MOLECULAR CLONING AND CHARACTERIZATION OF ALPHA LACTALBUMIN GENE IN VECHUR CATTLE

RAJEEV. M.

Thesis submitted in partial fulfilment of the requirement for the degree of

Master of Veterinary Science

Faculty of Veterinary and Animal Sciences
Kerala Agricultural University, Thrissur

2007

Department of Animal Breeding and Genetics
COLLEGE OF VETERINARY AND ANIMAL SCIENCES
MANNUTHY, THRISSUR-680651
KERALA, INDIA

Dedicated to

My Parents & Brother

DECLARATION

I hereby declare that this thesis, entitled **"Molecular Cloning and Characterization of Alpha Lactalbumin Gene in Vechur Cattle"** is a bonafide record of research work done by me during the course of research and that the thesis has not previously formed the basis for the award to me of any degree, diploma, associateship, fellowship or other similar title, of any other University or Society.

Mannuthy
17.03.2008.

RAJEEV. M

CERTIFICATE

Certified that this thesis, entitled **"Molecular Cloning and Characterization of Alpha Lactalbumin Gene in Vechur Cattle"** is a record of research work done independently by **Dr. Rajeev. M,** under my guidance and supervision and it has not previously formed the basis for the award of any degree, diploma, fellowship or associateship to him.

Mannuthy
17.3.08

Dr. T. V. Aravindakshan
(Chairman, Advisory Committee)
Professor
Centre for Advanced Studies in
Animal Genetics and Breeding
College of Veterinary and Animal Sciences,
Mannuthy, Thrissur

CERTIFICATE

We, the undersigned members of the Advisory Committee of **Dr. Rajeev. M,** a candidate for the degree of Master of Veterinary Science in Animal Breeding and Genetics, agree that this thesis entitled **"Molecular Cloning and Characterization of Alpha Lactalbumin Gene in Vechur Cattle"** may be submitted by **Dr. Rajeev. M,** in partial fulfilment of the requirement for the degree.

Dr. T. V. Aravindakshan
(Chairman, Advisory Committee)
Professor
Centre for Advanced Studies in Animal Genetics and Breeding
College of Veterinary and Animal Sciences,
Mannuthy, Thrissur

Dr. K. V. Raghunandanan,
Director
Centre for Advanced Studies in
Animal Genetics and Breeding,
College of Veterinary and
Animal Sciences, Mannuthy
(Member)

Dr. K. C. Raghavan,
Professor
Department of Animal Breeding,
Genetics & Biostatistics.
College of Veterinary and
Animal Sciences, Mannuthy.
(Member)

Dr. G. Krishnan Nair,
Professor & Head,
Department of Veterinary Microbiology
College of Veterinary and Animal Sciences.
(Member)

External Examiner
P. S. RAMAMATHULLA

Acknowledgement

I express my sincere and heartfelt sense of obligation and gratefulness to ***Dr. T.V. Aravindakshan,*** *Professor, Centre for Advanced Studies in Animal Genetics and Breeding and Chairman of the Advisory Committee, without whose common-sense, knowledge, perceptiveness, invaluable guidance, constant supervision, unlimited patience and timely help, I would never have finished.*

I am thankful to ***Dr. K.V. Raghunandanan,*** *Director, Centre for Advanced Studies in Animal Genetics and Breeding and Member of the Advisory Committee, for his words of inspiration, kindness and professional guidance during the entire period of research.*

There is no word to pay my gratefulness to ***Dr. K.C. Raghavan,*** *Professor, Centre for Advanced Studies in Animal Genetics and Breeding and Member of the Advisory Committee, for his wholehearted cooperation, incessant encouragement and affection.*

I owe my sincere thanks to ***Dr. G. Krishnan Nair,*** *Professor & Head, Department of Veterinary Microbiology and Member of the Advisory Committee, for his wholehearted help and suggestions offered in bacteriological works carried out through out the work.*

I am extremely greatful to ***Dr. Stephen Mathew,*** *Professor,* ***Dr. K.A. Bindu,*** *Associate Professor, of Department of Animal Breeding & Genetics and* ***Dr. Koshy John,*** *Associate Professor, Department of Veterinary*

Microbiology, for their suggestions and support during difficult times in the period of study.

I wish to place on record the invaluable help and understanding rendered to me by my colleagues and friends ***Dr. Jimcy Joseph, Dr. Nisha Valsan*** *and* ***Dr. Reshmi R.Chandran*** *throughout the course of the academic program.*

I have benefited a lot from the interaction with my seniors, ***Dr. Ariprasath and Dr. Seena*** *during the course of study. I also warmly acknowledge the timely help and support of my junior colleagues* ***Dr. Rojan, Dr. Abraham, and Dr. Bipin,*** *especially of* ***Rojan*** *who provided me valuable help in compiling the photographs.*

I wish to place my profound gratitude to ***Dr. Sunil G & Dr. Rajagopal*** *who helped me during the bacteriological experiments. My sincere thanks to* ***Drs. Hamza, Biju, Abhilash, Kishore, Binesh, Glady, Vivek, Prathiush, Vinod, Benoj, Roymon, Jinesh, Bibu, Harish, Prasath, Mahesh, Shaiby, Nishand, Prince, Aslam and Jesto*** *for their help and cooperation.*

I am thankful to my friends ***Dr. Manoj, Dr. Binoj & Dr. Benjamin*** *for providing me some key research articles for my work.*

Special thanks to ***Mr. Binoy & Mr. Arun*** *for their valuable help, suggestions and keen interest shown at every stage of this research work.*

I express my heartfelt gratitude to my friends ***Drs. Jose, Rajendran, George Sherin, Magnus, Bejoy, Siju, Pradeep*** *and* ***Lalu*** *for their moral support and encouragement.*

It is a pleasure to gratefully acknowledge the assistance and lively support of all my friends at COVAS.

I would like to acknowledge the ***staff and laborers of Vechur cattle farm*** *in my study sites for their wholehearted cooperation during the study.*

My sincere thanks are due to all ***Staff members*** *of the Department of Animal Breeding and Genetics.*

I am grateful to the ***Dean,*** *College of Veterinary and Animal Sciences, Mannuthy and* ***Kerala Agricultural University*** *for the facilities provided for this research work.*

I would like to say a big 'thank-you' to my ***Achan, Amma, Sajeev, Manju, Meenu, Manju's parents,*** *Manoj, Asha and* ***Ganesh*** *who provided me a constant source of love, affection and happiness.*

Above all, I thank my favorite Lord Kottarakara Ganapathy for being with me and strengthening me at all times...

Rajeev. M

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PLATES

Introduction

1. INTRODUCTION

Cattle play a vital role in the rural economy of India. Cow milk is the major source of animal protein in the diet of a large number of people. According to National Bureau of Animal Genetic Resources (NBAGR), there are thirty breeds of cattle in India and the latest addition is a dwarf cattle from Kerala called the Vechur cattle. The Vechur cattle became the first among the native cattle of Kerala to be approved as a distinct breed.

The Vechur cow has historical value and the celebrated work *Travancore state manual* has a particular mention of Vechur cows (Pillai, 1940). In order to increase milk production in the least possible time, attempts were made to transform all the local cows to crossbreds of exotic breeds and in the process, the populations and genetic base of several valuable indigenous cattle including Vechur began to shrink rapidly. However, the conservation efforts taken by Kerala Agricultural University (KAU) saved the Vechur cattle from the brink of extinction and eventually led to its recognition as one of the smallest among the described breeds of cattle in the world. The Vechur cattle, named after a village in Kottayam district where they were originally found, have now attracted international recognition and attention among the scientific community because of their smaller size and the considered medicinal value of their milk. The smaller size of the fat globules in the milk makes it easily digestible and hence it is widely recommended for children and convalescents. However, scientific studies on Vechur milk components and their beneficial effects are scarce.

The immunomodulatory properties of bovine milk and whey have long been documented. Alpha lactalbumin (α-LA) is one of the major whey proteins found in all milks studied so far. This protein has been described to have several physiological functions in the neonatal period. In the mammary gland, it participates in lactose

synthesis, and facilitates milk production. Scientific studies proved that the gene for alpha lactalbumin can be used as a genetic marker for milk production (Bleck and Bremel, 1993). Alpha-lactalbumin binds divalent cations (Ca, Zn) and facilitate the absorption of essential minerals, and it provides a well-balanced supply of essential amino acids to the growing infant. During its digestion, peptides appear to be transiently formed and these have antibacterial and immunostimulatory properties, thereby possibly aiding in the protection against infection. The partially folded form of the human α-LA has been found to have apoptotic effects (Svensson *et al.*, 2000). It is the primary protein in human milk. Cow milk contains α-LA, albeit less than human milk. As the α-LA protein is a major ingredient in infant formulae, information on the degree of homology of Vechur α-LA with human α-LA would help to determine its proportion of incorporation.

