCATTAT ATGTATACGG GCATTAATCT ATATTCCACA TTTCTCCCAA TGTCCATTCC ATGCATGATC
AP3317   .......... .......... .......... .......... .......... .........T ..........
AP3318   .......... .......... .......... .......... .......... .........T ..........
AP3319   .......... .......... .......... .......... .......... .........T ..........
AP3321   .......... .......... .......... .......... .......... .........T ..........
AP3322   .......... .......... .......... .......... .......... .........T ..........
AP3323   .......... .......... .......... .......... .......... .........T ..........
AP6741   .......... .......... .......C.. .......C.. ........C. .......CT. ...A...G..
AP3324   .......... .......... A......C.. .C.....C.. .......... .......... ...A......
AP3325   .......... .......... .......... .C.....C.. .......... .......CT. ...A...G..
```

Reapeat Sequence

```
                                 stop
our-rjf  CAGGACA-TAC TCATTCACCC TCCCCATAGA CAGCTCCAAA CCACTACCAA GTCACCTAAC TATGAATGGT
AP3317   .......-C.. .......... .......... .......... .......... .......... ..........
AP3318   .......-C.. .......... .......... .......... .......... .......... ..........
AP3319   .......-C.. .......... .......... .......... .......... .......... ..........
AP3321   ..A....-... .......... .......... ...T..T... ......T... .C........ ..........
AP3322   ....... ... .......... .......... .......... .......... .......... ..........
AP3323   .......-... .......... .......... .......... .......... .......... ..........
AP6741   ..A.T..-.-. CT........ CA......C. TG...TAT.T .........G .C........ ..........
AP3324   ..A.T..-.-T A.G.CGT..T A.......TC ..A...T.T. ......A..G .....T.... ..........
AP3325   ....T..A.-. CT..CT.T.T .A......C. T.ATCTTT.. ......A... .......... ..........
```

```
                                                                  Repeat Sequence
our-rjf   TGCAGGACAT AAATCTCACT CTCATGCTCT TCCCCCAACA AG-------- ---------- ----------
AP3317    .A........ .......... ......T... .......... ..-------- ---------- ----------
AP3318    .A........ .......... ......T... .......... ..-------- ---------- ----------
AP3319    .A........ .......... ......T... .......... ..-------- ---------- ----------
AP3321    .A........ .......... ......T... C....T.... ..-------- ---------- ----------
AP3322    .......... ......T... .......... .......... ..-------- ---------- ----------
AP3323    .......... .......... .......... .......... ..-------- ---------- ----------
AP6741    .......... .C.CT.A... T.A....... ...T..C... G.TCACCTAA TCATGAATGA TCATAGGTCA
AP3324    .A........ GCC...A.TA ..AG...... A....T.... ..-------- ---------- ----------
AP3325    .A........ .C....A..C T.A....... ...T.T.... ..CCACCTAA CTATGAATGG TTACAGGACA
```

```
                                                  Repeat Sequence
our-rjf   ---------- ---------- ---------- ---TCACCTAA CTATGAATGG TTGCAGGACA TACATTTAAC
AP3317    ---------- ---------- ---------- ---......... .......... .......... ..........
AP3318    ---------- ---------- ---------- ---......... .......... .......... ..........
AP3319    ---------- ---------- ---------- ---......... .......... .......... ..........
AP3321    ---------- ---------- ---------- ---......... .......... .......... ..........
AP3322    ---------- ---------- ---------- ---......... .......... .......... ..........
AP3323    ---------- ---------- ---------- ---......... .......... .......... ..........
AP6741    TACGTTTAAT TCCAATGCTC GTCCCCTTAC AAG......... .......... .......... ..........
AP3324    ---------- ---------- ---------- ---......... .......... .......... ..........
AP3325    TACATCTAAC CTTAATGCTC TTCCCCTGAC AGG......... .......... .......... ..........
```

Domain II = = =>

```
our-rjf  TTGCAGGACA TACATTTAAC TACT-ATGCC TCTAACCCAT TTGGTTATGC TCGCCGTATC AGATGGATTT
AP3317   ..A....... .......... ...C-....T .......... .......... .......... ..........
AP3318   ..A....... .......... ...C-....T .......... .......... .......... ..........
AP3319   .......... .......... .....C-... T......... .......... .......... ..........
AP3321   ..A....... .......... ...C-....T .......... .......... .......... ..........
AP3322   .......... .......... ...C-....T .......... .......... .......... ..........
AP3323   .......... .......... ...C-....T .......... .......... .......... ..........
AP6741   .CA....... .GA.CC...A .C..T....T .......... .......... ...T...... ..........
AP3324   ..A....... .....C.... ...C-....A .......... .......... ...T....C. .A........
AP3325   ..A....... ....C.CC.A .C..T....T .......... .......... ...T...... ..........
```

F BOX

E BOX

```
our-rjf  ATTGATCGTC CACCTCACGA GAGATCAGCA ACCCCTGCCT GTAATGTACT TCATGACCAG TCTCAGGCCC
AP3317   .......... .......... .......... ........T. .......... .......... ..........
AP3318   .......... .......... .......... ........T. .......... .......... ..........
AP3319   .......... .......... .......... ........T. .......... .......... ..........
AP3321   .......... .......... .......... .......... .......... .......... ..........
AP3322   .......... .......... .......... .......... .......... .......... ..........
AP3323   .......... .......... .......... .......... .......... .......... ..........
AP6741   .........T .......... .......... .......... .......... .......... ..........
AP3324   .......... .......... .......... .........C .......... .......... ..........
AP3325   .........T .......... .......... .......... .......... .......... ..........
```

```
                                                                  D BOX
our-rjf  ATTCTTTCCC CCTACACCCC TCGCCCTACT TGCCTTCCAC CGTGCCTCTG GTTCCTCGGT CAGGCACATC
AP3317   .......... .......... .......... .......... ...A...... .......... ..........
AP3318   .......... .......... .......... .......... ...A...... .......... ..........
AP3319   .......... .......... .......... .......... ...A...... .......... ..........
AP3321   .......... .......... .......... .......... ...A...... .......... ..........
AP3322   .......... .......... .......... .......... ...A...... .......... ..........
AP3323   .......... .......... .......... .......... ...A...... .......... ..........
AP6741   .......... .......... .......... .......... .......... .-........ ..........
AP3324   .......... .......... .......... .......... ...A...... .......... ..........
AP3325   .......... .......... .......... .......... .......... .......... ..........

                                      C BOX
our-rjf  CCATGCATAA CTCCTGAACT TTCTCACTTT TCACGAAGTC ATCTGTGGAT TATCTTCCCC TCTTTAGTCC
AP3317   .......... .......... .......... .......... .......... .......... ..........
AP3318   .......... .......... .......... .......... .......... .......... ..........
AP3319   .......... .......... .......... .......... .......... .......... ..........
AP3321   .......... .......... .......... .......... .......... .......... ..........
AP3322   .......... .......... .......... .......... .......... .......... ..........
AP3323   .......... .......... .......... .......... .......... .......... ..........
AP6741   .......... .......... .......... .......... .......... .......... ..........
AP3324   .......... .......... .......... .......... .......... .......... ..........
AP3325   .......... .......... .......... .......... .......... .......... .T........
```

```
our-rjf  GTGATCGCGG CATCTTCTCT CTTCTATTGC TGTTGGTTCC TTCTCTTTTT GGGGCTTCTT CACAGGTTGC
AP3317   .......... .......... .......... .......... .......... .......... ..........
AP3318   .......... .......... .......... .......... .......... .......... ..........
AP3319   .......... .......... .......... .......... .......... .......... .........A.
AP3321   .......... .......... .......... .......... .......... .......... ..........
AP3322   .......... .......... .......... .......... .......... .......... .........A.
AP3323   .......... .......... .......... .......... .......... .......... ..........
AP6741   .......... .......... .......... .......... .......... .......... ..........
AP3324   .......... .......CT. .....T.... ..G....... .......... .......... ..........
AP3325   .......... ........T. .......... .......... ......C... .......... ..........
```

```
our-rjf  CCTTCACAGT GCGGGTGCGG AGTGCTATTC AAGTGAAGCC TGGACTACAC CTGCGTTGCG TCCTATCCTA
AP3317   .......... .......... .......... .......... .......... .......... ..........
AP3318   ....----.. .......... .......G.. .......... .......... .......... ..........
AP3319   .......... .......... .......... .......... .......... .......... ..........
AP3321   .......... .......... .......... .......... .......... .......... ..........
AP3322   .......... .......... .......... .......... .......... .......... ..........
AP3323   .......... .......... .......... .......... .......... .......... ..........
AP6741   .......... .......... .......... .......... .......... .......... ..........
AP3324   ...A...... .......... ...A...... .......... .......... .......... ..........
AP3325   .......... .......... .......... .......... .......... .......... ..........
```

```
our-rjf  GTCCTCTCGT GTCCCTCGAT GAGACGGTTT GCGTGTATGG GGAATCATCT TGACACTGAT GCACTTTGGA
AP3317   .......... .......... .......... .......... .......... .......... ..........
AP3318   .......... .......... .......... .......... .......... .......... ..........
AP3319   .......... .......... .......... .......... .......... .......... ..........
AP3321   .......... .......... .......... .......... .......... .......... ..........
AP3322   .......... .......... .......... .......... .......... .......... ..........
AP3323   .......... .......... .......... .......... .......... .......... ..........
AP6741   .......... .......... .......... .......... .......... .......... ..........
AP3324   .......... .......... .......... .......... .......... .......... ..........
AP3325   .......... .......... .......... .......... .......... .......... ..........
```

```
                                  ][Domain III= = = ►                    CSB- I
                                     oH
our-rjf  TCGCATTTGG TTATGGTTCT TCCACCCCCC CCGGTAAATG GTGC-TATTT AGTGAATGCT TGTCGGACAT
AP3317   .......... .......... .......... .......... ....-..... .......... ..........
AP3318   .......... .......... .......... .......... ....-..... .......... ..........
AP3319   .......... .......... .......... .......... ....-..... .......... ..........
AP3321   .......... .......... .......... .......... ....-..... .......... ..........
AP3322   .......... .......... .......... .......... ....-..... .......... ..........
AP3323   .......... .......... .......... .......... ....-..... .......... ..........
AP6741   .......... .......... .......... .......... ....C..... .......... ..........
AP3324   .......... ......C... .......... .......... ....-..... .......... ..........
AP3325   .......... .......... .......... .......... ....-..... .......... ..........
910
```

```
          ┐
our-rjf   ATTTTTATCA ATTTTCACTT CCTCTATTTT CTTCACAAAA CTAGGAAATT CACCACAATT TTTTCTTTGT
AP3317    .......... .......... .......... .......... .......... .......... ..........
AP3318    .......... .......... .......... .......... .......... .......... ..........
AP3319    .......... .......... .......... .......... .......... .......... ..........
AP3321    .......... .......... .......... .......... .......... .......... ..........
AP3322    .......... .......... .......... .......... .......... .......... ..........
AP3323    .......... .......... .......... .......... .......... .......... ..........
AP6741    .......... .......... .......... ..C....... ......G... .......... ....T.....
AP3324    ......TATC .A...T.... .......... ..C....... .......... .......... ..........
AP3325    .......... .......... .......... .......... .......... .......... .........G


our-rjf   TATTTTTTAA TTTTTTTTTT ATTTTTTAAA AACATTTTTT AAAAAACTAA ATTACATACA AACTACCGCA
AP3317    .......... .......... .......... .......... .......... .......... ..........
AP3318    .......... .......... .......... .......... .......... .......... ..........
AP3319    .......... .......... .......... .......... .......... .......... ..........
AP3321    .......... .......... .......... .......... .......... .......... ..........
AP3322    .......... .......... .......... .......... .......... .......... ..........
AP3323    .......... .......... .......... .......... .......... .......... ..........
AP6741    .T........ .....A...C .......... .......... .......... ......C.T. ..........
AP3324    .T........ .....A...C G......... .......... .......... ......C.T. ..........
AP3325    .T........ .....A...C .......... .......... .......... ....T.A.T. ..........
```

LSP/HSP

```
our-rjf  TAAAATCCCT CAAACTATAC AAACGTTTAT CGTATAATAT ATATACATTA TTGTTTATTC TATCATTATT
AP3317   .......... .......... .......... .......... .......... .......... ..........
AP3318   .......... .......... .......... .......... .......... .......... ..........
AP3319   .......... .......... .......... .......... .......... .......... ..........
AP3321   .......... .......... .......... .......... .......... .......... ..........
AP3322   .......... .......... .......... .......... .......... .......... ..........
AP3323   .......... .......... .......... .......... .......... .......... ..........
AP6741   .......... ...G...... .......... ......G... .......... ......G..T A.........
AP3324   .......... .......... .......... ......G... .......... ......G..T C.........
AP3325   .......... .......... .......... ......G... .......... ......G..T C.........
```

```
our-rjf  AGAGAAACTC CACTACCAAA ACCATCATT- AAAACAAAAA TTTACATGCC ACTTAACTCC CCTCACAAAC
AP3317   .......... .......... .........- .......... .......... .......... ..........
AP3318   .......... .......... .........- .......... .......... .......... ..........
AP3319   .......... .......... .........- .......... .......... .......... ..........
AP3321   .......... .......... .........- .......... .......... .......... ..........
AP3322   .......... .......... .........- .......... .......... .......... ..........
AP3323   .......... .......... .........- .......... .......... .......... ..........
AP6741   .......... .......... .T.......T .......... ...G..CA.. ........T. ..A.......
AP3324   .......... .......... .T.......- .......... .......A.. ........T. ..A.......
AP3325   .......... .......... .........- .......... ......CA.. ........T. ..CT......
```

```
our-rjf  AATCGTTATT TATATTGTTA ATTAGCAAAC ACAAAACCCG CCTTCTACCA CTATAAA
AP3317   .......... .......... .......... ........T. .......... .......
AP3318   .......... .......... .......... ........T. .......... .......
AP3319   .......... .......... .......... ........T. .......... .......
AP3321   .......... .......... .......... .........A .......... .......
AP3322   .......... .......... .......... .......... .......... .......
AP3323   .......... .......... .......... .......... .......... .......
AP6741   ....A..... .......... ....A..... .......T.A .......... .C.....
AP3324   ....A..... .......... ....A..... .......TT. ..C....... .C.....
AP3325   .......... .......... .......... ........T. .......... .C..T..
```

Fig 4.12. Aligned sequences of the genus Gallus. Dot indicates identity of nucleotides to the reference sequence (Our-RJF) Conserved blocks of putative functional importance, as discussed in the text are high lighted. The putative ends of H strand synthesis at TCCC motifs are indicated as "stop". Sequence corresponding to goose hairpin has been indicated. Stem and loop of this hairpin are under lined with double lines and dotted line, respectively. Sequence similar to mammalian ETAS1 and ETAS 2 are indicated. TAS sequences and GYRCAT motif can produce two hairpins (Stem and loop of this hairpin are under lined with double lines and dotted line, respectively). Repeat sequences has been marked. In the central domain, the sequences similar to F, E, D and C boxes have been indicated. The putative origin of heavy strand replication (O_H) is underlined. The CSB1 sequences has been marked. The LSP and HSP (L,Abbe et al, 1991) is underlined.

Gallus species showed the insertion of 59 nts after 360 nt position in comparison to *G. gallus*. However among the other three jungle fowls, a deletion of 52 nts was observed after 308 nt position in *G. varius*.

Within *G. g. domesticus*, a total of 6 polymorphic sites were observed. Apart from addition and deletions at 4 sites, there were 2 nucleotide substitutions and all were transitions (G/A). Within *G. gallus* subspecies excluding *G. g. murghi*, the polymorphic sites were 13 and all of these were nucleotide substitutions. All these substitutions were transitions (8 T/C and 5 G/A).

Between *G. g. murghi* and different chicken breeds, a total of 14 polymorphic sites were observed. While 5 sites were due to additions and deletions of nucleotides, 9 were due to nucleotide substitutions. All these substitutions were transitions (5 C/T and 4 G/A). Between *G. g. murghi* and other *G. gallus* subspecies, the polymorphic sites were 17 and all of these were nucleotide substitutions. All the substitutions were transitions (11 T/C and 6 G/A). Between the Gallus species, high polymorphism was observed. Apart from major addition / deletions D loop region of *G. sonneratii*, *G. lafayettei* and *G. varius*, several additions and deletions were also observed. Apart from these additions/ deletions, there were 127 nucleotide substitutions between Gallus species. Out of these substitutions,83 were transitions and 44 were transversions.

4.3.3.1. Structure of mt D loop

A plot of number of variable sites in non-overlapping 50 bp segments was used to examine the distribution of variation across the control region (Fig 4.13). As described by Saccone *et al.* (1991), three distinct regions or domains were evident on the basis of nucleotide variability i.e. highly variable left domain, which (1– 450 nt), highly conserved central domain (451-900 nt) and right domain (901-1298). The domain I showed maximum size variation from 386 bp to 448 bp, while the domain II and domain III showed very narrow range of size variation i.e. 458 bp to 462 bp and from 383 bp to 384 bp, respectively (Table 4.4).

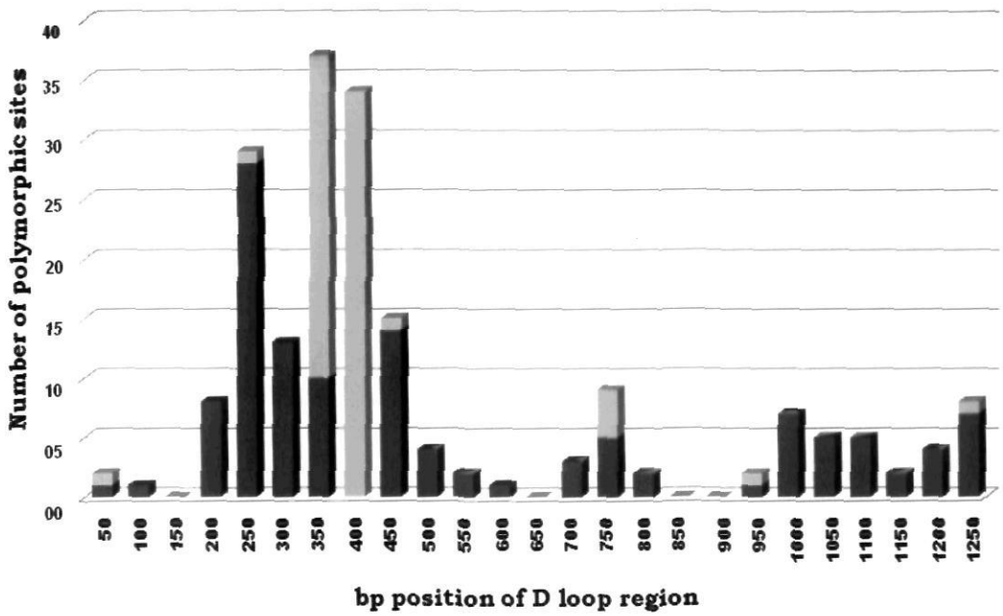

Fig 4.13. Plot of Gallus species control region variability in 50 base windows. Substitutions are indicated with black bars and indels with grey bars.

Table 4.4. The sizes of the different domain of d loop region in different gallus species.

Species	Accession number	Domain I	Domain II	Domain III	Total
G. g. murghi		386	462	384	1232
G. g. domesticus	AP003317	386	462	383	1231
G. g. domesticus	AP003318	386	458	383	1227
G. g. domesticus	AP003319	386	462	383	1231
G. g. spadiceus	AP003321	386	462	384	1232
G. g. gallus	AP003322	386	462	384	1232
G. g. bankiva	AP003323	386	462	384	1232
G. sonneratii	AP006741	448	457	383	1288
G. varius	AP003324	385	462	381	1228
G. lafayettei	AP003325	447	462	383	1292

Table 4.5. Base composition of different domains of D loop region in Indian Red Jungle Fowl

Domain	Proportion of different bases			
	A	T	G	C
D I (1-400 bases)	29.5	30.3	12.3	28.0
D II (401-871 bases)	16.7	32.4	20.8	30.0
D III (872-1232 bases)	37.1	37.7	4.7	20.5

The base composition of these domains in Indian Red Jungle Fowl is presented in Table 4.5. In domain I, the proportion of A+T was higher than G+C i.e. 59.8 % vs 40.3, while domain II was CT rich (62.4 %). In domain III, very low proportion of G i.e. 4.7 % was found, while it was highly rich in AT (78.4 %).

All the three domains possess the different motifs of functional importance. While the Left domain, possess the TAS (Terminal Associated Sequences) as well as tandem repeats, the central domain possess different conserved segments such as F box, D box and C box. The right domain possesses the CSB (Conserved Sequence Block, origin of replication and promoters for Heavy strand and Light Strand transcription (HSP and LSP). Between the Gallus species showed the conserved primary sequence motifs of the three domains in other vertebrates.

Domain I or Left Domain (ETAS domain)

There are different blocks of conserved sequences which are similar to motifs previously identified in other avian and mammalian species. The first block is perfectly conserved in Gallus species and has sequence capable to produce a hairpin structure. The secondary structure of this hairpin is determined by a stem of seven complementary Cs/Gs and a loop containing a TCCC motif. There are two copies of TCCC mapping at nt 23-26 and 244-247, which flanked the cloverleaf putative secondary structure formed by single strand domain I. The second block perfectly conserved in Gallus species has sequence similarity to mammalian TASs including the highly conserved motif GRYCAT. These GRYCAT and TAS motifs are included in the putative ETAS1 sequence block of Gallus species.

There are repeats of the sequences, which resembles with CSB1 and these repeats showed the GGACAT motif. These repeats varied between the Gallus species. All the *G. gallus* sub species as well as *G. varius* showed two repeats, while the *G. sonneratii* and *G. lafayettei* had three repeats. The single stranded domain I can form potential secondary cloverleaf structures and this structure is highly conserved between Gallus species.

Domain II or Central Domain

In the domain II, the conserved F, D and C boxes are readily recognizable. The F box showed absolute conservation among *G. gallus* subspecies, however only *G. sonneratii* and *G. lafayettei* showed only one nucleotide substitution (C to T). The D box and C box showed absolute sequence homology in all the Gallus species.

Domain III or Right Domain (CSB domain)

In domain III, the CSB-I was absolutely conserved between Gallus species,, however the CSB-II and CSB-III sequences could not be unambiguously identified. A poly (C) sequence similar to the O_H of mammals was identified just a few nucleotide downstream from the putative CSB-I. The bi-directional LSP/ HSP promoters, as identified by L'Abbe *et al*. (1991) were perfectly conserved in Gallus species. This sequence contains an almost perfect inverted repeat, which can be folded into stable hairpin. Farther upstream, is the most variable part of the CSB domain, which is rich in poly(T) and poly(A) strings.

4.3.3.2. Genetic distances between the Gallus species

The genetic distances was estimated between different chicken breeds, red jungle fowl subspecies and jungle fowls as Jukes-Cantor genetic distance, based on transitions as well as transversions in d loop region and are presented in Table 4.6. The *G. g. murghi* showed very low genetic distances with the chicken breeds (0.008) as well as with other *G. gallus* sub-species (0.003-0.012), however with other jungle fowl species, it showed higher distance (0.049-0.060). The Jukes-Cantor genetic distance were quite low between chicken breeds and ranged from 0.000 to 0.002. Similarly, the genetic distances between the chicken breeds and other *G. gallus* subspecies, were quite low (0.005 – 0.010). Other *G. gallus* subspecies also showed very low genetic distances among themselves (0.002-0.010). However, other jungle fowl species i.e. CyJF, GyJF and GJF showed comparatively much higher genetic distances not only from RJF subspecies including chicken breeds (0.049-0.060), but also from each other (0.042-0.051).

Table 4.6. Genetic distances (Jukes-Cantor) between different Gallus species based on nucleotide sequence variability in d loop region.

	Ind. RJF	AP003317	AP003318	AP003319	AP003321	AP003322	AP003323	AP006741	AP003324	AP003325
Ind. RJF	-									
AP003317	0.008	-								
AP003318	0.008	0.000	-							
AP003319	0.008	0.002	0.002	-						
AP003321	0.012	0.008	0.008	0.010	-					
AP003322	0.005	0.007	0.007	0.005	0.010	-				
AP003323	0.003	0.005	0.005	0.005	0.008	0.002	-			
AP006741	0.054	0.058	0.058	0.059	0.055	0.055	0.055	-		
AP003324	0.060	0.060	0.060	0.061	0.058	0.060	0.060	0.050	-	
AP003325	0.049	0.051	0.051	0.053	0.053	0.053	0.052	0.042	0.051	-

4.3.3.3. Phylogenetic relatedness among Gallus species

The phylogenetic tree construced using NJ method based on nucleotide sequence variability of complete D loop region has been shown in Fig 4.14. Among the different jungle fowls, three jungle fowls i.e. *G. sonneratii*, *G. lafayattei* and *G. varius* seemed to be well separated out from the *G. gallus* as all these three made a separate cluster. Among the *G. gallus* subspecies, *G. g. spadiceus* falls as separate branch, while *G. g. murghi* and *G. g. domesticus* falls in one sub-cluster, while the other two sub species i.e. *G. g. bankiva* and *G. g. gallus* made another subcluster

4.4 Polymorphism between first 400 nucleotides of D loop region in *Gallus gallus* species

A total of 13 nucleotide sequences (Table 3.4) of first 400 nucleotide of mt D loop region from different subspecies of Red Jungle Fowl, including those from Indian Red Jungle Fowl (*G. g. murghi*) were used to study the nucleotide variability in this region. The sequences were aligned using ClustalX programme and the nucleotide variability was studied by using MEGA 4 software (Fig 4.15). A total of 50 out of 400 sites were polymorphic between sub species of Red Jungle Fowl. Out of these 50 polymorphic sites, 5 were insertions and deletions while, 45 were nucleotide substitutions. Out of the 45 nucleotide substitutions, 27 were transitions (22 T/C and 5 G/A), while 18 were transversions (4 T/G, 10 Y/A, 1 C/G and 3 C/A).

4.4.1. Genetic distances between the Gallus species

The genetic distances was estimated between different red jungle fowl subspecies as Jukes-Cantor genetic distance, based on transitions as well as transversions in first 400 nucleotides of D loop region and are presented in Table 4.7. The *G. g. murghi* showed very low genetic distances with the *GG_gallus*8. *GG_gallus*10 and *GG_gallus3322* (0.008-0.013), while with other *G. g. gallus* samples i.e. *GG_gallus*11, *GG_gallus*39 and *GG-gallus*58, the genetic distances were comparatively higher (0.018-0.034). While, *G. g. murghi* showed the genetic distance ranging from 0.031-0.034 with *G. g. spadiceus*, it ranged from 0.061-0.066 with *G. g. bankiva*.

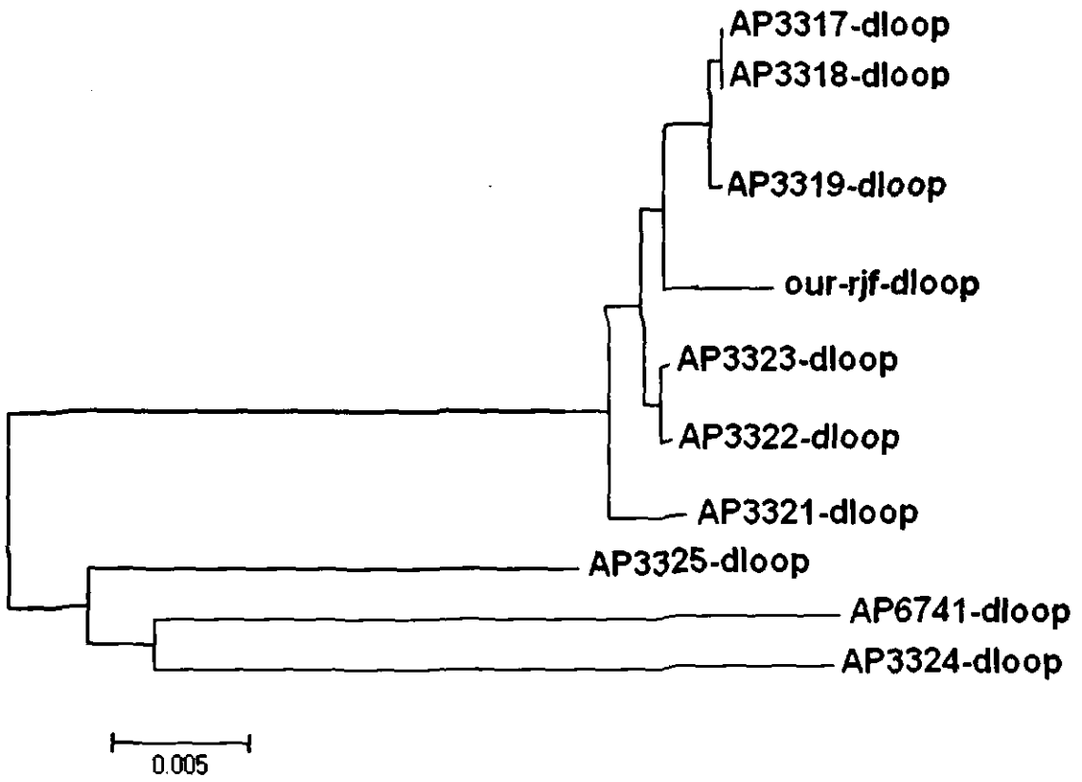

Fig 4.14. Phylogenetic tree constructed by NJ method based on the nucleotide sequences of D loop region in Indian RJF, other subspecies of RJF other jungle fowls an domesticated fowl.