A study on nucleotide and amino acid sequence analysis would throw light on the Vechur α-LA gene architecture and the extent of functional and structural similarity with that of other species including human. The ease and automation of DNA sequencing have greatly facilitated the characterization of mammary cDNA and genes and, consequently, of milk proteins in various species. The expression analysis of native or modified genes using *in vitro* transcription systems, mammary cell lines, and transgenic animals, has greatly improved the knowledge of the functioning of milk protein genes. A wider knowledge of both the structure and function of this milk protein gene, will provide potential for genetic modification of milk composition.

This is an attempt to study the main features of the sequence, nature and patterns of exon-intron length distributions of the gene and the protein structure of α-LA of Vechur cattle and to compare its features with that of other species. The interspecies comparison of the α-LA protein sequence will also give an idea about

the evolutionary status of the animal. This study will be an initial step to reveal the hidden facts about the milk of the native dwarf cattle of Kerala.

The present study has been taken up with the following objectives:

1) To clone the α-LA gene from Vechur cattle.

2) To completely sequence the cloned α-LA gene.

3) To compare the sequence in Vechur cattle with that of human sequence to analyze the extent of similarities/differences between them.

Review of Literature

2. REVIEW OF LITERATURE

2.1 MILK PROTEIN GENES

The activity of milk protein genes is essential for the survival of mammals. They are also of paramount importance for human communities for which milk is an essential source of proteins and lipids. From a basic point of view, milk protein genes constitute an interesting model. They are expressed in a controlled manner and at a very high level quite specifically in mammary cells.

Banerjee (1976) suggested that milk protein genes (encoding a number of caseins and whey protein alpha lactalbumin) provide an attractive system for studying hormonal control since, during lactation, their expression is modulated in the epithelial cells of the mammary gland by an intricate combination of peptide and steroid hormones.

2.1.1 Alpha-Lactalbumin (α-LA) Gene

The product of α-LA gene is a small globular protein, approximately 14 kDa in weight, which in the majority of mammals consists of 123 amino acids and is present in the milk of all mammals.

Qasba and Safaya (1984) first reported the sequences of genes encoding whey proteins for rat α-LA, a 2 kb transcriptional unit divided into four exons.

Davies *et al.* (1987) described the organization, sequence and chromosomal localization of the human α-LA gene, which is assigned to chromosome 12q13.

According to Vilotte *et al.* (1987) the α-LA gene is localized on chromosome BTA 5q21 in cattle and is composed of four exons and three introns spanning 2023 bp in length.

Vilotte *et al.* (1991) described the isolation and characterization of a genomic clone encoding goat α-LA, and its sequence was aligned with that of its bovine counterpart. Southern analysis indicated the presence of several α-LA related sequences in the goat genome, as observed in cow and sheep, which suggests that the relevant duplication event(s) occurred before divergence of these ruminant species.

Mercier and Vilotte (1993) explained the advantage of better knowledge of both the structure and function of milk protein genes. This has already allowed the use of powerful techniques for the rapid identification of alleles and offers the potential for the genetic modification of milk composition.

2.2 ALPHA-LACTALBUMIN PROTEIN

According to Forsum (1973) α-LA, commercially once known as "lactalbumin" or whey protein, was traditionally considered a highly digestible protein and, owing to its balanced amino acid composition, it is of high biologic value.

Tucker (1981) reported that lactose is synthesized in the golgi apparatus of the mammary epithelial cells and is released into the milk along with α-LA. Dairy cows produce milk that contains approximately 1.2 mg/ml of α-LA and five per cent lactose.

Alpha-lactalbumin is the second most abundant whey protein component, making up approximately 20-25 per cent of the whey protein. It is the primary protein found in human breast milk and the important whey protein in cow milk and many other mammalian species. Alpha-lactalbumin is an excellent source of essential amino acids especially tryptophan and high in branched chain amino acids (Permyakov and Berliner, 2000).

Jackson *et al.* (2004) suggested that α-LA plays a nutritional role for the rapidly growing neonate as the protein in highest concentration in human milk. They reported that on average, α-LA contributed 16 per cent of the total nitrogen content of human milk and consequently an important part of the amino acid content.

2.2.1 Structure of α-LA

The primary structure of α-LA is very similar to that of c-type lysozyme and it has been suggested that the two proteins have evolved from a common ancestral gene (Brew *et al.*, 1968).

Fitzgerald *et al.* (1970) explained α-LA as the major protein in human milk whey and is secreted by the mammary gland epithelium. The native protein is one of two components in the lactose synthase complex which catalyses the final step in lactose biosynthesis in the lactating mammary gland.

Findlay and Brew (1972) observed that bovine α-LA has a molecular weight of 14.2 kDa, and a high degree of homology with the human form of the protein. Indeed 74 per cent of the amino acids are identical, and a further six per cent have a high degree of chemical similarity.

The bovine α-LA protein has four disulphide bonds, no free thiol, and is a calcium-binding metalloprotein, a structure stabilised by the calcium. Bovine α-LA consists of 123 amino acids and around 10 per cent of the bovine form of the protein is glycosylated. There are 4 disulphide bonds: Cys6-Cys120, Cys28-Cys111, Cys61-Cys77, and Cys73-Cys91 (Hiraoka *et al.*, 1980).

Vilotte *et al.* (1987) described α-LA as a calcium metalloprotein with 123 amino acids in most of the species with exceptions of 122 amino acids in rabbit and 140 in rat.

Acharya *et al.* (1989) studied Baboon α-LA which has a high degree of homology with the bovine and human forms of the protein, therefore obtained an indication of these protein structures. They determined the tertiary structures of baboon and human α-LA proteins by X-ray crystallographic methods. This protein consists of four α-helices, five 3_{10} helices, two β-structures, and an appreciable amount of random coil, representing 30, 20, 6 and 4 per cent, respectively.

Rao and Brew (1989) reported that α-LA has got high affinity for Ca^{2+} and other divalent cations and Ca^{2+} is essential for the folding and structural stability of the protein.

Acharya *et al.* (1991) reported that human α-LA has a molecular weight of 14 kDa and consists of 122-123 amino acids divided into two domains; a large α-helical domain and a small β-domain. The two domains are connected by a calcium binding loop and the native conformation is stabilized by four disulphide bonds.

Brew and Grobler (1992) observed that α-LA is a single-chain polypeptide of 123 amino acids in both human and bovine corresponding to a molecular mass of 14.070 kDa in human and 14.178 kDa in cattle. They also reported that both proteins contain four disulfide bonds and have 74 per cent homology between the amino acid sequences.

Permyakov and Berliner (2000) explained the structure of native α-LA which consists of two domains, a large α-helical domain and a small β-sheet domain, which are connected by a calcium binding loop. The α-helical domain is composed of three major α -helices (residues 5-11, 23-24, and 86-98) and two short 3_{10} helices (residues 18-20, and 115-118). The small β-sheet domain is composed of a series of loops, a small three-stranded antiparallel β-pleated sheet (residues 41-44, 47-50, and 55-56) and a short 3_{10} helix (three residues per turn and an intrachain hydrogen bond loop containing 10 atoms; residues 77-80). The two domains are divided by a deep cleft

between them. At the same time, the two domains are held together by the cysteine bridge between residues 73 and 91, forming the Ca^{2+} binding loop. A second important disulphide bridge 61-77 connects the two domains as well. Overall, the structure of α-LA is stabilized by four disulphide bridges (6-120, 61-77, 73-91, and 28-111).

2.2.2 Role of α-LA in Milk Synthesis

Brodbeck *et al.* (1967) observed that lactose synthase consists of two components, the A and B proteins, and they identified the B protein as alpha-lactalbumin.

Brew *et al.* (1968) reported that the A protein acts as a galactosyltransferase utilizing a variety of acceptors and found that partially purified bovine A protein had a very low endogenous lactose synthase activity in the absence of α-LA.

Andrews (1969) observed that α-LA of human and bovine milk was equally effective in stimulating the lactose synthase reaction, each with a K_m of 3 μ*M* (43μg/ml) in the presence of 80 m*M* glucose. The K_m for glucose was lowered in the presence of human α-LA, being 15, 8 and 3 m*M* at 100, 200 and 400 μg/ml, respectively.

Fitzgerald *et al.* (1970) suggested that the physiological function of α-LA is to lower the K_m of glucose so that it may be used maximally for the synthesis of lactose.

Hill and Brew (1975) found that α-LA is present in mammalian milks and plays a role in the biosynthesis of lactose by modulating the carbohydrate-binding properties of galactosyltransferase enzyme in the lactating mammary gland through a protein-protein interaction.