```
our-rjf-400     AATTTTATTT TTTAACCTAA CTCCCCTACT AAGTGTACCC CCCCTTTCCC CCCCAGGGGG GGTATACTAT GCATAATCGT
GG_bankiva15    .......... .......... .......... .......... .......... .......... .......... ..........
GG_bankiva18    .......... .......... .......... .......... .......... ...-...... .......... ..........
GG_bankiva19    --AA...... .......... .......... .......... .......... .......... .......... ..........
GG_gallus11     .......... .......... .......... .......... .......... .......... .......... ..........
GG_gallus39     .......... .......... .......... .......... .......... .......... .......... ..........
GG_gallus58     .......... A..T...... .......... .......... .......... .......... .......... ..........
GG_gallus8      .......... .......... .......... .......... .......... .......... .......... ..........
GG_gallus10     .......... .......... .......... .......... .......... .......... .......... ..........
GG_gallus-3322  .......... .......... .......... .......... .......... .......... .......... ..........
GG_spadiceus3   .......... .......... .......... .......... .......... .......... .......... ..........
GG_spadiceus4   .......... ..-....... .......... .......... .......... .......... .......... ..........
GG_spadiceus5   .......... .......... .......... .......... .......... .......... .......... ..........
80
```

```
our-rjf-400     GCATACATTT ATATACCACA TATATTATGG TACCGGTAAT ATATACTATA TATGTACTAA ACCCATTATA TGTATACGGG
GG_bankiva15    .......... .......... .......... .......... .......... .......... .......... ..........
GG_bankiva18    .......... .......... .......... .......... .......... .......... .......... ..........
GG_bankiva19    .......... .......... .......... .......... .......... .......... .......... ..........
GG_gallus11     .......... .......... .......... .......... .......... .......... .......... ..........
GG_gallus39     .......... .......... .......... .......... .......... .......... .......... ..........
GG_gallus58     .......... .......... .......... .......... .......... .......... .......... ..........
GG_gallus8      .......... .......... .......... .......... .......... .......... .......... ..........
GG_gallus10     .......... .......... .......... .......... .......... .......... .......... ..........
GG_gallus-3322  .......... .......... .......... .......... .......... .......... .......... ..........
GG_spadiceus3   .......... .......... .......... .......... .......... .......... .......... ..........
GG_spadiceus4   .......... .......... .......... .......... .......... .......... .......... ....A.....
GG_spadiceus5   .......... .......... .......... .......... .......... .......... .......... ..........
160
```

```
our-rjf-400     CATTAATCTA TATTCCACAT TTCTCCCAAT GTCCATTCCA TGCATGATCC AGGACATACT CATTCACCCT CCCCATAGAC
GG_bankiva15    .......... C.....T... .......... .......... .......... .A.T...T.. -.G...TAT. .......AGT
GG_bankiva18    .......... C.....T... .......... .......... .......... .A.T...T.. -.G...TAT. .......AGT
GG_bankiva19    .......... C.....T... .......... .......... .......... .A.T...T.. -.G...TAT. .......AGT
GG_gallus11     .......... .......... .......... .......T. .......... .......... .......... .........T
GG_gallus39     .......... .......... .......... .......T. .......... .A........ ........T. ..........
GG_gallus58     .......... .......... .......... .......T. .......... .A........ ........T. ..........
GG_gallus8      .......... C......... .......... .......T. .......... .......... .......... ..........
GG_gallus10     .......... .......... .......... .......T. .......... .......... .......... ..........
GG_gallus-3322  .......... .......... .......... .......T. .......... .......... .......... ..........
GG_spadiceus3   .......... .......... .......... .......T. .......... .......... .......... ...T.C....
GG_spadiceus4   .......... .......... .......... .......T. .......... .......... .......... ...T.C....
GG_spadiceus5   .......... .......... .......... .......T. .......... .......... .......... ...T.C....
240

our-rjf-400     AGCTCCAAAC CACTACCAAG TCACCTAACT ATGAATGGTT GCAGGACATA AATCTCACTC TCATGCTCTT CCCCCAACAA
GG_bankiva15    .....T..-. .......... A......... .......... A......... .C...T.... ........C. ..........
GG_bankiva18    .....T..-. .......... A......... .........C A......... .C........ .....T.... ..........
GG_bankiva19    .....T..-. .......... A......... .......... A......... .C...T.... .....T.... ..........
GG_gallus11     .......... .......... .......... .......... .......... .......... .......... ..........
GG_gallus39     .......... .......... A......... .......... A......... .......... .......... ..........
GG_gallus58     ...C...... .......... A......... .......... A......... .......... .......... ..........
GG_gallus8      .......... .......... .......... .......... .T........ .......... .......... ..........
GG_gallus10     .......... .......... .......... .......... .......... .......... .......... ..........
GG_gallus-3322  .......... .......... .......... .......... .......... .....T.... .......... ..........
GG_spadiceus3   .......... ...C...... ....T..... .......... A......... .......... .....T.... ....T.C...
GG_spadiceus4   .......... ...C...... ....T..... .......... A......... .......... .....T.... ....T.C...
GG_spadiceus5   .......... ...C...... ....T..... .......... A......... .......... .....T.... ....T.C...
320
```

```
our-rjf-400     GTCACCTAAC TATGAATGGT TGCAGGACAT ACATTTAACT ACTATG-CTC TAACCCATTT GGTTATGCTC GCCGTATCAG
GG_bankiva15    .......... .......... .A........ .......... ..C...AT.. .......... .......... .TG-......
GG_bankiva18    .......... .......... .A........ .......... .TC...AT.. .......... .......... .TG.......
GG_bankiva19    .......... .......... .......... .......... .TC...AT.. .......... .......... .TG-......
GG_gallus11     ......C... .......... .A........ .......... ..C...-T.. .......... .......... .T........
GG_gallus39     .......... .......... .A........ .T.C...... ..C...-T.. .......... .......... .T........
GG_gallus58     .......... .......... .A........ .T.C...... .......-T.. .......... .......... .T........
GG_gallus8      .......... .......... .......... .......... ..C...-T.. .......... .......... ..........
GG_gallus10     .......... .......... .......... .......... .T....-T.. .......... .......... ..........
GG_gallus-3322  .......... .......... .......... .......... ..C...-T.. .......... .......... ..........
GG_spadiceus3   .......... .......... .......... .T........ ..C...-T.. .......... .......... ..........
GG_spadiceus4   .......... .......... .......... .T........ ..C...-T.. .......... .......... ..........
GG_spadiceus5   .......... .......... .......... .T........ ..C...-T.. .........G T......... ..........
400
our-rjf-400     A----
GG_bankiva15    .T---
GG_bankiva18    .TG-
GG_bankiva19    .TGGA
GG_gallus11     .----
GG_gallus39     .----
GG_gallus58     .----
GG_gallus8      .----
GG_gallus10     .----
GG_gallus-3322  .----
GG_spadiceus3   .----
GG_spadiceus4   .T---
GG_spadiceus5   .----
405
```

Fig 4.15. Sequence alignment of first 400 bp of mitochondrial D loop region in different subspecies of Red Jungle Fowl

Table 4.7. Genetic distances (Jukes-Cantor) between different Gallus gallus subspecies based on first 400 nucleotide sequence of D loop region.

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
[2]	0.061											
[3]	0.064	0.013										
[4]	0.066	0.015	0.013									
[5]	0.018	0.053	0.055	0.064								
[6]	0.029	0.047	0.050	0.058	0.021							
[7]	0.034	0.058	0.061	0.069	0.031	0.010						
[8]	0.013	0.058	0.061	0.064	0.015	0.026	0.036					
[9]	0.008	0.064	0.061	0.064	0.015	0.026	0.031	0.010				
[10]	0.010	0.055	0.064	0.061	0.013	0.023	0.034	0.008	0.008			
[11]	0.031	0.078	0.075	0.078	0.034	0.034	0.045	0.029	0.029	0.026		
[12]	0.034	0.081	0.078	0.081	0.036	0.036	0.047	0.031	0.031	0.029	0.003	
[13]	0.036	0.083	0.081	0.083	0.039	0.039	0.050	0.034	0.034	0.031	0.005	0.008

[1] our-rjf [2] *GG_bankiva*5 [3] *GG_bankiva*18 [4] *GG_bankiva*19 [5] *GG_gallus*11
[6] GG_gallus39 [7] *G_gallus*58 [8] *GG_gallus*8 [9] *GG_gallus*10 [10] *GG_gallus*3322
[11] *GG_spadiceus*3 [12] GG_spadiceus4 [13] *GG_spadiceus*

Though very low genetic distances were observed within *G. g. spadiceus* (0.003-0.008) and *G. g. bankiva* (0.013-0.015), but within *G. g. gallus*, the genetic distances were comparatively higher and showed much wider range (0.069).

4.4.2. Phylogenetic relatedness among Gallus species

The species *G. gallus* is composed of four sub-species i.e. *G. g. murghi*, *G. g. gallus*, *G. g. spadiceus* and *G. g. banikva*. The phylogenetic tree construced using NJ method based on nucleotide sequence variability in the first 400 nucleotides of D loop region has been shown in Fig 4.16. Among the different red jungle fowl subspecies, *G. g. banikva* seemed to be well separated from the other three subspecies as it made a separate cluster. Among these three subspecies, *G. g. gallus* (*GG_gallus*39 and *GG_gallus*58) made separate subcluster. In other sub-cluster, while *GG_gallus*11, made a separate branch, *G. g. spadiceus* falls in one sub-cluster, but as separate branch, while *G. g. murghi* and *G. g. gallus* (*GG_gallus8*. *GG_gallus*10 and *GG_gallus*3322) falls together.

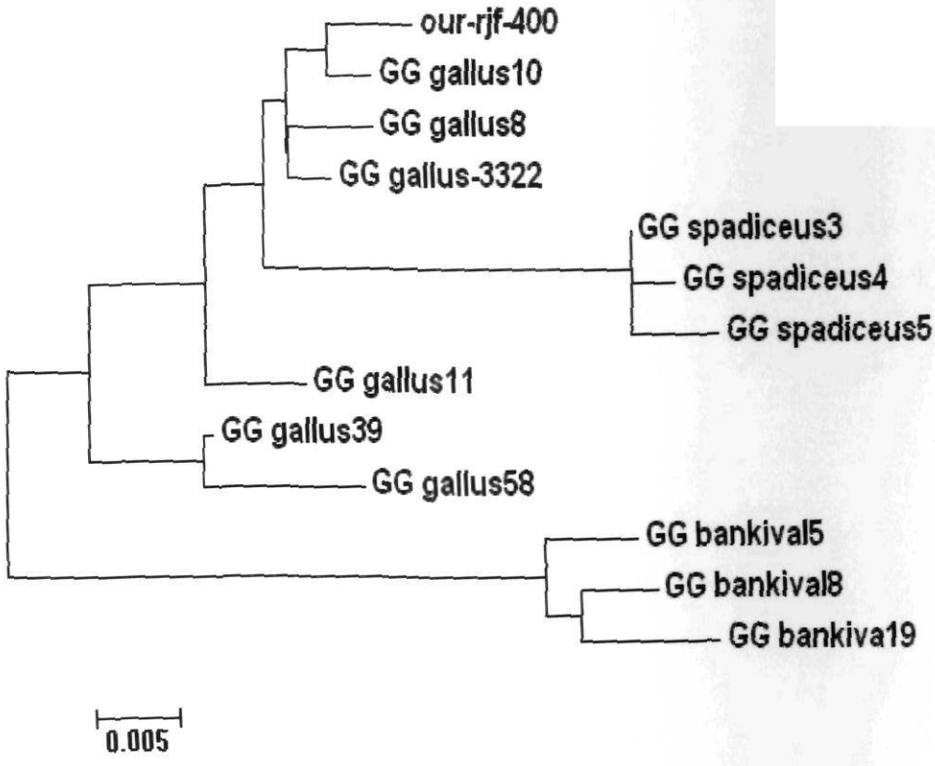

Fig 4.16. Phylogenetic tree constructed by NJ method based on the sequence of first 400 nucleotides of D loop region in Indian RJF and other subspecies of RJF.

Discussion

5.1. Identifying the probable SNPs in mitochondrial genome

The total mitochondrial genome ranged from 16784 bp to 16788 bp in different *G. gallus* subspecies i.e. *G. g. domesticus*, *G. g. gallus*, *G. g. spadiceus* and *G. g. bankiva*. On sequence comparisons of different genes identified in total 79 (~0.5 %) polymorphic sites and of these 12 were due to insertion and deletions, while 67 were due to nucleotide substitutions. It suggests very low nucleotide sequence variability for mitochondrial genome with *G. gallus* subspecies. Gupta *et al.* (2005) reported very high genetic similarity (> 99 %) between *G. g. murghi* and White Leghorn based on nucleotide sequence variability for 12S rRNA gene. There are several reports on genetic similarity between Indian red jungle fowl and domestic chicken using different type of DNA markers. Very close genetic relatedness between Indian red jungle fowl (*G. g. murghi*) and domestic fowl was also revealed by other DNA markers such as randomly amplified polymorphic DNA, micro/mini satellite associated sequence amplification (MASA) and microsatellite markers (Tomar *et al.*, 2007).

In present study all the 67 nucleotide substitution sites were SNPs and of these 51 were in coding region and 16 were in non coding region. Guan *et al.* (2007) identified 91 SNPs, with 70 in the coding region and 21 in the non-coding region of mt genome by in silico analysis. Out of the 67 SNPs, only 19 were identified as population specific SNPs. Restriction enzyme mapping for these SNPs could not detect any polymorphic restriction enzyme, hence PCR-RFLP could not be done. Alternatively, Allele Specific PCR was used for genotyping the RJF and DF populations for the probale SNPs.

5.2. SNP genotyping

Effective primers with good score could be designed for six SNPs only i.e. COI-114-C/T, COI-255-G/A, COII-279-C/T, ATP6-294-G/A, ND4-969-C/T and CYTB-543-T/C, hence the genotyping of the RJF an population was done for these six SNPs. While the general primers, expected to amplify the common band in both the populations were designed from the total conserved regions encompassing the SNP in both the populations, the allele specific primers were designed with a specific base at last three terminal nucleotides at 3' end corresponding to one SNP allele. Newton *et al.* (1989) showed that a specific base at 3' end corresponding to one SNP allele was able to amplify the specific allele (match base allele) more effectively than the non-specific one (mismatch base allele). Furthermore an additional mismatch within three or four bases from the 3' end enhances primer specificity (Kwok *et al.*, 1990 ; Drenkand *et al.*, 2000). In present study, while the general primers amplified the common band in both the population, the allele specific primers also amplified the specific and in respective populations.

At COI-114-C/T SNP, while the COI-F and COI-R primers amplified a 603 bp fragment in all the individuals from both the populations, COI-F and COI-114-R-DF primers amplified the 334 bp fragment in 17 out of 18 individuals from DF population and COI-F and COI-114-R-RJF primers amplified the 334 bp band in 16 out of 18 individuals from RJF population. Similarly at COI-255-G/A SNP site, using COI-F and COI-R primers, a 603 bp fragment was amplified in all the individuals from both the populations, while COI-F and COI-255-R-DF primers amplified a 475 bp band in 15 out of 18 individuals in DF population and COI-F and COI-255-R-RJF primers amplified the 475 bp fragment in 16 out of 18 individuals in RJF population.

Using COII-F and COII-R primers, a 523 bp fragment was amplified in all the individuals from both the populations, while COII-279-F-DF and COII-R primers amplified the 268 bp fragment in 14 out of 18 individuals in DF population and COII-279-F-RJF and COII-R primers amplified 268 bp band in 17 out of 18 individuals from RJF population. At ATP6-294-G/A SNP, ATP6-F and ATP6-R primers amplified a 496 bp fragment in all the individuals from

both the populations, however ATP6-294-F-DF and ATP6-R primers amplified the 270 bp in 16 out of 18 individuals from Df population. Similarly, a 270 bp band was amplified in 15 out of 18 individuals from RJF population with ATP6-294-F-RJF and ATP6-R primers. Using ND4-F and ND4-R primers, a 521 bp fragment was amplified in all the individuals from both the populations, while ND4-F and ND4-969-R-DF primers could amplify a 335 bp fragment in 15 out of 18 individuals from DF population and ND4-F and ND4-969-R-RJF primers amplified the 335 bp fragment in 17 out of 18 individuals from RJF population. At CYTB-543-T/C SNP, while the CYTB-F and CYTB-R primers amplified the 514 bp fragment in all the individuals from both the populations, CYTB-F and CYTB-543-R-DF primers amplified 363 bp band in 17 out of 18 individuals and CYTB-F and CYTB-543-R-RJF primers amplified 363 bp band in 14 out of 18 individuals from RJF population.

At each SNP site, the specific allele showed very high frequency in the respective populations, but these alleles were not fixed. At COI-114-C/T snp, allelic frequency of C allele in DF population and T allele in RJF population was found to be 0.94 and 0.89, respectively. At COI-255-G/A snp, allelic frequency of G allele in DF population and A allele in RJF population was found to be 0.83 and 0.89, respectively. Similarly the allelic frequency of C allele in DF population and T allele in RJF population was found to be 0.78 and 0.94, respectively at COII-279-C/T snp. At ATP6-294-G/A SNP, allelic frequency of G allele in DF population and A allele in RJF population was found to be 0.94 and 0.83, respectively and allelic frequency of C and T alleles were 0.83 and 0.94 in DF and RJF populations respectively. At CYTB-543-T/C SNP, allelic frequency of T allele in DF population and C allele in RJF population was found to be 0.94 and 0.78, respectively. In spite of these high frequency of specific alleles in respective populations, these alleles were not fixed in the populations hence these SNPs could not be used as population specific markers.

Soleimani *et al.* (2003) reported the Allele Specific PCR with 3 primers as an efficient, cost effective, and reliable way for SNP validation in barley cultivars. Subsequently, Bundock *et al.* (2005) reported an alternative cheap, simple, transferable and robust method of SNP genotyping, based on a three primer nested allele specific PCR. They developed markers for 40 SNP sites in 28 barley

genes based on the nested allele specific PCR method. We obtained a 95% success rate for the development of allele specific markers (for one or both alleles) with this approach. Markers for both alleles at a SNP site were developed in 84% of cases.

Harumi *et al.* (2004) identified 6 SNPs in mt D loop regions from sequence alignment of D loop region sequences from six chickens. For six of these, they designed the PCR primers that had each base (A,C G and T) as the penultimate base at the 3' end (N2 base) and bases specific to both alleles at the 3' end. These primers were effective in typing the SNP. Twenty one out of 96 primers succeeded in distinguishing the SNP by the presence or absence f the PCR product. Sano *et al.* (2005) used polymerase chain reaction (PCR) based screening using single nucleotide polymorphism (SNP) at 686 position in mt D loop (#AB091008) for testing male putative germline chimeric chickens. In this study, the type of SNP in all BPR at position 020 was fixed as base G, and as base A in all WL. When the PCR product was typed as both donor (base G)- and recipient (base A)-derived, the male was determined as the germline chimera. The results of the male screening test showed good agreement with the progeny test for detecting germline chimeric chickens.

Bharani Kumar *et al.* (2007) identified 10 SNPs identified in Myostatin gene in the nine breeds/populations of indigenous poultry and Red Jungle Fowl. The five SNPs were in promoter region, while one each in intron 1 and 2 and three in exon 1. PCR-RFLP was carried out to genotype the populations at seven SNPs while three SNPs were genotyped by SNaPshot method using automated DNA sequencer as no restriction sites were found at these SNP sites. The correspondence analysis separated out Punjab Brown, Red jungle fowl and Nicobari populations from the rest of the indigenous chicken populations. It is postulated that the populations of Punjab Brown and Nicobari showed the recent common ancestry with Red jungle fowl.

Gaudet *et al.* (2007) reported a new method for polymerase chain reaction (PCR) amplification of multiple specific alleles (PAMSA), which allows efficient discrimination of SNP polymorphisms in one reaction tube with standard PCR conditions. To validate the accuracy of this method, 94 full-sib individuals

were genotyped with three SNPs and compared to the genotypes obtained by cleaved amplified polymorphic sequence (CAPS) or derived CAPS. This method is flexible, inexpensive, and well suited for high throughput and automated genotyping.

5.3. Polymorphism in D loop region

The 1325 bp fragment was sequenced on ABI DNA sequencer. The nucleotide sequence revealed that the 1325 bp fragment is comprised of 45 nucleotide from 3' end of tRNA-Glu, complete nucleotide of D loop region (1232) and 48 nucleotides from 5' end of tRNA-Phe. The mitochondrial D loop size seemed to be highly conserved within red jungle fowl subspecies, however in *G. g. domesticus*, the size ranged from 1227 in White Plymouth Rock to 1232 in New Hampshire. However, the size variability was very high when co mpared among the Gallus species as it was 1288 nt in *G. sonnerattii*, 1228 nt in *G. varius* and 1292 nt in *G. lafeyettii*. Sequence homology comparisons revealed that the much higher size in *G. sonnerattii* (1288 bp) and *G. lafeyettii* (1292 bp) is due to an additional tandem repeat.

For the D loop region, very low variability was observed within *G. g. domesticus* as only 6 polymorphic sites were observed. Among these 2 were nucleotide substitutions and all were transitions. Similarly, very low nucleotide variability was observed within *G. gallus* subspecies including *G. g. murghi* as there were 17 polymorphic sites and all of these were transitions. Between *G. g. murghi* and different chicken breeds, a total of 14 polymorphic sites were observed and 9 of these were due to nucleotide substitutions, which were transitions only. Between the Gallus species, high polymorphism was observed. Apart from major addition / deletions D loop region of *G. sonnerattii*, *G. lafeyettii* and *G. varius*, several additions and deletions were also observed. Apart from these additions / deletions, there were 127 nucleotide substitutions between Gallus species. Out of these substitutions, 83 were transitions and 44 were transversions. These results showed a transition bias in the nucleotide substitutions among Gallus genus, especially in *G. gallus* species.

Shen *et al.* (2002) used complete cytochrome b (cyt b) gene (1140 bp) nucleotide sequences and reported high homologies observed among the chicken