An increase in α-LA expression could lead to increased formation of the lactose synthase complex and an increase in lactose production. Lactose synthase is an enzyme of the transferase class that catalyses the transfer of galactose from UDP-galactose to glucose, forming lactose. The enzyme is a complex of the enzyme N-acetyllactosamine synthase and α-LA, which catalyzes the final phase of the lactose biosynthesis in milk (Kuhn *et al.*, 1980).

Larson (1985) explained the importance of higher concentrations of α-LA in the regulation of milk volume, by increasing the amount of lactose secreted into the milk.

According to Mincheva-Nilsson *et al.* (1990) α-LA is responsible for the protective effect of whey protein isolate, being four fold more effective than whey protein isolates itself.

Bleck and Bremel (1993) showed that the α-LA locus can be used as a genetic marker for milk production, and the marker may be directly responsible for the increased milk production. The α-LA variation serves as a useful marker on chromosome 5 for Holsteins. They suggested that levels of α-LA expression could be used as an indicator for selection of animals to produce high volumes of milk.

The α-LA protein takes part in lactose biosynthesis and indirectly in the regulation of water circulation and milk secretion by the secretary cells of mammary gland. Within the mammary epithelial cells, α-LA forms a complex with the membrane bound enzyme β-1,4-galactosyltransferase to form lactose synthase (Voelker *et al.*, 1997).

Boston *et al.* (2001) concluded that increased expression of α-LA in the mammary gland can increase milk production by increasing lactose synthesis. The

lactose synthase enzyme, a complex of α-LA and β1,4-galactosyltransferase, is apparently limited by the availability of α-LA.

Lonnerdal and Lien (2003) concluded from their studies that α-LA is the predominant protein in human milk and it appears that during evolution several roles have developed for this protein. Besides its roles of being part of the lactose synthase complex and a source of amino acids, recent research suggests that α-LA can be physiologically active by affecting gut microflora, by enhancing mineral absorption, by stimulating immune function, and possibly by having a role in apoptosis.

2.2.3 Beneficial Effects of α-LA

Heine *et al.* (1995) explained that the tryptophan-large neutral amino acids (Trp-LNAA) ratio in the plasma is essential for the Trp availability and thus for the serotonin synthesis in the brain. Human milk protein provides optimal conditions for the availability of the neurotransmitter serotonin due to its high Trp concentration. A low protein cow milk-based infant formula supplemented with α-LA containing 5.8 per cent Trp resembles human milk to a much higher degree.

Hakansson *et al.* (1995) described about a protein fraction of human milk inducing apoptosis in tumor cells but not in mature healthy cells. The main constituent of this protein fraction was α-LA.

Svensson *et al.* (1999) reported that α-LA, together with an unnamed component of human breast milk, killed all types of breast cancer cells tested to date. They suggested that the protein changed in configuration during digestion and resulted in a modified product. This "transformed protein", in the presence of cancer cells and a co-factor, resulted in the death of the cancer cells.

Hakansonn *et al.* (2000) described an α-LA folding variant from human milk with bactericidal activity against antibiotic resistant and susceptible strains of *Streptococcus pneumoniae*. This folding variant of α-LA is a new example of naturally occurring molecules with antimicrobial activity.

Markus *et al.* (2000) observed that α-LA increases the plasma Trp-LNAA ratio, and in vulnerable subjects raises brain serotonin activity, reduces cortisol concentration, and improves mood under stress.

Svensson *et al.* (2000) observed that HAMLET (human α-LA made lethal to tumor cells), is a tumoricidal complex of apo α-LA and oleic acid which kills tumor cells while healthy cells are spared.

Matsumoto *et al.* (2001) indicated towards a new biological function of bovine α-LA, its protective effect against ethanol- and stress- induced gastric mucosal injury in rats. They explained that α-LA has marked anti ulcer activity as an active component of cow's milk protein, and suggested that α-LA intake may serve to protect against gastric mucosal injury, in part through endogenous prostaglandin synthesis.

Markus *et al.* (2002) suggested that dietary protein rich in α-LA improves cognitive performance in stress subjects via increased brain Trp and serotonin activities. Alpha-lactalbumin would increase the plasma Trp-LNAA ratio and improve cognitive performance which often declines under chronic stress.

Orosco *et al.* (2004) reported that ingestion of α-LA induced anxiolytic-like and rewarding effects possibly related to serotonergic activation and suggested that shifting the commonly consumed casein enriched to α-LA enriched diets may induce beneficial effects on mood.

Booij *et al.* (2006) showed that α-LA improved abstract visual memory and impaired simple motor performance. They suggested that supplements of α-LA may be useful for nutrition research in relation to age- or disease- related memory decline and the long-term effects of α-LA should also be investigated.

2.2.3.1. Immunomodulator effect

Wong *et al.* (1997) reported that α-LA has physiological properties of whey proteins including immunoenhancing effects. This protein has an enhancing effect on the synthesis of the 1β interleukin (IL-1β) which stimulates immunological response at many levels, so have an important effect on cellular immune functions in ruminants.

Kaczmarczyk *et al.* (2005) studied the relation between the polymorphism of the α-LA gene and populations and subpopulations of peripheral blood lymphocytes in young heifers naturally infected or not infected with bovine leukaemia virus. Their result suggested an incomplete expression of the immunological function of α-LA gene product in heifers in their first month of life.

2.2.3.2. An ingredient in infant formula

Infant formula is a modern artificial substitute for human breast milk. Formulae are designed for infant consumption and are usually based on either cow milk or soy milk. The ideal "humanization" of milk substitutes should include the creation of an amino acid pattern closely resembling that of human milk. The differences in the amino acid compositions of bovine and human milks are largely attributable to differences in their α-LA contents.

Heine *et al.* (1991) reported that the amino acid composition of bovine α-LA and its 72 per cent sequence identity to human α-LA makes it an ideal protein for the nutrition of human infants.

Boston *et al.* (2001) studied on transgenic mice that expressed bovine α-LA and suggested that α-LA is a limiting component in Lactose synthase complex and augmenting α-LA expression in the dam increased the growth of suckling offspring.

According to Matsumoto *et al.* (2001), α-LA is the most abundant component in human whey protein and has strong similarity in its amino acid sequence to the bovine type. It is currently known that bovine α-LA has low immunogenicity for human beings and is an ideal nutrient for infants and young children because of its high essential amino acid content especially of tryptophan. The high concentration of α-LA in human milk fulfils a biological role in the gastrointestinal protection of newborn infants.

Human and bovine milk differ substantially in the ratio of whey proteins to casein being 60:40 in human and 20:80 in bovine milks, respectively and in the proportions of specific proteins. The concentration of α-LA is relatively low in current infant formulae although they closely mimic the ratio of total whey proteins to casein in human milk (Lien, 2003).

Alpha-lactalbumin has an amino acid composition that contributes substantially to meeting the essential amino acid requirements of newborn infants and so important in infant nutrition. The intact protein and various forms of α-LA resulting from partial digestion, such as multimers and fragments, have various biologic activities (Lonnerdal and Lien, 2003).

2.3 SEQUENCE COMPARISON OF α-LA

Mercier *et al.* (1978) deduced a 19-amino acid leader sequence in bovines based on nucleotide sequence and found that it is identical to that reported for ovine α-LA except at amino acid position 5 where *His* is substituted by *Trp* in the ovine sequence.

Comparison of the complete human α-LA gene and its 5′ and 3′ flanking sequences with the rat α-LA gene shows identical structural organization and identifies extensive homology within the 5′ flanking regions of the two genes. The comparison of the position of exon-intron boundaries in the human and rat α-LA genes reveals that all three introns occur at identical positions in the two genes (Qasba and Safaya, 1984).

Goat α-LA, which shares 92 per cent homology with its bovine counterpart contains two short additional nucleotide stretches 305-328 and 641-648 upstream from the transcription unit. Interestingly, the latter being a 'milk-box' consensus sequence common to α-LA and Ca-sensitive casein-encoding genes (Hall *et al.*, 1987).

Hurley and Schuler (1987) identified a cDNA clone for bovine α-LA, using a rat cDNA probe. The bovine α-LA mRNA sequence has 74 per cent similarity with that of rat and guinea pig and 79 per cent similarity with human α-LA mRNA sequence. The amino acid sequence of bovine α-LA showed a similarity of 76 per cent to the rat, 85 per cent to the human and 78 per cent to the guinea pig α-LA.

Hall *et al.* (1987) determined and compared the complete nucleotide sequence of the human α-LA gene and its immediate flanking sequences with those of the rat α-LA gene. A high degree of homology (67 per cent) was observed in the 5′ flanking regions, extending as far as 655 nucleotide residues upstream of the

transcriptional initiation site. Another significant finding was a run of about 30 nucleotide residues highly conserved in casein and α-LA genes of rat and human both in terms of sequence and position.