breeds (egg-purpose). Similarly, Wada *et al.* (2004) reported 99.77 % similarity between silkie and white leghorn as they reported 39 nucleotide differences in 16, 784 bp of complete mitochondrial DNA. Some other studies, based on nucleotide variation in hypervariable D loop region of mtDNA reported comparatively higher sequence divergence between different chicken breeds. Fu *et al.* (2001b) reported 4.45 % sequence divergence (24 polymorphic sites out of 539 bp of D –loop) among the 30 DNA sequences from native chicken breeds of Zhejiang province and the white leghorn chicken. Further, Liu *et al.* (2004) investigated the genetic variability of 544 bp mt DNA hyper variable HVSI) in a total of 48 birds belonging to 12 Chinese native chicken breeds and identified 16 from 35 polymorphic nucleotide sites, which accounted for 6.4 %.

Earlier studies based nucleotide diversity in 539 bp of D loop region also revealed very low sequence divergence between *G. g. gallus* and *G. g. spadiceus* (Fu *et al.,* 2001 ; Niu *et al.,* 2002), however Fumihito *et al.* (1996) observed comparatively more nucleotides diversity between *G. g. bankiva* and other two *G. gallus* subspecies i.e. *G. g. gallus* and *G. g. spadiceus*. A high nucleotide divergence was observed by earlier workers also (Munechika *et al.,* 1997; Fu *et al.,* 2001 and Niu *et al.,* 2002).

Structure of mt D loop

Sequencing of the 1325 bp fragment and its alignment with mitochondrial genome of other galliformes as well other Aves, showed that the D loop has tRNA Glu at its 5' end and tRNA PHe at 3' end. This gene order in the flanking region of D loop is very well conserved in all the Galliformes as well as other Avian order (Desjardins and Moris, 1990; Desjardins *et al.,* 1990 and Quinn and Wilson, 1993) and is different from that observed in mammals, where tRNA Pro is present at 5" end of D loop (Saccone *et al.*, 1991; Douzery and Randi, 1997).

Distribution of variable sites in non-overlapping 50 bp segments of the mt D loop sequences from Genus Gallus revealed that the whole D loop region could be divided into three domains as suggested by Saccone *et al.* (1991) i.e. highly variable left domain, which (1– 450 nt), highly conserved central domain (451-900 nt) and right domain (901-1298). This three domain structure of D

loop was very much evident in different avian species. Marshal and Baker (1997) delimited domains in D loop region to encompass regions of differing variability as follows: 5' domain I-bases l-400; central domain II-bases 401-873; and 3' domain III-bases 874-1252 in Fringilline Finches (Fringilla spp.) and the Greenfinch (*Carduelis chloris)* under the order Passeriformes. Randi and Lucchini (1998) showed that D-loops of *Alectoris* can be descriptively divided into three domains: a variable domain I (nt 1–316), with 22% variable sites (including indels) among *Alectoris;* a central conserved domain II (nt 317–784), with only 5% variable sites among *Alectoris;* and a variable domain III, spanning the origin of replication OH to the end of the D-loop (nt 785–1160), with 14% variable sites. Roques *et al.* (2004) reported three distinct domains in the control region i.e. left domain I (DI), central domain (DII) and the right domain III (DIII) in two vulture species i.e. *Gypaetus barbatus* and *Neophron percnopterus*

The base composition of these domains in Indian Red Jungle Fowl (Table 4.5) showed that in domain I, the proportion of A+T was higher than G+C i.e. 59.8 % vs 40.3, while domain II was CT rich (62.4 %). In domain III, very low proportion of G i.e. 4.7 % was found, while it was highly rich in AT (78.4 %). Roques *et al.* (2004) reported that though the control region DI of *N. percnopterus* is not AC rich, as in most bird species, but was exceedingly rich in thymidine (33.7%) due to the presence of a tandemly repeated region. However both the species had A + T > G + C, a CT-rich D-II, and a DIII very poor in G and with high proportions of AT nucleotides.

Domain I or Left Domain (ETAS domain)

There are different blocks of conserved sequences which are similar to motifs previously identified in other avian and mammalian species. The first block is perfectly conserved in Gallus species and has sequence similarity to the goose hairpin as described by Quinn and Wilson (1993) in goose. The secondary structure of this hairpin is determined by a stem of seven complementary Cs/Gs and a loop containing a TCCC motif, also present in mammalian D loop (Douzery and Randi, 1997), which was associated expermimentally with termination of H strands (Dufresne *et al.*, 1996). There are two copies of TCCC mapping at nt 23-26 and 244-247, which flanked the

cloverleaf putative secondary structure formed by single strand domain I. The presence of this 'Goose hairpin' was also reported in the D-loops of 16 species of phasianids (Fumihito *et al.*, 1994; 1995), and in *Anas platyrhyncos* (Ramirez *et al.*, 1993) and in *Alectoris* (Randi and Lucchini, 1998). At the 5′ end of the DI, an interrupted poly-C sequence was observed in the Gallus sequences including *G. g. murghi*. This structure seems to be a conserved feature across many species and was described in birds including Struthioniformes, Galliformes, Falconiformes, and Sphenisciformes (Haring *et al.*, 2001; Ritchie and Lambert,2000 & Roques *et al.*, 2004).

The second block perfectly conserved in Gallus species has sequence similarity to mammalian TASs including. Sbisa *et al.* (1997) identified two long conserved blocks in DI i.e. ETAS1 and ETAS2, potentially involved in the control of H strand synthesis. Interestingly, in second block of domain I in Gallus genus, a sequence similar to ETAS1, however, the ETAS2 sequence could not be found without ambiguity. This sequence has the highly conserved motif GRYCAT as well as TAS motifs and are well conserved in all the species of Gallus genus.

A putative TAS sequence has been identified in domain I of chaffinches and greenfinch (Marshall and Baker, 1997). The motif GYRCAT (Y 4 C/T; R 4 A/G) is widespread in domain I of a range of mammalian D-loops (Douzery and Randi, 1997). Its functional importance is suggested by both comparative (Douzery and Randi, 1997) and experimental (Dufresne *et al.*, 1996) findings. These GYRCAT and TAS motifs are included in the putative ETAS1 sequence block of *Alectoris* which has 67% similarity to the mammalian consensus ETAS1 sequence (Sbisa *et al.*,1997).

Roques *et al.* (2004) also reported that in both the DI of *G. barbatus* and *N. percnopterus*, a 31 bp long sequence of high similarity (71%) with the ETAS1 sequence could be identified, however, the ETAS2 sequence could not be found without ambiguity.

There are repeats of the sequences, which resembles with CSB1 and these repeats showed the GGACAT motif. These repeats varied between the gallus species. All the *G. gallus* sub species as well as G. varius showed two

repeats, while the *G. sonneratti* and *G. lefeyettei* had three repeats. The single stranded domain I can form potential secondary cloverleaf structures and this structure is highly conserved between Gallus species.

The CSB1 like sequences were also reported in other avian species like *Alectoris* and the motif GGACAT is very conserved among galliforms, *Anas* and *Anser (* Desjardins and Morais, 1991; Quinn and Wilson, 1993; Ramirez *et al.*, 1993; Fumihito *et al.*,1994 &1995 ; Randi and Lucchini, 1998). The potential secondary cloverleaf structures formed by single-stranded domain I is also well conserved.

Domain II or Central Domain

The domain II showed very high degree of conserved sequences in the genus *Gallus* and showed the presence of the F, E, D and C boxes. The F box showed absolute conservation among *G. gallus* subspecies, however only *G. sonneratii* and *G. lafayettei* showed only one nucleotide substitution (C to T). The D box and C box showed absolute sequence homology in all the Gallus species. These conserved blocks were also reported in other avian species like goose (Quinn and Wilson, 1993), finches (Marshall and Baker, 1997), Alectoris (Randi and Lucchini, 1998) and vultures (Roques *et al.*, 2004) as well as mammals (Sbisa *et al.*, 1997).

Roques *et al.* (2004) also reported the CSB1 region in domain III in both species, however, the conserved blocks CSB2 and CSB3, previously reported in other vertebrates and believed to be involved in the initiation of DNA synthesis, could not be mapped unambiguously. Similarly these boxes were absent in other Falconiformes and Ciconiiformes species such as *Falco peregrinus* (Mindell *et al.*, 1998), *Buteo buteo* (Haring *et al.*, 2001) and *Ciconia ciconia* (Yamamoto *et al.*, 2000).

Domain III or Right Domain (CSB domain)

The domain III also showed several conserved functional sequences in the genus *Gallus*. A poly (C) sequence similar to the O_H of mammals was identified just a few nucleotide downstream from the poutative CSB-I. The CSB-I was absolutely conserved between Gallus species,, however the CSB-II and CSB-III

sequences could not be unambiguously identified. Randi and Lucchini (1998) also reported this poly C stretch sequence similar to the O_H of mammals in *Alectoris*. The conserved CSB1 block has been identified in the domain III, but CSB2 and CSB 3 could not be identified in this region of other avian species like Alectoris (Randi and Lucchini, 1998), chaffinches and greenfinch (Marshall and Baker, 1997).

The bi-directional LSP/HSP promters, as identified by L'Abbe *et al.* (1991) were perfectly conserved in the genus *Gallus*. This sequence contain an almost perfect inverted repeat, which can be folded into stable hairpin. Farther upstream, is the most variable part of the CSB domain, which is rich in poly(T) and poly(A) strings. The LSP/HSP sequences were also identified in Alectoris and were well conserved. A stable hairpin is possible immediately upstream from the promoters. (Randi and Lucchini, 1998).

Genetic distances and phylogenetic relationship between Gallus species

The domestication of chicken seems to have a long history. Earliest evidences of domestication of the chicken date back to nearly 8000 years. From 16 neolithic sites along the yellow river in Northeast China, the remains of domesticated chickens were evident and some of them dated back to 7500 years (West and Zhou, 1989). However, very strong indication of domestication of chicken was found in the Mohenjo-Doro in the Indus valley (Zeuner, 1963). The signs of domestication were also found in unlikely places like Ukraine and Spain (West and Zhou, 1989). Thus the question of single domestication site or multiple and independent domestication sites is always debatable. Earlier population studies based on isozyme polymorphism suggested the multiple and independent site of domestication (Hashiguchi *et al.,* 1983). Among the different jungle fowls, red jungle fowl (*G. gallus*) is believed to be the sole progenitor of domestic fowl. There are five subspecies of red jungle fowl inhabiting the Indian sub-continent eastwards across Myanmar, South China, Indonesia to Java. Among these five, in India, two subspecies namely *G. g. murghi* (Indian red Jungle Fowl) and *G. g. spadiceus* (Bermese red jungle fowl) are found. While the former is distributed in the north and central part of India, extending eastwards to Orissa and West Bengal, the lateris confined to

the Northeastern parts of India. Apart from red jungle fowl, there are three other jungle fowl species i.e. Green jungle fowl (*G. varius*), Ceylon jungle fowl (*G. lafayetei*) and Grey Jungle fowl (*G. sonerratii*). Therefore, estimation of genetic diversity between domestic fowl, red jungle fowl sub-species and other jungle fowl species and establishing phylogenetic relationship is important to have understanding of monophyletic or polyphyletic origin of domestic chicken origin.

Overall low genetic distances were observed within as well between the species under genus Gallus. The *G. g. murghi* showed very low genetic distances with the chicken breeds (0.008) as well as with other *G. g.* sub-species (0.003-0.012), however with other jungle fowl species, it showed higher distance (0.049-0.060). In general all the other three jungle fowl species i.e. CyJF, GyJF and GJF showed comparatively much higher genetic distances not only from RJF subspecies including chicken breeds (0.049-0.060), but also from each other (0.042-0.051). The phylogenetic tree based on nucleotide sequence variability of complete D loop region showed the distantness of the *G. gallus* with other three jungle fowls i.e. *G. sonneratii, G. lafayattei* and *G. varius* as they make a separate cluster. Between *G. gallus* subspecies, *G. g. bankiva* and *G. g. gallus* seemed to be distant from *G. g. spadiceus*, *G. g. murghi* and *G. g. domesticus*. *G. g. murghi* and *G. g. domesticus* showed maximum closeness. These phylogenetic analysis also clearly support very high genetic relatedness of *G. g. domesticus* with *G. gallus* however, the relative closeness of *G. g. domesticus* with other *G. gallus* subspecies is not conclusive due to almost equal distance of all these red jungle fowl subspecies from most of the domestic chicken breeds.

Earlier reports also showed the much closeness of domestic chicken (*G. g. domesticus*) with red jungle fowl in comparison to other jungle fowl species, which were clearly well distant from domestic chicken, hence exclude the possibilities of their contribution towards evolution of domestic chicken. Fu *et al.* (2001) and Niu *et al.* (2002), based on first 539 bases of mitochondrial DNA D-loop region, reported that among the four species of the genus *Gallus*, *G. g. domesticus* was closest to the red jungle fowl. Similarly, Dong *et al.*(2002) reported that *G. g. domesticus* was closest the red jungle fowl (*G. gallus*) based on mt DNA D-loop polymorphism. Nishibori *et al.* (2005), based on mitochondrial

D-loop analysis further supports the hypothesis developed from morphology and progeny production that red jungle fowl (RJF) is the direct ancestor of the chicken. Based on RFLP and sequence of first 400 bp of control region of mitochondrial DNA, Fumihito *et al*. (1994) reported very low sequence divergence (0.5-3.0 %) between various domestic breeds of chicken and *G. g. gallus* (from Thailand), indicating a single domestication event in the area inhibited by this sub-species of the red jungle fowl. Later Fumihito *et al*. (1996) made phylogenic analysis, on the basis of nucleotide divergence 480 bp of D loop region and observed that all domestic fowls including Indonesian races belong to same clusters with *G. g. gallus* and *G. g. spadiceus* from Thailand and its adjacent areas. In contrast to these reports, Liu *et al.* (2004) based on the genetic variability of 544 bp mt DNA hyper variable HVSI), reported close genetic relationship between the Chinese native chicken breeds with *G. g. gallus* as well as G. *g. spadiceus* from different areas. Similarly, phylogenetic analyses based on mtDNA hyper variable segment I (HVS-I) revealed nine highly divergent mtDNA clades (A–I) in which seven clades contained both the red jungle fowls and domestic chickens and there was no breed-specific clade in the chickens (Liu *et al.*, 2006). Oka *et al.* (2007) analyzed the mitochondrial DNA D-loop region of Japanese native chickens. These results indicate that Chinese and Korean chickens were derived from Southeast Asia. Following the domestication of red jungle fowl, a non-game type chicken was developed, and it spread to China. A game type chicken was developed in each area. Both non-game and game chickens formed the foundation of Japanese native chickens.