Petterson *et al.* (2006) explained that human α-LA has high sequence homology with the bovine, equine, porcine, and caprine proteins. The caprine and bovine proteins are the most closely related, with 95 per cent sequence homology with amino acid differences in the α-helical region at position 10, 11, 17, 30, 70 and 90. The porcine protein lacks one amino acid at position 68 and has seven unique amino acid residues throughout the sequence at positions 20, 23, 33, 39, 44, 67, 89 and 123.

2.4 VECHUR CATTLE

Girija (1995) reported the diploid chromosome number in Vechur cattle as 60. The acrocentric nature of the Y-chromosome establishes that the Vechur cattle belong to the species of the Zebu cattle (*Bos indicus*) as different from the European cattle (*Bos taurus*), which has metacentric Y-chromosome

The percentage of fat and total solids in the milk of Vechur cows is higher compared to crossbred cows. The mean size of fat globule in the milk of the Vechur cow (3.21 μm) is higher than that of the goat (2.60 μm), but considerably smaller than that of the crossbred cows (4.87 μm) and of Murrah buffalo (5.85 μm). The small size of fat globules means high phospholipid content because of greater surface area. Phospholipids are important in the development of brain and nerve tissues and also play a vital role in the absorption and digestion of fat (Venkatachalapathy, 1996).

Vechur cows were considered the pride of Kerala. These animals derived their name from the village of origin, Vechur, a small place by the side of

Vembanattu lake near Vaikom in Kottayam district of Central Kerala. The FAO has listed this breed in their Domestic Animal diversity Information System (Iype and Venkatachalapathy, 1997).

Venkatachalapathy and Iype (1997) reported that Vechur cattle has unique characteristics of its own and have separate identity from other breeds of cattle not only by its small size but also due to its milk protein variants, composition of milk, size of fat globules and level of saturated fatty acids.

Vechur cattle of Kerala are small sized with an average weight of about 130 kg for cows and 170 kg for bulls. These animals need less feed as compared to crossbreds with an yield of 2-3 kg milk per day and they are adapted to hot and humid local environment (Iype and Venkatachalapathy, 2001).

2.4.1 Molecular Studies on Vechur cattle

Aravindakshan *et al.* (2003) analysed the genetic polymorphism at the third exon of bovine prolactin (PRL) gene in Vechur cattle by PCR based RFLP. They observed that the different PRL genotypes did not vary significantly ($p>0.05$) for different traits such as milk fat percentage, peak milk yield, first lactation yield and lactation length.

Aravindakshan *et al.* (2004) carried out investigations on the genetic polymorphism at the β-lactoglobulin locus in Vechur cattle using PCR-RFLP. The analysis revealed two alleles, namely allele A and allele B and the frequency of the desirable B allele was very high.

Suprabha *et al.* (2005) used the band sharing (BS) values in RAPD-PCR analysis of dwarf cattle of Kerala of each group and between groups as an analytical tool to find out genetic variation among animals. The Vechur, Vatakara and

Vechur-Vatakara combinations were found to be more uniform with high BS value compared to other combinations. Band sharing values of Kasargode-Highrange dwarf animals were significantly different from Vechur, Vatakara and their combinations.

The Vechur cattle were screened for the presence of the bovine leukocyte adhesion deficiency (BLAD) genetic defect using the PCR-RFLP and DNA sequencing methods and found that none of the 129 Vechur animals under study was found to carry the BLAD allele (Aravindakshan *et al.* 2006).

Aravindakshan and James (personal communication, 10 July 2007) investigated the DNA sequence variations at the *kappa*-casein (κ-CN) gene locus carrying nucleotide substitutions of A or B alleles in Vechur cattle by the PCR-RFLP and DNA sequencing methods. They identified that the two alleles differed at four additional positions which may facilitate more accurate typing and offering more flexibility in the selection of enzymes for genotyping κ-CN locus.

Aravindakshan *et al.* (personal communication, 10 July 2007) reported the gene and genotype frequencies of GH/*Msp*I polymorphism in Vechur cattle. A 768 bp fragment enclosing the polymorphic *Msp*I site within the third intron of growth hormone gene was amplified by specific primers. The digestion of the PCR product with *Msp*I enzyme revealed two alleles namely *Msp*I (+) and *Msp*I (-) and among these *Msp*(-) allele was found to be more frequent as in the case of other Indian cattle.

2.5 MOLECULAR CLONING

Genetic engineering is a new and powerful technology which allows scientists to 'reprogram' animals, altering their genetic make-up or even mixing genetic material between different species. From a handful of experiments in the

early 1980s, animal genetic engineering has mushroomed into a major area of scientific and commercial interest, leading to a rapid increase in animal use.

Transgenic technologies combined with advances in mammary gland biology led to the development of transgenic mice, which secreted a pharmaceutical protein in their milk. The application of these technologies in cattle has been restricted largely to produce potential pharmaceutical or nutriceutical products in the mammary gland (Gordon *et al.*, 1980).

Alexander *et al.* (1988) isolated the bovine κ-casein gene as a series of overlapping lambda phage clones and showed that it consist of five exons distributed over a total length of approximately 13 kb. They reported that most of the mature protein-coding sequence was contained in a single large exon.

Maga and Murray (1995) explained the possibility of transgenic technology to alter the functional and physical properties of milk resulting in novel manufacturing properties. The properties of milk have been altered by adding a new protein with the aim of improving the milk, not of recovering the protein for other uses.

Dong and Kurland (1996) stated that generally experiments for the overproduction of recombinant proteins begin by the cloning of a gene for the desired protein into a multicopy plasmid under the control of a strong inducible promoter.

Wall *et al.* (1997) suggested that fundamental knowledge generated from transgenic animal models enabled scientists to produce transgenic livestock with goals of improving animal production efficiency and generating new products. The ability to modify mammary-specific genes provides an opportunity to pursue several distinctly different avenues of research. The objective of the emerging gene

"pharming" industry is to produce pharmaceuticals for treating human diseases. It is argued that mammary glands are an ideal site for producing complex bioactive proteins that can be cost effectively harvested and purified.

Laird *et al.* (1998) reported that transfection of the guinea-pig α-LA gene cloned in a bovine papilloma virus vector into the mouse C127 and human MCF-7 mammary tumour cell-lines gave rise to stable but seemingly constitutive expression of α-LA. The expression was from the correct transcriptional start point, resulting in the accumulation of correctly processed mRNA and the secretion of α-LA into the culture medium.

According to Zuelke (1998) transgenic techniques were quickly extended into sheep, goats, pigs and cattle, and the prospect of producing pharmaceutical proteins in the mammary gland of livestock species became a reality.

Boston *et al.* (2001) demonstrated that an increase in α-LA expression in transgenic mice may increase pup growth and milk production marginally. They suggested that use of transgenic techniques to increase α-LA expression may provide a mechanism by which milk production maybe enhanced.

Bhure and Sharma (2007) characterized the regulatory elements in ovine αS1-casein gene promoter. They cloned and sequenced region extending from - 2136 to + 49 bp containing 5'-flanking region and exon1. Computational analysis of the sequence showed presence of core promoter elements namely TATA box, CAAT box and initiator sequence.

2.5.1 Isolation of Genomic DNA

The genomic DNA serves as a unique, stable, heritable representation of an organism's genetic makeup. Methodologies that enable rapid isolation and analysis

of DNA have become the fundamental to a variety of research techniques, especially polymerase chain reaction (PCR) based methods to exponentially increase the sensitivity of analysis. So the isolation of pure, high molecular weight genomic DNA is the first and foremost requirement to carryout genetic analysis such as characterization of gene, DNA polymorphisms and linkage studies.

Brawerman *et al.* (1972) pointed out that although phenol denatures protein efficiently, it does not completely inhibit RNase activity, and it is a solvent for RNA molecules that contain long tracts of poly (A). Both of these problems can be circumvented by using a mixture of phenol:chloroform:isoamyl alcohol (25:24:1). The subsequent extraction with chloroform removes any lingering traces of phenol from the nucleic acid preparation.

Sambrook *et al.* (1989) described that the standard way to remove proteins from nucleic acid solution is to extract first with phenol chloroform and then with chloroform. This procedure takes advantage of the fact that deproteinization is more efficient when two different organic solvents are used instead of one.

Aravindakshan *et al.* (1998) compared three methods of DNA extraction namely, the guanidine hydrochloride method, the high salt method and the phenol-chloroform method and demonstrated that both high salt method and phenol chloroform method produced good yields of high molecular weight DNA from cattle white blood cells whereas the guanidine hydrochloride method failed to yield clean DNA.

Arranz *et al.* (2001) used Proteinase–K digestion followed by salting out for obtaining genomic DNA from blood and semen.