5.4 Polymorphism between first 400 nucleotides of D loop region in *Gallus gallus* species

The Domain I of the D loop region was found to be most polymorphic. Hence the nucleotide sequence variability in this region was used to establish the phylogenic relationship of Indian red jungle fowl (*G. gallus murghi*) with the other three subspecies of Red Jungle Fowl. While very low genetic distances were observed within *G. g. spadicieus* (0.003-0.008) and *G. g. bankiva* (0.013 - 0.015), *G. gallus* showed higher genetic distances range (0.008-0.069) within themselves. The *G g. murghi* showed very low genetic distances (0.008-0.013) with the one group of *G. g. gallus* (*GG_gallus*8.

*GG_gallus*10 and *GG_gallus*3322) in comparison to other *G. g. gallus* group i.e. *GG_gallus*11, *GG_gallus*39 and *GG-gallus*58 (0.018-0.034). The genetic distances between *G. g. murghi* and *G. g. bankiva* were higher (0.061-0.066) than those observed between *G. g. murghi* and *G. g. spadiceus* (0.031-0.034).

The phylogenetic analysis revealed that among the different red jungle fowl subspecies, *G. g. banikva* seemed to be well separated from the other three subspecies as it made a separate cluster. Among these three subspecies, *G. g. gallus* makes two separate clusters i.e. one those from Thailand (*GG_gallus*8. *GG_gallus*10 and *GG_gallus*3322) and other those from Japan and Indonesia (*GG_gallus*11 and *GG_gallus*39 and *GG_gallus*58). All the *G. g . spadiceus* falls in one sub-cluster. The *G. g . murghi* falls in one sub-cluster along with the *G. g. gallus* from Thailand.

Fumihito *et al.* (1996) made phylogenic analysis, on the basis of nucleotide divergence 480 bp of D loop region and showed that *G. g. gallus* and *G. g. spadiceus* do not form two separate clusters in the phylogenetic tree. It seems that *G. g. gallus* is divided into two separate populations: (i) an island population from Sumatra and a (ii) continental population. *G. g. spadiceus* is more closely allied with a continental population of *G. g. gallus* than an island population of *G. g. gallus* is to its continental relation. Thus, the subspecies status given to *G. g. spadiceus* might be questioned. On the other hand, *G. g. bankiva*, which inhabits the Indonesian islands of Sumatra, Java, and Bali, is a very distinct entity that is clearly separate from *G. g. gallus* as well as *G. g. spadiceus*; thus, its subspecies status appears to be well-deserved. The reports on the phylogenetic relationship of *G. g. murghi* are very few. Gupta *et al.* (2005), who reported very high genetic similarity (> 99 %) between *G. g. murghi* and White Leghorn. However there are several reports on genetic similarity between Indian red jungle fowl and domestic chicken using different type of DNA markers. Very close genetic relatedness between Indian red jungle fowl (*G. g. murghi*) and domestic fowl was also revealed by other DNA markers such as randomly amplified polymorphic DNA, micro/mini satellite associated sequence amplification (MASA) and microsatellite markers (Tomar *et al.* 2007). However very recently, Kanginakudru *et al.* (2008) analyzed 76 Indian birds that included 56 *G. g. murghi* (RJF), 16 *G. g. domesticus* (domestic chicken) and 4 *G. sonneratii*

(Grey JF) using both microsatellite markers and mitochondrial D-loop sequences. The microsatellite-based phylogenetic trees showed a clear separation of *G. g. domesticus* from *G. g.murghi*, and *G. sonneratii*. Mitochondrial DNA based mismatch distribution analyses showed a lower Harpending's raggedness index in both *G. g. murghi* (0.001515) and in Indian *G. g. domesticus* (0.0149) birds indicating population expansion. When meta analysis of global populations of 855 birds was carried out using median joining haplotype network, 43 Indian birds of *G. g. domesticus* (19 haplotypes) were distributed throughout the network sharing haplotypes with the RJFs of different origins. These results suggested the evidence for domestication of Indian birds from *G. g. spadiceus* and *G. g. gallus* as well as from *G. g. murghi*, corroborating multiple domestication of Indian and other domestic chicken.

Hence the following conclusions may be drawn from these results

- Sequence homology comparisons of whole mitochondrial genome between DF and RJF revealed 19 probable SNPs.
- None of the SNPs could generate PCR-RFLP.
- Effective primers could be designed for 6 of the 19 SNPs.
- Allele specific PCR was effective in genotyping the specific alleles.
- The specific alleles at all the 6 SNP sites, though were present in very high frequency in respective populations, but were not fixed. Hence could not be used for population identification.
- Complete mitochondrial D loop was amplified, cloned and sequenced in Indian RJF. The size was 1232 bp.
- The complete D loop could be divided into three domains i.e. Domain I, II and III.
- The D loop showed most of the conserved primary and secondary structures in the genus *Gallus* found in mammalian D loop region.
- Within Gallus genus, the D loop was highly conserved.
- Variable number of the tandem repeats in the D loop of the different Gallus species is the reason for size difference in D loop.

- Based on nucleotide sequence variability of complete D loop region, Indian RJF showed very high similarity with domestic chicken breeds as well as other RJF subspecies.
- Other JF (CyJF, GyJF, GJF) showed higher genetic distances with Indian RJF, other RJF subspecies as well as domestic chicken breeds.
- Based on nucleotide sequence variability of first 400 nt of mt D loop, Indian RJF showed very high similarity with the *G. g. gallus* from Thailand.
- The *G. g. gallus* showed two clusters well separated from each other suggesting independent origin and multiplication of this subspecies at these places.

Summary

Present study has been mainly aimed for genotyping the potential SNPs in mitochondrial genome by allele specific PCR and/or PCR RFLP in red jungle fowl and domestic fowl population and estimating genetic diversity between red jungle fowl and domestic fowl using the polymorphism in mitochondrial D loop.

A total of eight complete mitochondrial genome sequences (four from *G. g. domesticus* and four from other subspecies of red jungle fowl i.e. *G. g. gallus*, *G. g. spadiceus* and *G. g. bankiva*) were identified. Sequence comparisons of all the 37 genes and the control region revealed that size variation was observed only in control region, 12S rRNA gene, 16S rRNA gene, COIII gene and ND4 gene. The nucleotide sequence variability was observed in the control region (D loop), both the rRNA genes, all the mRNA genes except ATPase8 and only two tRNA genes i.e. tRNA-Arg and tRNA-Pro. The number of total polymorphic sites were 17 in D loop, 5 in 12S rRNA, 8 in 16S rRNA, 3 in NDI, 2 in ND2, 5 in COI, 3 in COII, 2 in ATPase 6, 6 in COIII, 2 in ND3, 1 in tRNA-Arg, 8 in ND4, 4 in ND4L, 4 in ND5, 6 in Cyt b, 1 in tRNA-Pro and 2 in ND6.

Out of these, total of 19 population specific nucleotide substitutions were identified. The number of SNPs specific to *G. g. domesticus* were 3 in D loop, 1 in 16S rRNA, 1 in NDI, 4 in COI, 2 in COII, 1 in ATPase 6, 1 in COIII, 1 in ND3, 2 in ND4, 1 in ND4L, 1 in Cyt b and 1 in tRNA-Pro. None of these substitutions created or abolished the restriction site of the commonly available restriction enzymes. However, for 6 SNPs i.e. COI-114-C/T SNP, COI-255-G/A

SNP, COII-279-C/T SNP, ATP6-294-G/A SNP, ND4-969-C/T SNP and CYTB-543-T/C SNP, primers with good score could be designed. For each SNP, two sets of primers were designed. The first set i.e. general set, was designed for amplification of the common fragment containing the SNP. Other set of primers was allele specific, having one common primer, which used to be one of the primer of general set of primers, while the second primer of this set used to be allele specific, having the SNP as either the most terminal nucleotide at 3' end or one of the 3 terminal nucleotides at 3' end.

A resource population of 18 RJF individuals and 18 DF individuals were genotyped for these 6 SNPs. The allele specific primers were effective in amplifying the specific alleles. At each SNP site, the specific allele showed very high frequency in the respective populations, but these alleles were not fixed. At COI-114-C/T snp, allelic frequency of C allele in DF population and T allele in RJF population was found to be 0.94 and 0.89, respectively. At COI-255-G/A snp, allelic frequency of G allele in DF population and A allele in RJF population was found to be 0.83 and 0.89, respectively. Similarly the allelic frequency of C allele in DF population and T allele in RJF population was found to be 0.78 and 0.94, respectively at COII-279-C/T snp. At ATP6-294-G/A SNP, allelic frequency of G allele in DF population and A allele in RJF population was found to be 0.94 and 0.83, respectively and allelic frequency of C and T alleles were 0.83 and 0.94 in DF and RJF populations respectively. At CYTB-543-T/C SNP, allelic frequency of T allele in DF population and C allele in RJF population was found to be 0.94 and 0.78, respectively. In spite of these high frequency of specific alleles in respective populations, these alleles were not fixed in the populations hence these SNPs could not be used as population specific markers.

A 1325 bp fragment having the complete D loop region was amplified using the forward primer from tRNA-Glu and the reverse primer from tRNA-Phe. This fragment was cloned in pTZ57R/T vector and sequenced. The nucleotide sequencing revealed that this 1325 bp fragment was comprising of 45 nucleotide from 3' end of tRNA-Glu, complete D loop region (1232) and 48 nucleotides from 5' end of tRNA-Phe was amplified using the primers were designed to amplify the complete D loop region. The 1232 bp nucleotide sequence was aligned with the related D loop sequences from from other Gallus species,

G. gallus subspecies including *G. g. domesticus*. The length of D loop region showed ranged from 1227 in WPR to 1292 bp in CyJF.

For the D loop region, very low variability was observed within *G. g. domesticus* as only 6 polymorphic sites were observed. Among these 2 were nucleotide substitutions and all were transitions. Similarly, very low nucleotide variability was observed within *G. gallus* subspecies including *G. g. murghi* as there were 17 polymorphic sites and all of these were transitions. Between *G. g. murghi* and different chicken breeds, a total of 14 polymorphic sites were observed and 9 of these were due to nucleotide substitutions, which were transitions only. Between the Gallus species, high polymorphism was observed. Apart from major addition / deletions in D loop region of *G. sonnerattii*, *G. lafeyettii* and *G. varius*, several additions and deletions were also observed. Apart from these additions/ deletions, there were 127 nucleotide substitutions between Gallus species. Out of these substitutions, 83 were transitions and 44 were transversions. These results showed a transition bias in the nucleotide substitutions among Gallus genus, especially in *G. gallus* species.

Based on the distribution of variation across the control region, three distinct regions or domains i.e. highly variable left domain, which (1– 450 nt), highly conserved central domain (451-900 nt) and right domain (901-1298) were observed. In domain I, first block is perfectly conserved in Gallus species and has sequence capable to produce a hairpin structure. The secondary structure of this hairpin is determined by a stem of seven complementary Cs/ Gs and a loop containing a TCCC motif. There are two copies of TCCC mapping at nt 23-26 and 244-247, which flanked the cloverleaf putative secondary structure formed by single strand domain I. The second block perfectly conserved in Gallus species has sequence similarity to mammalian TASs including the highly conserved motif GRYCAT. These GRYCAT and TAS motifs are included in the putative ETAS1 sequence block of Gallus species. There are repeats of the sequences, which resembles with CSB1 and these repeats showed the GGACAT motif. These repeats varied between the Gallus species. All the *G. gallus* sub species as well as *G. varius* showed two repeats, while the *G. sonneratii* and *G. lafayettei* had three repeats. The single stranded domain I can form

potential secondary cloverleaf structures and this structure is highly conserved between Gallus species.

In the domain II, the conserved F, D and C boxes are readily recognizable. The F box showed absolute conservation among *G. gallus* subspecies, however only *G. sonneratii* and *G. lafayettei* showed only one nucleotide substitution (C to T). The D box and C box showed absolute sequence homology in all the Gallus species.

In domain III, the CSB-I was absolutely conserved between Gallus species,, however the CSB-II and CSB-III sequences could not be unambiguously identified. A poly (C) sequence similar to the O_H of mammals was identified just a few nucleotide downstream from the putative CSB-I. The bi-directional LSP/ HSP promoters were perfectly conserved in Gallus species. This sequence contains an almost perfect inverted repeat, which can be folded into stable hairpin. Farther upstream, is the most variable part of the CSB domain, which is rich in poly(T) and poly(A) strings.

Overall low genetic distances were observed within as well between the species under genus Gallus. The *G. g. murghi* showed very low genetic distances with the chicken breeds (0.008) as well as with other *G. gallus* sub-species (0.003-0.012), however with other jungle fowl species, it showed higher distance (0.049-0.060). In general all the other three jungle fowl species i.e. CyJF, GyJF and GJF showed comparatively much higher genetic distances not only from RJF subspecies including chicken breeds (0.049-0.060), but also from each other (0.042-0.051). The phylogenetic tree based on nucleotide sequence variability of complete D loop region showed the distantness of the *G. gallus* with other three jungle fowls i.e. *G. sonneratii*, *G. lafayattei* and *G. varius* as they make a separate cluster. Between *G. gallus* subspecies, *G. g. bankiva* and *G. g. gallus* seemed to be distant from *G. gallus* spadiceus, *G. g. murghi* and *G. g. domesticus*. *G. g. murghi* and *G. g. domesticus* showed maximum closeness. These phylogenetic analysis also clearly support very high genetic relatedness of *G. g. domesticus* with *G. gallus* however, the relative closeness of *G. g. domesticus* with other *G. gallus* subspecies is not conclusive due to almost equal distance of all these red jungle fowl subspecies from most of the domestic chicken breeds.

The Domain I of the D loop region was found to be most polymorphic. Hence the nucleotide sequence variability in this region was used to establish the phylogenic relationship of Indian red jungle fowl (*G. g. murghi*) with the other three subspecies of Red Jungle Fowl. While very low genetic distances were observed within *G. g. spadicieus* (0.003-0.008) and *G. g. bankiva* (0.013-0.015), *G. gallus* showed higher genetic distances range (0.008-0.069) within themselves. The *G. g. murghi* showed very low genetic distances (0.008-0.013) with the one group of *G. g. gallus* (*GG_gallus*8. *GG_gallus*10 and *GG_gallus*3322) in comparison to other *G. g. gallus* group i.e. *GG_gallus*11, *GG_gallus*39 and *GG-gallus*58 (0.018-0.034). The genetic distances between *G. g. murghi* and *G. g. bankiva* were higher (0.061-0.066) than those observed between *G. g. murghi* and *G. g. spadiceus* (0.031-0.034).