Barker *et al.* (2001) described the different steps involved in DNA extraction from white blood cells of goats using phenol-chloroform extraction procedure.

2.5.2 Yield and Purity of DNA

Apparao *et al.* (1994) extracted genomic DNA from cattle, buffalo, sheep, goat and swine using a modified phenol-chloroform extraction method and obtained yields of 250 to 300 µg DNA from 15 ml of whole blood.

Ahmad *et al.* (1995) reported a modification of the widely used standard proteinase K/phenol DNA isolation method for improving the yield and purity of DNA from frozen blood samples, by an initial trypsinisation of whole blood before cell lysis to obtain lymphocytic nuclei and subsequent DNA purification. This modification to improve yield and purity of DNA from frozen blood samples should be useful to laboratories performing DNA based diagnostic work or studying molecular genetic mechanisms of disease.

Aravindakshan *et al.* (1998) reported that the mean yields of DNA extracted from 10 ml of whole blood of cattle and buffalo by phenol and high salt methods were 394.50/446.16 µg and 344.25/432.83 µg, respectively. The ratio of optical densities at 260 and 280 nm was consistent between 1.75 and 1.91 indicating good deproteinisation.

2.5.3 Polymerase Chain Reaction (PCR)

The polymerase chain reaction (PCR), which involves the *in vitro* enzymatic synthesis of millions of copies of specific DNA segments, has transformed the way DNA analysis is carried out in molecular studies. The ability to reproduce a target section of a DNA sequence through the use of the PCR has facilitated a wide array of amplification techniques.

The PCR was developed by Mullis *et al.* (1986). This technique involves enzymatic amplification of a specific sequence of DNA using two oligonucleotide

primers that flank DNA segment to be amplified resulting in several fold increase in the amount of target sequence.

According to Saiki *et al.* (1988), a thermostable DNA polymerase isolated from *Thermus aquaticus* (*Taq*) enable *in vitro* DNA amplification to be performed at higher temperatures and significantly improves the specificity, yield, sensitivity and the length of the product amplified.

Innis and Gelfand (1990) pointed out that "time at temperature" is the main reason for denaturation/loss of activity of *Taq* DNA polymerase. They suggested that reducing the denaturation time to 30 s for short template sequences and increasing the denaturation temperature up to 96 °C and decreasing time may prolong the activity of the enzyme.

Yap and McGee (1991) suggested the possibility of reducing the denaturation temperature after about 10 rounds of amplification, as the mean length of target DNA is decreased. For templates of 300 bp or less, denaturation temperature may be reduced to as low as 88 °C for templates with 50 per cent (G+C) content, which means one may do as many as 40 cycles without much decrease in enzyme efficiency.

White (2005) explained that optimization of the PCR includes the consideration of buffer, additives, magnesium chloride, deoxyribonucleotides, primer design, thermostable DNA polymerase, DNA template, reaction volume, denaturation time, annealing temperature and extension time. Nonspecific products can be further minimized by using a 'hot start' and a variable annealing temperature.

2.5.3.1 Design of Oligonucleotide Primers

The success of the PCR strategy is highly dependent on the small synthetic

oligonucleotides that hybridize to the cDNA sequences. These short nucleotides function in pairs known as the forward and reverse primers, which amplify a specific DNA sequence and, more importantly, anneal exclusively to that DNA target locus (Lexa et al., 2001).

2.5.3.1.1 Melting Temperature, Primer Length and GC Content

Primer length and sequence are of critical importance in designing the parameters of a successful amplification. The melting temperature (Tm) of DNA duplex increases both with its length and with increasing (G+C) content of the primers. The Tm value can be roughly calculated using the following formula.

$$\mathbf{Tm = 4(G + C) + 2(A + T)\ ^{o}C}$$

Thus, the annealing temperature chosen for a PCR depends directly on the length and composition of the primer(s). One should aim at using an annealing temperature (Ta) about 5 °C below the lowest Tm of the pair of primers to be used (Innis and Gelfand, 1990).

A more rigorous treatment of Ta is given by Rychlik *et al.* (1990). They suggested that if the Ta is increased by 1 °C every other cycle, specificity of amplification and yield of products of less than 1 kb in length both are increased. One consequence of having too low a Ta is that one or both primers will anneal to sequences other than the true target, as internal single-base mismatches or partial annealing may be tolerated, leading to "non-specific" amplification and consequent reduction in yield of the desired product.

2.5.3.1.2 Primer Specificity

According to Wu *et al.* (1991) Primer specificity is an important parameter in PCR primer design. The primers should bind to the target sequence only but not

somewhere else to ensure the amplification of the intended fragment alone. In other words the target sequence should occur only once in the template. Primer length not only affects the T_m, but also the uniqueness of the sequence in the template.

2.5.3.1.3. Hairpin and Self Dimer

Rychlik (1995) described the importance of avoiding primer complementarity and primer-dimer formation. The hardest part in PCR primer design is to avoid primer complementarity, especially at the 3′ ends, which may result in folding of the primer in half and form a so-called *hairpin* structure, which is stabilized by the complementary base pairing. Primer-dimer formation causes problems to PCR reaction as it may act as a competitor to amplification of the target DNA.

2.5.3.2 General Rules for PCR Primer Design

According to Innis and Gelfand (1990), the rules for primer design are:

1. Primers should be 17-28 bases in length.
2. The (G+C) content of the primers should be 50-60 per cent.
3. The 3′ of the primers should end in a G or C or CG or GC so as to prevent "breathing" of ends and to increase the efficiency of priming.
4. T_m values between 55-80 °C are preferred.
5. Avoid primers with 3′ complementarity, as otherwise primer dimers will be synthesized preferentially to any other product.
6. Primer self-complementarity (ability to form secondary structures such as hairpins) should be avoided.
7. Runs of three or more Cs or Gs at the 3′-ends of primers may promote mispriming at G or C-rich sequences because of stability of annealing and should be avoided.

Lexa *et al.* (2001) described Virtual PCR as a promising programme, which accepts user-given primers and then conducts a similarity search using NCBI Blast to identify sequences in public databases that are complementary to any two primers. Input is limited to only two primers at a time, making batch analysis a difficult job.

2.5.4 Cloning of PCR products

A successful cloning of PCR products is a key step for further analysis of DNA fragments, though it is a difficult task.

Icho *et al.* (1985) reported the construction of plasmids which carry *cds gene* of *E. coli* which codes for the enzyme CDP-diglyceride synthetase. Using these plasmids they have sequenced 1274 bp of DNA, including a 750-bp open reading frame which is the coding region of the *cds gene*. They suggested that gene cloning is an essential prerequisite for many regulatory and structural studies.

Cloning methods are separated into two groups, namely, ligation-dependent cloning and ligation-independent cloning. The ligation-dependent cloning methods for PCR products can be further classified into three types based on the ends of the DNA as blunt-end cloning, sticky end cloning and T-A cloning (Clark, 1988).

Abed *et al.* (1995) reported that direct sequencing of the products of PCR presents difficulties and often requires special manipulations such as the generation of excess single-stranded DNA using asymmetric PCR. The alternative methods involve cloning PCR products into a vector DNA suitable for sequence analysis.

2.5.4.1 T-A Cloning

The T-A cloning was designed to clone PCR products produced by *Taq* DNA polymerase. It takes advantage of the terminal transferase activity of this polymerase which adds a single 3'-dA overhang to each end of the PCR product. A PCR product

with 3′-dA overhangs can be directly cloned into a linearized plasmid vector with complementary single 5′-deoxythymidine (dT) overhangs at both ends without modification by restriction enzymes or any other enzymes. The dT overhang containing vector is called a T-vector and the corresponding cloning is called T-A cloning. T-vector based methods have emerged as a popular means for cloning PCR products.

Taq DNA polymerase is the most widely used thermostable DNA polymerase and it has terminal transferase activity. In the presence of four deoxyribonucleotide triphosphates (dNTPs), deoxyadenosine (dA) is preferentially added to the 3′ termini of PCR-amplified duplex DNA molecules leaving a single 3′-dA overhang (Clark, 1988).

Marchuk *et al.* (1990) explained construction of T-vectors as a rapid and general system for direct cloning of unmodified PCR products. PCR products are gel purified to avoid cloning spurious bands, and ligated to the vector at 14 °C. The vector and PCR products have complementary single base 3′ overhangs. Vector self-ligation events are prohibited by the 3′-dT overhang and concatamerization of the insert is prohibited by the unphosphorylated 5′ end contributed by the oligonucleotide primer as well as the 3′-dA overhang added by *Taq* DNA polymerase during the PCR reaction.

Many DNA polymerases, including *Taq* DNA polymerase, are capable of adding an additional non-template directed nucleotide to the 3′-ends of a blunt-ended DNA fragment. *Taq* polymerase preferentially adds a single 3′-dA to blunt-ended double stranded DNA through this terminal transferase-like activity (Mead *et al.*, 1991).

Zhou *et al.* (1995) explained that TA cloning is one of the simplest and most efficient methods for the cloning of PCR products. The procedure exploits the

terminal transferase activity of certain thermophilic DNA polymerases, including *Taq* DNA polymerase.

Novy *et al.* (1996) suggested that performing excessive PCR cycles in the absence of sufficient *Taq* DNA polymerase can lead to conditions in which ragged ends are produced due to incomplete elongation of the amplicons. Therefore the number of PCR cycles performed and the amount of active *Taq* DNA polymerase used to produce the insert affect the T-A cloning efficiency.

Guo and Bi (2002) described T-A cloning as simple, reliable and more efficient than blunt-end cloning. They observed that as long as DNA polymerases without 3'-5' exonuclease activity are used, T-A cloning is applicable and the PCR products may need no post-PCR processing steps.

Lee *et al.* (2007) tested the ability of *Taq* DNA polymerase to amplify long DNA fragments and showed that under proper conditions *Taq* could successfully perform long PCR up to 24 kb. They proposed that the most important requirements are the survival rate of *Taq* polymerase at high temperatures and that of the primers against the 5' to 3' exonuclease activity of *Taq* polymerase.

Liu *et al.* (2007) studied the relationship between variable T-A cloning efficiency and the different 5' end nucleotide base of primers used in PCR amplification. They obtained different cloning efficiency with different primer pairs containing dA, dT, dC and dG at the 5' terminus, respectively. The data showed that when the 5' end base of primer pair was dA more white colonies could be obtained in cloning the corresponding PCR product in comparison with other bases. This presumably is a consequence of variability in 3'-dA addition to PCR products mediated by *Taq* polymerase.

2.5.5 *E. coli* as a Host Organism

In early 1970's Cohen *et al.* (1973) successfully cloned a DNA fragment for the first time into a plasmid, which was propagated in *E. coli* cells using a calcium chloride method. Since then DNA cloning has developed to a key method in molecular biology and *E. coli* is used in every molecular genetics laboratory for many methods including DNA cloning, protein over expression and purification.

Oliner *et al.* (1993) described an efficient method to clone PCR products exploiting endogenous *E. coli* enzymatic activities. PCR products are engineered to contain terminal sequences identical to sequences at the two ends of a linearized vector. PCR products and vector DNA are then simply co-transfected into *E. coli* strain JC8679, obviating the requirement for enzymatic treatment of the PCR product or *in vitro* ligation. The high rate of homologous recombination in this strain results in efficient incorporation of the insert into the vector, a process refer to as *in vivo* cloning .

E. coli has been the workhorse for the production of recombinant proteins because of fast growth, well-known genetic characteristics and the availability of various tools for gene expression (Lee, 1996).

The two main methods for transformation of competent *E. coli* cells are the calcium chloride method and the electroporation method (Okamoto *et al.*, 1997; Topcu, 2000).

2.5.5.1 Preparation of Competent E. coli for transformation

E. coli cells are more likely to incorporate foreign DNA if their cell walls are altered so that DNA can pass through more easily. Such cells are said to be

competent. Cells are made competent by a process that uses calcium chloride and heat shock.

Morrison (1977) described that all of the methods of rendering cells competent for uptake of DNA involve, briefly : (i) several successive washes in cold NaCl, $MgCl_2$, and/or $CaCl_2$ solutions, (ii) incubation in $CaCl_2$ at 0 °C with DNA, and (iii) a short incubation at some higher temperature before (iv) dilution into growth medium and/or selection of transformants. His data showed that, $CaCl_2$ treated *E. coli* cells can be preserved without loss of viability and with good competence by freezing and storage at -82 °C and the results are reproducible from tube to tube within such preparations for at least three months.

Sambrook *et al.* (1989) explained that the Lac+ bacteria resulting from α-complementation can be easily recognized because they form blue colonies in the presence of the chromogenic substrate 5-bromo-4-chloro-3-indolyl-beta-D-galactoside (X-gal). However insertion of a fragment of foreign DNA into the polycloning site of the plasmid results in production of an amino-terminal fragment that is not capable of α-complementation. Bacteria carrying recombinant plasmids therefore form white colonies.

Tsen *et al.* (2002) reported that even though *E. coli* does not have a natural transformation process, strains of *E. coli* can incorporate extra cellular plasmids into cytoplasm 'naturally' at low frequencies. Their results showed that natural transformation with plasmid exists in *E. coli.*

2.5.5.2 Lysogeny Broth (LB)

Lysogeny broth, a nutritionally rich medium which favours lysogeny is used as a general medium for *E. coli* in molecular genetic studies. It is also known as Luria broth or Luria-Bertani broth.

According to Miller (1987) LB media are nutritionally rich suitable for the growth of pure cultures like recombinant strains. *E. coli* and derived strains are deficient in vitamin B synthesis and they are not able to grow on nutritionally poor media. Tryptone and Yeast extract serve as a source of nitrogen, sulfur and carbon while Yeast extract also contains Vitamin B complex. Sodium chloride provides sodium ions for the membrane transport and maintains osmotic equilibrium of the medium.

2.5.6. DNA Sequencing

Sequencing is the determination of the order of bases in a sample of DNA. Modern DNA sequencing began in 1977, with development of the chemical method of Maxam and Gilbert and the dideoxy method of Sanger, Nicklen and Coulson. Nowadays the dideoxy method is more popular. Various modifications have been developed for this approach and it has been automated for very large-scale sequencing such as the whole genome sequencing projects.

Sanger and Coulson (1975) described a relatively rapid and simple technique than other available methods, called the "plus and minus" method. This method depends on the use of DNA polymerase to transcribe specific regions of the DNA under controlled conditions.

Sanger *et al.* (1977) explained a new method of DNA sequencing which is similar, but more rapid and accurate than the "plus and minus" method. This technique makes use of the 2′, 3′-dideoxy analogues of the normal dideoxynucleoside triphosphates, which act as specific chain-terminating inhibitors of DNA polymerase.

According to Chan (2005) the completion of the human genome sequence has given a stimulus to focus on developing new sequencing methodologies that will

enable "personal genomics", or the routine study of our individual genomes.

2.5.6.1 Basic Local Alignment Search Tool (BLAST)

Basic Local Alignment Search Tool is a search algorithm developed by Altschul *et al.* (1990). They explained BLAST as a sequence similarity search program that can be used via a web interface or as a stand alone tool. Sequence similarity may reflect homology and homology may imply functional similarity. There are several types of BLAST to compare all combinations of nucleotide or protein queries with nucleotide or protein databases.

Tatusova and Madden (1999) described about 'BLAST 2 Sequences' which is a BLAST-based tool for aligning two known homologous protein or nucleotide sequences. This tool utilizes the BLAST algorithm for pair wise DNA-DNA or protein-protein sequence comparison. The entire database search by the standard BLAST program would be unnecessary and time-consuming.

The National Center for Biotechnology Information (NCBI) maintains a BLAST server with a home page at http://www.ncbi.nlm.nih.gov/BLAST/. BLAST is a heuristic that finds short matches between two sequences and attempts to start alignments from these 'hot spots'. In addition to performing alignments, BLAST provides statistical information to help decipher the biological significance of the alignment (McGinnis and Madden, 2004).

Ye *et al.* (2006) made a report on recent enhancements to the results produced by the BLAST server at the NCBI. These include features to highlight mismatches between similar sequences, show where the query was masked for low-complexity sequence and integrate information about the database sequences from the NCBI Entrez system into the BLAST display.

2.5.6.1.2 ClustalW

ClustalW is an alignment program for DNA and proteins with improved sensitivity for the alignment of divergent protein sequences (Chenna *et al.,* 2003).

The sequence alignment method ClustalW produces both sequence alignments and phylogenetic tree or *phylogram* which indicates the relationships between the taxa and also conveys a sense of time or rate of evolution (Knapp, 2003).

2.5.6.1.3 Software for Three Dimensional Structure Prediction of Proteins

Lambert *et al.* opined that (2002) homology or comparative modeling is currently considered as the most accurate method to predict the three-dimensional (3D) structure of proteins. They discussed the automated homology modelling program ESyPred3D. The important steps in modelling are searching databanks to identify the structural homolog, alignment of target–template, model building and optimization, and model evaluation. The most critical step in homology modeling is target–template alignment. They designed a modelling package known as MODELLER, which builds the final 3D structure of the protein.