The phylogenetic analysis revealed that among the different red jungle fowl subspecies, *G. g. banikva* seemed to be well separated from the other three subspecies as it made a separate cluster. Among these three subspecies, *G. g. gallus* makes two separate clusters i.e. one those from Thailand (*GG_gallus*8. *GG_gallus*10 and *GG_gallus*3322) and other those from Japan and Indonesia (*GG_gallus*11 and *GG_gallus*39 and *GG_gallus*58). All the *G. g. spadiceus* falls in one sub-cluster. The *G. g. murghi* falls in one sub-cluster along with the *G. g. gallus* from Thailand.

Mini Abstract

Present study has been aimed for genotyping the potential SNPs in mitochondrial genome by allele specific PCR in red jungle fowl (RJF) and domestic fowl population and estimating genetic diversity between RJF and domestic fowl using the polymorphism in mitochondrial D loop. On sequence homology comparisons of eight complete mitochondrial genome sequences *G. g. domesticus* and subspecies of red jungle fowl, 19 population specific SNPs were identified and for six of these primers with good score could be designed. For each SNP, two sets of primers were designed. A resource population of 18 RJF individuals and 18 DF individuals were genotyped for these 6 SNPs. The allele specific primers were effective in amplifying the specific alleles. At each SNP site, the specific allele showed very high frequency in the respective populations, but these alleles were not fixed.

A 1325 bp fragment having the complete D loop region comprising of 45 nt from 3' end of tRNA-Glu, complete D loop region (1232 nt) and 48 nt from 5' end of tRNA-Phe was amplified, cloned and sequenced in Indian Red Jungle Fowl (In-RJF). The nucleotide sequence of D loop from In-RJF was aligned and compared with the related sequences from other Gallus species, *G. gallus* subspecies including *G. g. domesticus*. The length of D loop region showed ranged from 1227 in WPR to 1292 bp in CyJF. For the D loop region, very low variability was observed within as well as between the *G.g. domesticus* and *G. gallus* subspecies (0.5 %-1 %) and all the nt substitutions observed were transitions only. Between the *Gallus* species, high polymorphism was observed. Apart from major addition / deletions in D loop region of other jungle fowls, 127 nt substitutions (83 transitions and 44 transversions) were observed between *Gallus* species.

The D loop region showed three distinct domains i.e. highly variable left domain or domain I (DI), highly conserved central domain or domain II (DII) (451-900 nt) and right domain or domain III (DIII) in all the Gallus sequences including In-RJF. In DI, sequence capable to produce a goose hairpin like structure, two copies of TCCC motifs flanking the cloverleaf putative secondary structure formed by single strand domain I, sequences similarity to mammalian TASs including the highly conserved GRYCAT and TAS motifs and the repeats of the CSB1 like sequences having the GGACAT motif were found. These repeats varied between the *Gallus* species from two in all the *G. gallus* sub species and *G. varius* to 3 in *G. sonneratii* and *G. lafayettei*. In the domain II, the conserved F, D and C boxes are readily recognizable. In domain III, a poly (C) sequence similar to the O_H of mammals, the CSB-I sequence and the bi-directional LSP/HSP promoters were identified and found to be perfectly conserved in *Gallus* species.

The *G. g. murghi* showed very low genetic distances within as well as between the different subspecies of RJF, while between the jungle fowl species, the genetic distances were high (0.042-0.060). The phylogenetic tree based on nt sequence variation in D loop, In-RJF showed distinct closeness with domestic chicken. All other three jungle fowls made a separate cluster. The nt sequence variation in the first 400 nucleotides of D loop region was studied between the different RJF subspecies including In-RJF. In the phylogenetic tree, *G.g. banikva* seemed to be well separated from the other three subspecies. Among these three subspecies, *G.g. gallus* makes two separate clusters i.e. one those from Thailand and other those from Japan and Indonesia. All the *G.g.* spadiceus falls in one sub-cluster. The *G.g.* murghi species in one sub-cluster along with the *G.g. gallus* from Thailand.

लघु सारांश

प्रस्तुत शोध मुख्यतः रेड जंगल फाऊल (आर जे एफ) और देसी मुर्गी में एलील विशेष पी.सी.आर. द्वारा माईटोकान्ड्रियल जीनोम में सम्भावित एस.एन.पी. के आंकलन के लिये और माईटोकोन्ड्रियल डी लूप में बहुरूपता के द्वारा रेड़ जंगल फाऊल और देसी मुर्गी की आनुवांशिक भिन्नता मापने के लिये किया गया। जी गैलस डोमेस्टिकस और रेड़ जंगल मुर्गी की उपप्रजातियों के सम्पूर्ण माईटोकान्ड्रियल जीनोम क्रम की समरूपता तुलना करने पर 19 जनसंख्या–विशेष एसएनपी की पहचान की गई और इनमें से छः के लिये उत्तम अंको के प्राईमरर्स निर्मित किये गये। प्रत्येक एस.एन.पी. के लिये दो संग्रह प्राईमर बनाये गये। 18 आर जे एफ पृथक नमूनों व 18 डी.एफ. नमूनो को इन छः एस.एन.पी. के लिये जीनोटाईप किया गया। एलील–विशेष प्राईमर इन एलील के संवर्धन के लिये संक्षम पाये गये। प्रत्येक एसएनपी स्थल विशेष एलील ने अपनी जनसंख्या में बहुत अधिक तीव्रता दर्शायी परन्तु ये एलील स्थिर नहीं थे। भारतीय रेड जंगल फाऊल के 1325 बेस युक्ष्मों के खण्ड का, जिसमें टी–आर.एन.ए. ग्लुटामिन के 3′ सिरे से 45 न्यूक्लियोटाईड सम्पूण डी लूप क्षेत्र (1232 न्यूक्लियोटाईड) और टी–आर.एन.ए.–फ्रिनाईल एलानिन के 5′ सिरे से 48 न्यूक्लियोटाइड थे, का सवंर्धन किया और इसे क्लोन करके क्रमित किया गया। आर.जे.एफ. के डी–लूप के न्यूक्लियोटाईड क्रम की अन्य गैलस प्रजातियों जी गैलस डोमेस्टिकस समेत जी गैलस उपप्रजाति संबंधित क्रमों के साथ मिलान करके तुलना की गई। डी–लूप क्षेत्र की लम्बाई 1227 से 1292 तक पाई गई। डी–लूप क्षेत्र के लिये जी गैलस डोमेस्टिकस व जी गैलस उपप्रजातियों में बहुत कम विभिन्नता (0.5% - 1%) पाई गई और सभी न्यूक्लियोटाईड स्थानान्तत्रण केवल ट्रासिशन थे। गैलस प्रजाति में अधिक बहुरूपता देखी गई। अन्य जंगल मुर्गियों के डी–लूप क्षेत्र में मुख्य योजन/घटन के अतिरिक्त, गैलस प्रजातियों में 127 न्यूकिलयोटाईड स्थानान्तरण (83 ट्रानजिशन व 44 ट्रान्सर्वशन) देखे गये। आर.जे.एफ. सहित सभी गैलस क्रमों में डी लूप क्षेत्र ने तीन विभिन्न प्रदेश दर्शाये जैसेकि अत्यधिक विभिन्नता वाला बाँया प्रक्षेत्र I (डी–I), अधिक समानता वाला मध्य प्रक्षेत्र II (डी–II) और दाँया प्रक्षेत्र या डी–III । डी–I में, हेयर पिन जैसी संरचना पैदा करने में सक्षम क्रम, सिंगल सट्रेडिड़ प्रक्षेत्र 1 द्वारा निर्मित प्रसिद्ध क्लौवरलीफ संरचना के बाहर TCCC मोटीफ, अत्याधिक समानता वाले GR YCAT और TAS मोटीफ सहित स्तनधारी TAS जैसे क्रम, और CSB1 जैसे क्रमों के दोहरान पाये गये। ये दोहराव गैलस प्रजातियों में भिन्न थे, जैसे जी गैलस उपप्रजातियों और गैलस वैरीयस में 2 से गैलस सानेराटि और गैलस लैफेयटि में 3 तक थे। प्रक्षेत्र 2 में समरूप एफ डी व सी बक्से आसानी से पहचाने जा सकते हैं। प्रक्षेत्र–III में स्तनधारियों के OH के जैसे बहुलक (सी) क्रम, सीएसवी–1 क्रम वद्धि दिशागामी एल.एस.पी. / एच.एस.पी. प्रोमोटर पहचाने गये और गैलस प्रजाति में संरक्षित पाये गये। जी गैलस मुरगी ने रेड जंगल फाऊल के विभिन्न उप–प्रजातियों में बहुत कम आनुवांशिक दूरी दर्शायी। जबकि जंगल फाऊल प्रजाति में आनुवांशिक दूरी अधिक (0.042-0.060) थी। डी–लूप में न्यूक्लियोटाई विभिन्नता पर आधारित वंश वृक्ष में, भारतीय रेड जंगल फाऊल ने घरेलू मुर्गी के साथ भिन्न समीपता दर्शायी। अन्य सभी तीन जंगली मुर्गियों ने पृथक समूह बनाया। डी लूप क्षेत्र के पहले 400 न्यूक्लिोटाईड में क्रम भिन्नता का भारतीय रेड जंगल फाऊल सहित सभी विभिन्न रेड जंगल फाऊल उप प्रजातियों में अध्ययन किया गया। वंश वृक्ष में जी गैलस बैनकिवा अन्य तीन उप प्रजातियों से अच्छी तरह पृथक प्रतीत होती है। इन तीन उपप्रजातियों में जी गैलस गैलस दो पृथक समूह बनाती है जैसेकि एक थाईलैंड से और दूसरा जापान व इन्डोनेशिया से। सभी जी.जी. प्रजातिया एक उप–समूह के अर्न्तगत आती है। जी गैलस मुरगी एक उप–समूह में थाईलैड से जी गैलस गैलस के साथ आती हैं।

Bibliography

Ali, S., Muller, C.R. and Epplen, J.T. (1986). DNA fingerprinting by oligonucleotide probes specific for simple repeats. *Human Genetics*, 74: 239-43.

Amills, M., N. Jimenez, D., Vilalba, M., Tor, E., Molina, D., Cubil, C., Marcos, A., Francesch, A., Sanchiz and Estany, J. (2003). Identification of 3 SNPs in the chicken Insulin like growth factor 1 and 2 gene and their association with growth and feeding traits. *Poultry Sci.* 82: 1486-1493

Anderson, S., Bankier, A.T, and Barrel,B.G. (1981). Sequence organigation of human mitochondrial DNA genome. *Nature*, 290 : 457-465.

Bellagamba, F., Valfre, F., Panseri, S and Moretti, V. M(2003). Polymerase chain reaction – based analysis to detect terrestrial animal protein in Fish meal. *J. Food Prot.* 66 (4) : 682-685.

Bharani, K.S.T., Neeraj. D., S.P.S. Ahlawat., Mishra. B., Tantia . M.S and R.K. Vijh (2007). Genetic Relationship among Chicken Populations of India Based on SNP Markers of Myostatin Gene (GDF 8). *International Journal of Poult. Sc.* 6 (9): 684-688.

Boles, T.C., Snow, C.C., Stover, E. (1995). Forensic DNA testing a remains from mass graver: A pilot project in Guatemala. *J. Forensic Sci.* 90: 349-355.

Bundock, P.C., Cross, M., Shapter, F and Henry, R.J. (2005). Development Of Robust Allele Specific PCR Based SNP Markers For Barley. *In Plant & Animal Genomes XIII Conference January, 15-19,Town & Country* Convention Center, San Diego, CA. P 182.

Colgan, S., Brien, L., Maher, M., Shilton, N., McDonnell, K. and Ward, S. (2001). Development of a DNA based assay for species identification in meat and bone meal. *Food Res. Int.*, 34(5) : 409-414.

Colombo, F., Viacava, R. and Griaretti, M. (2000). Differentiation of the ostrich (*Struthio camelus*) and emu (*Dromaius novaehollandiae*) species by polymerase chain reaction using an ostrich-specific primer pair. *Meat Sci.*, 56(1): 15-17.

Crochet, P. and Desmarais, E. (2000). Slow Rate of Evolution in the Mitochondrial Control Region of Gulls (Aves: Laridae). *Molecular Biology and Evolution* 17:1797-1806.

Desjardins, P and Morais, R (1991). Nucleotide sequence and Evolution of Coding and Noncoding Regions of a Quail Mitochondrial Genome. *J Mol Evol.* 32: 153-161.

Desjardins, P. , V. Ramirez, and R. Morais (1990). Gene organization of the Peking duck mitochondrial genome. *Curr Genet.* 17 : 515-518.

Desjardins, P. and Morais, R. (1990). Sequence and gene organization of the chicken mitochondrial gene : a novel gene order in higher vertebrates. *J. Mol. Biol.*, 212 : 529-534.

Dodgson, J.B., Cheng, H.H. and Okimoto, R. (1997). DNA marker technology: A revolution in animal genetics. *Poultry Science*, 76 : 1108-1114.

Dong, N., Yan, F.U., Jing, L., Ruan, H., Yu, X., Gong, C. and Zhang, Y. (2002). The origin and genetic diversity of chinese native chicken breeds. *Biochemical Genetics*, 40(5/6) : 163-174.

Douabin-Gicquel, V., Soriano,N., Ferran, H.,Wojcik,F., Palierne, E., Tamin, S., Joveli,T., Mc Kei, A.T., Le,Gall, J.Y., David,V. and Mosser, J. (2001). Identification of 96 SNPs in eight genes involved in iron metabolism: efficiency of bioiinformatic extraction compared with a systematic sequencing approach. *Human Genet.*, 109 : 393-401.

Douzery,E and Randi, E (1997).The mitochondrial control region of Cervidae: evolutionary patterns and phylogenetic contents.*Mol.Biol.Evol.* 14: 1154-1166.

Drenkard, E., Richter, B. G. Rozen, S., Stutius, L.M., Angel,N. A., Mindrions, M. , Chos, R. J., Oefner, P. J., Davis, R. W., Ausubel, F.M (2000). A simple procedure for the analysis of single nucleotide polymorphism facilities map- based cloning in Arabidopsis. *PlantPhysiol.* 124 : 1483-1492.

Dufrense, C., Mignotte, F., Gueride, M (1996). The presence of tandem repeats and the initiation of replication in rabbit mitochondrial DNA. *Eur. J. Biochem.* 235 : 593-600.

Ferrari, F., Stella, S. and Cantoni, C. (2001). Identification of chicken and turkey meat with PCR. *Ingegneria Alimentare le Conserve Animali,* 17(5): 28-29.

Fu, Y .D. , Niu and Ruan,H (2001). Studies of genetic diversity of native chicken breeds in Zhejiang province of China. *Acta Genetica Sinica.* 28 : 606-613.

Fumihito, A., Mayake, T., Takada, M., Ohno, S. and Kando, N. (1995). The genetic link between the Chinese bambro partridge (*Bambiusicola tholacica)* and the chicken and jungle fowls of the genus Gallus. *Proc.Natl.Acad. Sci, USA.* 92(24): 11053-11056.

Fumihito, A., Miyake, T., Takada, M., Shingu, R., En, T., Gojobori, T, Kondo, N., Ohno, S. (1996). Monophyletic origin and unique dispersal patterns of domestic fowls. *Proc.Natl. Acad. Sci, USA.* 93(13): C792-C795.

Fumihito, A.,Mayake, T., Sumi S. I. Takada,M., Ohno,S., and Kondo, N. (1994). One subspecies of the red jungle fowl (*Gallus gallus gallus*) suffices as the matriarchic ancestor of all domestic breeds. *Proc.Natl.Acad. Sci, USA.* 91 : 12505-12509.

Gaudet, M., Fara A., Sabatti M., Kuzminsky, E. and Mugnozza, G.S. (2007). Single-reaction for SNP Genotyping on Agarose Gel by Allele-specific PCR in Black Poplar (*Populus nigra* L.) *Plant Molecular Biology Reporter.* 25 (1-2) : 1-9

Girish, P, Anjaneyulu,A, Viswas,k, Sonthosh,F, Bhilegaankar,K, Agarwal,R, Kondaiah, N, Nagappa, K. (2007). PCR-RFLP of mitochondrial 12S rRNA gene: A simple method of identification of poultry meat species. *Vety. Research Communication,* 31(4): 447-455.

Glaus, K.R., Zassenhaus, H.P., Fechheimer, N.S. and Perlman, P.S (1980). In the organization and expression of the mitochondrial genome (Kroon A.M. and Saccone, C., eds) Vol.2. pp 131-135, Elsevier, North Holland Biochemical Press, Amsterdam.

Guan, X., Geng, T, Silva, P and Smith, E.J. (2007). Mitochondria) DNA Sequence and Haplotype Variation nalysis in the Chicken (*gallus gallus*). *Journal of Heredity*. 98 (7) : 723-726.

Gupta, J., Singh, A., Churchil, R.R., Singh, B.P.and Sharma D. (2005). Genetic divergence between red jungle fowl and other domesticated poultry species using 12S rRNA gene polymorphism. *India J Anim Genet. & Breeding,* 26 : 26-30.

Halushka, M.K., Fan, J.B., Bently, K., Hsie, L., Shen, N., Weder, A., Cooper,R., Lipshutz, R. and Chakravarti, A. (1999). Pattern of single nucleotide polymorphsms in candidate genes for blood-pressure homeostasis. *Nat Genet.,* 22 : 239-247.

Haring, E., Kruchenhauser, L., Gamauf, A and Riesing (2001).The complete sequence of mitochondrial genome of Buteobuteo Aves, Accipitridae indicates an early split in the phylogeny of raptors. *Mol.Biol.Evol.* 1810:1892-1904.

Harlid, A., Janke, A., Arnason, U. (1997). The mt DNA sequence of ostrich and the divergence between palcognathous and neognathous birds. *Mol. Bio. Evol.* 14(7): 754-761.

Harumi, T., Sano, A., Kagami, H., Tagami, T., Matsubara, Y. and Naito, M. (2004). PCR detection of SNP in the chicken mt D-loop region. *Animal-Science-Journal,* 75 (6) : 503-507.

Hashiguchi, T.T., Nishida,Y., Hayashi, Y., Maeda, and Manjoer, S.S (1983). Blood protein polymorphism of native and jungle fowls in Indonesia. *Asian-Aust.J.Anim. Sci.* 6 : 27-35.

Herman, L. (2001). Determination of the animal origin of raw food by species-specific PCR. *J. Dairy Res.,* 68(3) : 429-436.

Ingman, M., Kaessmam, H., Paabo,S. and Gyllensten, U (2001). Mitochondrial genome variation and the region of modern human. *Nature,* 408 : 708-713

Jeffrey, A.J, Wilson, V. and Thein, S.L. (1985). Hypervariable 'minisatellite' regions in humans. *Nature,* 314 : 67-73.

Kamiyama, T., Keo, K., and Gojabori, T. (2004). The evolutionary origin of long crowing chicken: its evolutionary relationship with fighting cocks disclosed by the mt DNA sequence analysis. *Gene* 333 (1/2) : 91-99.