Arnold *et al.* (2006) explained the features of SWISS-MODEL workspace, which is a web-based integrated service dedicated to protein structure homology modelling. It assists and guides the user in building protein homology models at different levels of complexity. A personal working environment is provided for each user where several modelling projects can be carried out in parallel. Protein sequence and structure databases necessary for modelling are accessible from the workspace and are updated in regular intervals. Tools for template selection, model building and structure quality evaluation can be invoked from within the workspace.

Materials and Methods

3. MATERIALS AND METHODS

3.1 EXPERIMENTAL ANIMALS

The study was carried out on Vechur cattle (*Bos indicus*) maintained at the Vechur cattle conservation unit of Kerala Agricultural University.

3.2 COLLECTION OF BLOOD SAMPLES

Blood samples (5 ml each) were collected, from the jugular vein aseptically using sterile disposable syringes and transferred immediately into sterile centrifuge tubes containing 0.5 *M* EDTA (1 mg/ml) as anticoagulant. After collection of blood, tubes were capped and shaken gently to facilitate thorough mixing of blood with anticoagulant. The blood samples were brought in ice to the laboratory without delay and stored at 4 °C till processed.

3.3 ISOLATION OF DNA FROM WHOLE BLOOD

Genomic DNA was extracted from whole blood using the standard phenol-chloroform extraction procedure (Sambrook *et al.*, 1989) with modifications. The procedure* was as follows:

1. Five ml of blood collected in a 15 ml centrifuge tube was centrifuged at 4000 rpm for 10 min and the plasma was discarded leaving erythrocytes and leucocytes.

2. Two to three volumes of ice-cold RBC lysis buffer was added and mixed well. The entire mixture was incubated in ice with occasional shaking for 10 min for the complete lysis of erythrocytes.

3. The tubes were centrifuged at 4500 rpm for 10 min to pellet the contents. The leucocytes were recovered as pellet and the red coloured supernatant, containing the lysed erythrocytes and haemoglobin, was discarded.

*Composition and methods of preparation of reagents and buffers are provided in Annexure-I.

4. Steps 2 & 3 were repeated, till the pellet was clear without any unlysed erythrocytes.

5. The white cell pellet was washed twice in 10 ml of Tris buffered saline by vigorous vortexing followed by centrifugation at 4000 rpm for 10 min.

6. The washed white cell pellet from 5 ml of blood was completely resuspended by vortexing in 5 ml of saline EDTA buffer so that no clumps remained. To this cell suspension, 25 µl of proteinase-K (20 mg/ml in water) and 0.25 ml of 20 per cent SDS were added and mixed gently and incubated at 50 °C in a water bath with occasional shaking for a period of at least three hours.

7. The digested samples were cooled to room temperature and 300 µl of 5 *M* sodium chloride solution was added and mixed well by vortexing. An equal volume of Tris saturated phenol (pH 7.8) was added and mixed gently by inversion of the tubes for 10 min and centrifuged at 4500 rpm for 15 min.

8. The aqueous phase containing DNA was collected in fresh tubes, to which an equal volume of saturated phenol:chloroform:isoamyl alcohol (25:24:1) was added. The contents were mixed thoroughly by inversion of the tube and centrifuged at 4500 rpm for 15 min.

9. The aqueous phase was collected in fresh tubes, to which an equal volume of chloroform:isoamyl alcohol (24 :1) was added, mixed and centrifuged at 4500 rpm for 15 min and this step was repeated.

10. The upper aqueous phase was carefully transferred to a sterile 50 ml beaker and $1/10^{th}$ volume of 3 *M* sodium acetate (pH 5.5) was added and mixed well.

11. An equal volume of chilled isopropyl alcohol was added and the precipitated DNA was spooled out using a fresh micropipette tip, washed in 70 per cent ethanol and air dried.

12 Dried DNA was resuspended in 0.5 ml of Tris EDTA buffer and stored at -20 °C.

3.4 DETERMINATION OF YIELD AND PURITY OF DNA

Twenty microlitres of the DNA stock solution was diluted to 2 ml with sterile distilled water giving a dilution of 100 times. Optical densities (OD) were measured at 260 nm and 280 nm using a 2 ml cuvette in a UV spectrophotometer (Jenway, UK). Sterile distilled water was used as blank. Yield and purity of DNA samples were estimated as follows.

3.4.1 Yield of DNA Samples

An OD of one at 260 nm wavelength corresponds to approximately 50 µg/ml of double stranded DNA. Concentration of DNA stock solution was calculated as follows.

Concentration of DNA stock solution (µg/ml) = OD_{260} x Dilution factor x 50

3.4.2 Purity of DNA Samples

The purity of DNA samples were assessed by estimating the ratio between the readings at 260 and 280 nm. Pure DNA samples have OD_{260}/OD_{280} ratios of 1.8 and above.

3.4.3 Checking Quality and Quantity of DNA

The quality of DNA samples was first assessed electrophoretically on one per cent agarose gel in 1X TAE buffer in a horizontal submarine gel

electrophoresis unit. The quantity of DNA was estimated by multiplying the concentration and the volume of DNA stock solution.

3.5 PCR AMPLIFICATION of α-LACTLBUMIN (α-LA) GENE

3.5.1 Template DNA Preparation for PCR

Working solutions of DNA samples were prepared from the DNA stock solution by diluting with sterile triple distilled water to get a final concentration of 50 ng/µl and were stored at -20 °C. Two microlitres of this working solution was used for every 50 µl PCR reaction.

3.5.2 Design of Primers

The primers for amplification of the complete α-LA gene in Vechur cattle were designed based on the bovine α-LA gene sequence available in the database. For this, a 3090 bp sequence (GenBank Accession X06366; Vilotte *et al.*, 1987) was down loaded from the National Centre for Biotechnology Information (NCBI) site (*http://www.ncbi.nlm.nih.gov*). This sequence covers the complete α-LA gene of *Bos taurus* cattle. The forward and reverse primers were designed online using the software Primer3 (v. 0.4.0) available free at the site *http://frodo.wi.mit.edu*. The primers designed for this study amplify a 1756 bp product from position 754 to 2509 of the downloaded sequence. The nucleotide sequences of the designed primers were the following:

Forward primer: 5′-GGG GTC ACC AAA ATG ATG TC-3′

Reverse primer: 5′-GCA AAG ACA GCA GGT GTT CA-3′

3.5.3 Synthesis and Dilution of the Primers

The primers were custom synthesized commercially (Integrated DNA Technologies, Inc., USA) and obtained in lyophilized form. They were reconstituted in sterile triple distilled water to make a stock solution of 200 p*M*/µl

concentration. The solutions were incubated at room temperature for one hour and then stored at -20 °C. Working solutions of the primers were prepared by dilution from the stocks.

3.5.4 PCR Conditions

The PCR assay was performed in 200 µl reaction tubes with a total volume of 50 µl, containing 100 ng of DNA template, 1 X PCR buffer, 1.5 m*M* $MgCl_2$, 30 p*M* of each primer, 200 µ*M* of each dNTP and one unit *Taq* DNA polymerase. A negative control, containing all reaction components except template DNA was also set up to check any contamination, if any.

The tubes were spun briefly and placed in the thermal cycler. The cycling conditions for DNA amplification included an initial denaturation at 94 °C and 35 cycles each of denaturation at 94 °C for 1 min, annealing at 57.4 °C for 1 min and extension at 72 °C for 1.5 min followed by final extension at 72 °C for 10 min. The PCR products were kept at -20 °C till further analysis.

3.5.5 Agarose Gel Electrophoresis of Amplified DNA

The PCR products were checked in one per cent agarose gel in 1 X TAE buffer in a horizontal submarine electrophoresis unit. Appropriately sized gel tray was cleaned and sealed. The tray was placed on a leveled surface and comb kept in proper position in the tray. One per cent agarose in TAE buffer was heated in a microwave oven until it was a clear solution, cooled to 60 °C and 0.5 µg/ml of ethidium bromide was added and mixed well. The mixture was cooled to hand bearable temperature and carefully poured into the gel tray avoiding air bubbles. After polymerization, the comb and sealing were removed gently and the tray was immersed in the buffer tank (Amersham Pharmacia Biotech., USA) containing 1X TAE buffer. Two microlitres of the PCR product was mixed with a small drop of 6 X gel loading buffer and loaded into the wells carefully. As a

molecular size standard, 1 kb DNA ladder (Bangalore Genei Pvt. Ltd.) was also mixed with gel loading buffer and loaded into one of the wells.

Electrophoresis was carried out at 2V/cm until the bromophenol blue dye migrated more than 2/3rd of length of the gel. The gel, which was stained with ethidium bromide, was visualized under a UV-trans illuminator (Hoefer, USA) and checked for amplification of target DNA.