Kanginakudru, S., Metta, M., Jakati R.D. and Nagaraju, J. (2008). Genetic evidence from Indian red jungle fowl corroborates multiple domestication of modern day chicken. *BMC Evolutionary Biology.* 8:174

Kim,H., Schmidt,C.J., Decker, K.S, Emara, M.G. (2002). Chicken SNP discovery by EST data mining, in : Plant, Animal & Microbe Genomes X , 12-16 January, San Diego, http:// www.intl-pag.org/pag/10/ abstracts/ PAGX_p246.html.

Kline, M.C., Vallone, P.M., Redman, J.W., Deuwer, D.L., Calloway, C.D. and Butler, J.M. (2005). Mitochondrial DNA typing screens with control region and coding region SNPs. *J. Forensic Sci.*, 50(2) : 377-385.

Kwok, S., Kellong,D.E., McKinney,N., Spasic,D., Goda, L., Lovenson, C., Sninsky, J.J (1990). Effect of primer-template mismatch on the polymerase chain reaction : human immunodeficiency virus type 1 model studies. *Nucleic Acis Res*. 18 : 999-1005.

L'Abbe, L.D., Duhaime, J.F., Lang, B.F. Morais, R. (1991). The transcription of DNA in chicken mitochondria initiates from one major bidirectional promoter. *J. Biol. Chem*. 266(17): 10844-10850.

Lee, V. J., M. S. A. Bhuiyan, H. J. Chung, W. Y. Jung, K. D. Choi, B.G.Jang,W.K.Paek, J. T. Jeon, C. S. Park and I. H. Lee (2007). Mitochondrial DNA Diversity of Korean Ogol Chicken. *Asian-Aust. J. Anim. Sci.* 20 : 477-481.

Lee, Y.H.W., Liaw, L.L., Yamg, T.T. Lo, S.J. (1989). Nucleotide sequence of the structural genes for the mt Asp, Cye, Ser, tRNA from chicken. *Nucleic. Acid Res*. 17(22): 9477.

Li, S., Aggrey, S.E., Zadworny, D., Fairfull, W., nKuhnlein, U. (1998). Evidence for a genetic variation in the mt genome affecting traits in white leghorn chickens. *J. of Heridity*, 89(3): 222-226.

Lie, M.M. Nie, Q.H., Pong, X., Zhag, D.X. and Zhang, X.Q. (2006). SNPs of chicken Insulin like factor binding protein 2 gene associated with chicken growth and carcass traits. *Poultry Sci.* 84: 1191-1198.

Lie, M.M., Nie, Q.H., Pong, X., Zhag, D.X. and Zhang, X.Q. (2005). SNPs of chicken Insulin like factor binding protein 2 genes associated with chicken growth and carcass traits. *Poultry Sci.*, 84 : 1191-1198.

Liu, D., Gai, X., Hu, B., Zhou, Y., nChen, X. (1998). Polymorphism in mt DNA of six fowl breeds revealed by restriction endonuclease. *J. Northeast Agri. Uni. English edition*. 5(2): 116-152.

Liu, H., Hu, Y., Wang, C., Lin, L. (1996). Sequence and comparisons of duck mtDNA control regions. *Comparative Biochemistry and Physiology-B-Bio and Mol. Bio.* 115(2): 209-214.

Liu, Y.P., Wu, G,S.,Yao, Y.G., Miao,Y.W., Luikart,G., Beja,P.A., Ding, Z.L., Palanichamy, M.G and Zhang,Y.P (2006). Multiple maternal origin of chickens: out of the Asian jungle. *Mol. Phylo. and Evol.*38 :112-119.

Liu, Z.W., Lie, C.Z., Luo, J., Ding, C., Chen, G.H., Chang, H., Wang, K.H., Liu, X.X., Zhang, X.Y., Xiao, X.J., Wu, S.L. (2004). Genetic variability of mt DNA sequence in chicken native breeds. *Asian Australian J. Ani. Sci.* 17(7): 903-909.

Marshall H. Dawn and Allan J. Baker (1997). Structural Conservation and Variation in the Mitochondrial Control Region of Fringilline Finches (*Fringilla spp.*) and the Greenfinch (*Carduelis chloris). Mol. Biol. Evol. 14: 173-194*

Martin I, Garcia T, Fajardo V, Lopez-Calleja I, Rojas M, Pavon M.A., Hernandez P.E., Gonzalez I, Martin, R. (2007). Technical note: detection of chicken, turkey, duck and goose tissues in feedstuffs using species-specific polymerase chain reaction. *J. Anim. Sci.,* 85(2): 452-458.

Matsunaga, T., Shibata, K., Yamada, J., Shinmura, Y. (1999). Effects of processing condition on species identification of meat products. *Journal of the Japanese Society for food Science and Technology,* 46(3): 187-194.

Meyer, R., Hoefelein, C., Luethy, J. and Candrian, U. (1995). Polymerase chain reaction-restriction fragment length polymorphism analysis: a simple method for species identification in food. *Journal of AOAC International*, 78(6): 1542-1551.

Mindell, D. P., Michael D. Sorenson and Derek E. D. (1998). Multiple independent origins of mitochondrial gene order in birds. *Proc. Natl. Acad. Sci. USA.* 95 : 10693–10697

Munechika, I., Suzuki, H., Wakana, S. (1997). Comparative analysis of the restriction endonuclease cleavage patterns of mtDNA in the genus Gallus Japanese. *Poultry Sci.* 34(3): 184-188.

Newton, C .R., Graham, L .E., Heptinstall, S. J., Poweu and Summers, C (1989). Analysis of any point mutation in DNA, the amplification refractory mutation system (ARMS). *Nucleic Acid Res*. 17 : 2503-2516.

Nishibori, M., Hayashi, T. and yasue, H. (2004). Complete nucleotide sequence of Numida maleagris (*Helmeted guineafowl*) mitochondrial genome. *J. Poultry Sci.,*41:259-268.

Nishibori, M., Hayashi, T., Tsudzuki, M., Yamamoto,Y. and Yasue, H (2003). Complete nucleotide sequence of mitochondrial DNA in chicken, white leghorn and Plymouth Rock. *Animal Sci. J.* 74 : 437-439.

Nishibori, M., Hayashi,T., Tsudzuki, M., Yamamoto,Y. and Yasue,H. (2001). Complete sequence of the Japanese quail (*Coturnix japonica*) mitochondrial genome and its genetic relationship with related species. *Animal Genetics*, 37: 380-385.

Nishibori, M., Shimogiri, T., Hayashi, T. and Yasue, H. (2005). Molecular evidence for hybridization of species in the genus Gallus except *G. varius*. *Animal Genet.*, 36(5) : 367-375.

Nishibori,M., Tsudzuki, M., Hayashi,T., Yamamoto,Y and Yasue,H (2002). Complete nucleotide sequence of *Coturnix chinensis* mitochondrial genome and phylogenetic analysis with related species. *J. of Heredity*, 93 : 439-444.

Niu, D. Y., Fu, J., Luo, H., Ruan, X. P., Yu, G., Chen, and Zhang, Y. P (2002). The origin and genetic diversity of Chinese native breeds. *Biochem. Genet.* 40(5/6) : 163-174.

Oka, T., Ino,Y., Nomura,K (2007). Analysis of mtDNA sequence shows Japanese native chickens have multiple origin. *Ani.Genetics*. 38: 287-293.

Pan, Y.E., Lee, Y.H.W., Wei, Y.H., Chiang, A.N. (1993). A gene for cytochrome C Oxidase Subunit II in duck mt DNA: Structural features and sequence evolution. *Biochem Mol. Biol. International* 30(3) : 479-489.

Peterson, A. T and Brisbin, I.L (1999). Genetic endangerment of wild red jungle fowl (Gallus gallus). *Bird Conservation International*, 8(4) : 387-394.

Qui, F. F., Nie, Q. H., Lud, C.L., Zhang, D.X., Lin, S.M. and Zhang, X.Q. (2006). Association of SNPs of insulin gene with chicken early growth and fat deposition. *Poult. Sci.*, 85 : 980-985

Quinn, T.W. and Wilso, A.C. (1993). Sequence Evolution in and Around the Ni itochondrial Control Region in Birds. *J. Mol. Evol.* 37 : 417-425.

Ramirez,V.,Savoie,P. and Morais, R. (1993). Molecular characterization of duck mitochondrial genome. *J. Mol. Evol.* 37 : 296-310

Randi, E. and Lucchini, V. (1998). Organization and Evolution of the mitochondrial DNA Control Region in the Avian Genus Alectoris. *J. Mol. Evol.* 47 : 449-462.

Ritchie, P and Lambert, D (2000). A repeat complex in the mitochondrial control region of Adelie penguins from Antartica. *Genome* 43 : 613-618.

Rodriguez, M.A., Garcia, T., Gonzalez, I., Asensio, L., Mayoral, B., Lopez-Calleja, I., Hernandez, P.E., Martin, R. (2003). Identification of goose, mule duck, chicken, turkey and swine in foie gras by species-specific polymerase chain reaction. *J. Agric Food Chem.*, 51(6):1524-1529.

Roques, S., Goddy, J.A. Negro, J.J and Hiraldo,F (2004).Organization and Variation of the Mitochondrial Control Region in Two Vulture Species, Gypaetus barbatus abd Neophron percnopteus. *J. of Heridity* 95 (4) : 332-337.

Saccone, C., Pesole. G and Sbisa. E (1991).The main regulatoru region of mammalian mitochondrial DNA : structure –function model and evolutionary pattern. *J. Mol.Evol.* 33 : 83-91.

Sachidanandam, R., Weissman, D., Schmidt, S.C., Kakol, J.M., Stein, L.D., Marth, G., Sherry, S., Mullikin, J.C., Mortimore, B.J., Willey, D.L. and others (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphism. *Nature,* 409: 928-933.

Saini, M., Das, D.K., Dhara, A., Swarup, D., Yadav, M.P., Gupta P.K. (2007). Characterization of peacocks (*Pava cristatus*) mitochondrial 12S rRNA sequence and its use in differentiation from closely related poultry species. *Br. Poult. Sci.*, 48(2) : 162-166.

Sambrook, J., Fritsch, E.E and Maniatis, T (1989).Molecular cloning . Cold Spring Harbor Laboratory, Cold Spring Harbor,N.Y.

Sano, A., Harumi, T., Hanzawa, S., Kawashima, T., (2005). New screening test for male germline chimeric chickens by polymerase chain reaction using single nucleotide polymorphism detection primers. *The Journal of Poult. Sc.* 42 : 152-157.

Sbisa, E., Tanzarello,F., Reyes,A., Pesole,G and Saccone, C (1997). Mammalian mitochondrial D-loop region structural analysis : identification of new conserved sequences and their functional and evolutionary implication. *Gene* 205 : 125-140.

Schmid,M., Nanda,I., Guttenbach,M., Steinlein, C., Hoehn, M., Schatr, M., Haaf, T., Weugend,S., Fries, R., Buerstedde, J.M., Wimmers, K., Burt, D.W., Smith, J. A., Hara,S., Law, A., Griffin, D.K., Bumstead, N., Kaufman, J., Thomson, P.A., Burke,T., Groenen, M.A., Crooijmans, R.P., Vignal,

A., Fillon,V., Morisson,M., Pitel,F., Tixier-Boichard, M., Ladjali-Mohammedi, K., Hillel, J., Maki-Tanila,A., Cheng, H.H., Delany,M.E., Burnside, J., Mizuno,S. (2000). First report on chicken genes and chromosomes, *Cytogenet. Cell Genet.*, 90 : 169-218.

Shen, X., Tsudzuki, M., Nakamira, T. (1999). Primer design and sequence information for Cyt b gene in the Chinese painted quail (*Excal factoria chinensis). Animal Sci. J.* 70(4): 240-242.

Shena, X., Kobayashi, Y., Mizutani, M., Ito, S. (2000). RFLP analysis of mtDNA in 14 laboratory sprains of Japanese quail. *Animal Sci. J.* 71(2): 123-129.

Singh, B.P. and Sharma, Deepak (2007). Present scenario of egg and meat production in India. In Model Training Course on "Alternate Poultry Production System for Diversified Needs" held on 18th-25th September 2007 at CARI, Izatnagar

Sironi, L., Lazzan, B,. Ranulli, P., Gorni, L. and Mariani, P. (2006). SNP discovery in avian Tapasin gene. *Poult.Sci*. 25: 606-612

Soleimani, V.D. and Baumi, B.R. and Johnson, D.A. (2003). Efficient Validation of Single Nucleotide Polymorphisms in Plants by Allele-Specific PCR, With an Example From Barley. *Plant Molecular Biology Reporter* 21: 281–288.

Thomson, J.D., Higgins,D.G and Gibson, T.J (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighing, position-specific gap penalties and weight matrix choice. *Nucleic Acid Res*. 22 : 4673-4680.

Tomar, A., Shukla, P.K., Gupta, J., Singh, A., Churchil, R.R. and Sharma D (2007). Genetic polymorphism among red jungle fowl and domesticated chicken detected by DNA markers. *J. Appl. Anim Res*. 31 : 205-208.

Verkaar, E. L. C., Nijman, I. J., Boutaga, K. and Lenstra, J. A. (2002). Differentiation of cattle species in beef by PCR-RFLP of mitochondrial and satellite DNA. *Meat Science*, 60(4): 365-369.

Vigilant, L., Stonekij, M., Harpending, H., Hawaker, K., Wilson, A.C. (1991). African population and the evolution of human mt DNA. *Science* 253: 1503-1507.

Wada, Y., Yamada, Y., Nishibori, M., and Yasue, H (2004). Complete nucleotide sequence of mitochondrial genome of silkie fowl (*Gallus gallus domesticus*). *J. Poult. Sci*., 41: 76-82.

Wang, q., Li, H., LI, N., Lez, L., Wag, Y and Tang, Z. (2006). Identification of SNPs of adipocyte fatty acid binding protein gene and its association with fatness traits in the chicken. *Poultry Sci.* 85: 429-434.

Weber, J.L. and May, P.E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal Human Genetics,* 44: 388-396.

West, B and Zhou, B.X (1989). Did chicken go north ? New evidence for domestication. *World's Poult. Sci.* J. 45 : 205-218.

Yamamoto, Y., Murata, K., Matsura, H., Hosoda, T., Tamura, K and Furuyama, J (2000). Determination of complete nucleotide sequence and haplotypes in the D-loop region of the mitochondrial genome in the riental white strik (Ciconia boyciana). *Genes Genet. Syst.* 75 : 25-32.

Yan, Fu, Dong, N., Ruan, H., Jing, L and Zhang, Y. (2002). Studies on genetic relationship between Silky chicken and other breeds of chicken. *Chinese J. Ani. Sci.* 38(1): 7-10.

Ye, X., McLood, S., Elfick, D., Dekkers, J.C.M. and Lamont, S.J. (2006). Rapid identification of SNPs and estimation of allele frequencies using sequence traces from DNA pool. *Poult. Sci.*, 85: 1165-1168.

Zeuner, F.E (1963). A history of domesticated Animal. London Hutchinson.

Zhon, P., Zhang, T., Wang, w., Chen, Y., Zhang, Y. (1997). Study on variation of restriction cleavage pattern of mt DNA in local chickens of yunnen. *Chinese J. Ani. Sci.* 33(6) : 26-27.

Zhiliang, G.U., Zhang, H.F., Zhu, D.H. and Li, H. (2002). Single nucleotide polymorphism analysis of the chicken Myostatin gene in different chicken lines. *Acta Genet. Sin.*, 29: 599-632.

Zhiliang, G.U., Dahai, Z.H.U. Ning, L.I., Hui, L.I., Xuemei, D.E.N.G. and Changxin, W.U. (2004). The single nucleotide polymorphisms of the chicken myostatin gene are associated with skeletal muscle adipose growth. *Sci China,* 47: 26-31.

Zimmermann, S., Zehmer, R. and Mebs, D. (1998). Identification of animal species in meat samples by DNA-analysis. *Fleischwirtschaft,* 78(5): 530-533.

Appendix

Ampicillin Stock (100 mg/ml)

Ampicillin (Sodium salt, Hi Media)	1000 mg
Distilled water	10 ml

Sterilized by filtration. Stored at -20°C.

Gel loading dye (6x)

Ethidium Bromide	0.25%
Sucrose (w/v)	40.0%

Stored at 4°C.

IPTG (isopropyl-β-D- thiogalactopyranoside) (1 M)

IPTG	238 mg
Distilled water	1 ml

Sterilized by filtration. Stored at 4°C.

Luria Bertani medium (LB medium)

Tryptone	10 gm
NaCl	5 gm
Yeast extract	5 gm
Distilled water to make	1 liter

Adjust pH to 7.5 using 10N NaOH Sterilize by autoclaving.

To prepare LB Agar add 1.5% agar to LB medium. Where indicated ampicillin is added to the medium 100 μg/ml.

Phosphate buffer saline (PBS)

	(1x	(10 x)
$Na_2HPO_4.2H_2O$	1.44 gm	14.4 gm
or		
$NaHPO_4$	1.10 gm	11.0 gm
KH_2PO_4	0.2 gm	2 gm
KCl	0.2 gm	2 gm
NaCl	8.0 gm	80 gm
Distilled water	1000 ml	1000 ml

SOB medium

Tryptone	2.0000 gm
NaCl	0.0500 gm
KCl	0.0186 gm
2M Mg^{++}	1.0000 ml
Yeast extract	0.5000 gm

Distilled water to make 100ml. Sterilized by Autoclaving.

SOC medium

SOB medium	980 ml
1M glucose	20 ml

Sterilized by filtration.

Solution I (Cell re-suspension solution)

Glucose	50 mM
Tris-HCl (pH 8.0)	25 mM
EDTA (pH 8.)	10 mM

Prepare solution I from standard stocks in batches of ~100 ml, autoclaved for 15 min at 15 psi on liquid cycle. Stored at 4°C.

Solution II (cell Lysis buffer)

Sodium Hydroxide	0.2 N
(Freshly diluted from a 10N stock)	
Sodium dodecyl sulfate	1% w/v

Prepare solution II fresh and stored at room temperature.

Solution III (Neutralization solution)

5 M Potassium acetate	60 ml
Glacial acetic acid	11.5 ml
Distilled water	28.5 ml

The resulting solution will be 3M respect to potassium and 5M with respect to acetate. Store the solution at 4°C and keep on ice just before use.

TBE (10X) (Tris Borate EDTA) Buffer, pH 8.3

Tris base	109.0 gm
Boric acid	55.0 gm
EDTA	9.3 gm

Add distilled water to make final volume up to 1000 ml.

TCM (Tris Calcium Magnesium) Solution

1 M Tris. HCI, pH 7.5	11 ml
2M $CaCl_2$	5.5 ml
1 M $MgCl_2$	11 ml
Sterile distilled water	972.5 ml

TE Buffer

Tris HCl (pH 8.0)	10 mM
EDTA	1 mM

TSS medium

2X LB Medium	5 ml
30% PEG +60mM $MgCl_2$	5 ml
DMSO	750 ml
Sterile Distilled water	4.250 ml

X-gal (5-bromo-4-Chloro-3-indolyl-β-D-galactopyranoside) (1000x)

X-gal	25 mg
N,N' Dimethyl formamide	1 ml

Stored at -20^{0}C.z