3.6 Gel Purification of PCR Product

For gel purification of PCR products, a preparative gel electrophoresis was carried out in the same manner as for checking the PCR product. The adjacent wells of the comb were combined with the help of cello tape so that approximately 50 µl of the PCR products could be loaded into a single well. The DNA band was cut from the agarose gel using a fine scalpel under UV light. The PCR product was purified form the gel slice using the Ultra Clean™ GelSpin™ DNA purification Kit (Catalog # 12400-50, MO BIO Laboratories, Inc. USA) as per the manufacturer's instructions as follows:

1. The gel band weight was determined using a separate tube. The steps were carried out directly in the spin filter as the gel band weighed only 0.186 g.
2. Placed the gel in the spin filter basket and ensured that it was resting on the white filter membrane so that the gel will be completely melted in the spin filter.
3. Added 0.56 ml gel bind buffer (i.e., three volumes of the gel band, 0.186 x 3 ml) to the gel slice, and ensured that the gel was submerged in the gel bind buffer and closed the lid.
4. Incubated for 2 min at 55 °C, inverted once and incubated until the gel was melted.

5. Inverted once to mix and centrifuged the spin filter for 10 s at 10,000 x *g* so that the melted gel and unwanted salts flew through the spin filter membrane as the DNA binds to the membrane.
6. Removed the spin filter and vortexed the collection tube for 5 s to mix the flow through so as to make the salt concentration homogenous.
7. Reloaded all the liquid from the collection tube back on to the spin filter, allowing any unbound DNA bind to the spin filter by passing the DNA through the spin filter second time, thus increasing the yield of DNA from the melted agarose.
8. Centrifuged for 10 s at 10000 x *g* and discarded flow through liquid and replaced spin filter basket, thus allowing the entire available DNA being bound to the spin filter. The discarded liquid containing only melted agarose and salt.
9. Added 300 µl of Gelwash buffer composed of ethanol solution that keeps the DNA bound to the spin filter while washing away residual salt and melted agarose.
10. Centrifuged for 10 s at 10000 x *g* and discarded flow through and centrifuged again for 30 s at 10000 x *g* to dry out any residual gel wash because the ethanol contained in the buffer can inhibit downstream applications for the recovered DNA.
11. Carefully transferred filter basket to a clean collection tube provided and added 50 µl of elution buffer (10 m*M* Tris) or water directly on to the center of the white spin filter membrane. The elution buffer gets in contact with the entire surface on the spin filter membrane for efficient recovery of the DNA. The DNA is released from the spin filter membrane at this point because there is no salt to keep it bound.
12. Centrifuged for 30 s at 10000 x *g* and discarded the filter basket. The DNA flows through the spin filter membrane and into the collection tube which was stored at -20 °C for further work.

3.7 CLONING α-LA GENE IN T-VECTOR

The gel purified PCR product was cloned using the GeNei™ Instant cloning kit (Catalog # KT63A, Bangalore Genei Pvt. Ltd). The TA cloning is a simple and an efficient method for cloning of PCR products which exploits the terminal transferase activity of certain thermophilic DNA polymerases including the *Taq* DNA polymerase. *Taq* DNA polymerase has non-template dependent activity which preferentially adds a single adenosine to the 3′-ends of a double stranded DNA molecule and thus most of the molecules amplified by *Taq* DNA polymerase possess single 3′-dA overhang. The use of a linearized "T-vector" which has single 3′-dT overhangs on both ends allows direct, high-efficiency cloning of PCR products carrying 3′–dA overhang, facilitated by complementarity between the PCR product 3′-dA overhangs and vector 3′-dT overhangs.

3.7.1 GeNei™ Instant Cloning Kit

The kit includes 2 X INSTANT ligation buffer and high activity INSTANT T4 DNA ligase to enhance single overhang ligation. The ligation reaction can be carried out for one hour at room temperature (25 °C). Ligation can be continued at 4 °C overnight to obtain maximum number of transformants.

The T-vector in the kit is a high copy number vector with pMB1 origin of replication, origin of replication from filamentous f1 phage for the preparation of ss-DNA and the ampicillin resistance marker. The multicloning site (MCS) is within the lacZα fragment that allows blue-white screening. The insert can be released using *Nco*I enzyme or by double digestion. The cloning region is flanked by SP6 and T7 RNA polymerase promoter sequences (Fig.1).

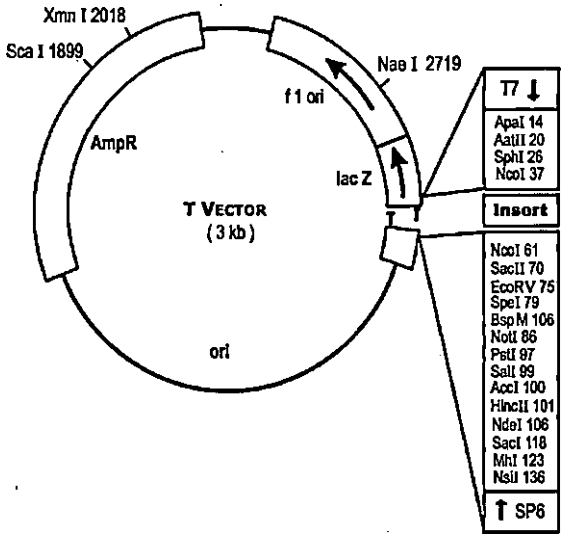


Fig. 1. The T-vector, showing positions of the marker gene, origin of replication, and the MCS.

3.7.1.1 Materials Provided in the Kit

Material	Quantity	Store at
Linearised T-vector (50 ng/µl in 10 m*M* Tris-Cl, 1 m*M* EDTA, pH 8.0)	12 µl (600 ng)	-20 °C
2 X INSTANT ligation buffer	2 x 50 µl	-20 °C
INSTANT T4 DNA Ligase	10 µl	-20 °C
Sterile water (Deionised, autoclaved)	1 ml	-20 °C
Control DNA insert (600 bp PCR product, 20 ng/ µl in 10 m*M* EDTA, pH 8.0)	10 µl (200 ng)	-20 °C

3.7.2 Additional Materials Required for Cloning

1. Competent cells
2. Isopropyl β-D-1-thiogalactopyraranoside (IPTG)

3. 5-Bromo-4-Chloro-3-indolyl-β-D-galactopyranoside (X-gal) (20 mg/ml in Dimethyl formamide)

4. Ampicillin (100 mg/ml)

5. TE buffer (10mM Tris Cl, 1m*M* EDTA, pH 8.0)

6. LB medium and agar

3.7.3 Ligation of α-LA Gene into the T- vector

The T-vector, the control insert and the gel purified PCR products were thawed and centrifuged briefly to collect the contents at the bottom of the vials. For every ligation reaction 50 ng of the vector was used and the remaining vector was stored back at -20 °C. A vector to insert ratio of 1:1 is ideal for efficient ligation and the amount of PCR product required was calculated using the formula:

Size of PCR product (bp) x 50 (amount of vector) / 3000 bp (size of the vector)

For the 1756 bp amplicon, the amount of PCR product required for a 1:1 ratio was calculated as 29.3 ng. After calculating the amount required, the PCR product was diluted to the required concentration in such a way that the volume of PCR product did not exceed 30 per cent of the total volume.

3.7.4 Ligation Reaction

The ligation reactions were performed in three different tubes, one each for the PCR product, the positive control insert and to check the self ligation of the linearised vector as follows.

Components	PCR product	Control insert	Self Ligation
PCR product	3 µl	-	-
Control insert	-	1.5 µl	-
2 X instant ligation buffer	5 µl	5 µl	5 µl
T- Vector	1 µl	1 µl	1 µl
Instant T4 DNA ligase	1 µl	1 µl	1 µl
Sterile distilled water	-	1.5 µl	3 µl

The reaction components were mixed and spun briefly to collect contents at the bottom. It was incubated at room temperature (25 °C) for one hour and stored at -20 °C till the time of transformation.

3.7.5 Transformation

The T-vector ligated PCR products were transformed to *E. coli* using the Genie™ Transformation Kit (Catalog # KT07, Bangalore Genei Pvt. Ltd) as per the manufacturer's instructions as follows.

3.7.5.1 Materials Provided in the Kit

Materials	Quantity	Store at
Ampicillin	100 mg	4 °C
0.1 *M* $CaCl_2$ (Sterile)	25 ml	4 °C
Host (modified *E. coli*)	1 vial	4 °C
IPTG	1 ml	-20 °C
Plasmid DNA	25 µl	-20 °C
X-gal	1 ml	-20 °C
LB Broth	15 g	RT
Agar	5 g	RT

3.7.5.2 Additional Materials Required

1. Equipment: Spectrophotometer, Refrigerated centrifuge, Incubator shaker and water bath
2. Glassware: Conical flask, Petri plates, Pipettes, Spreader
3. Distilled water, Capped Centrifuge tubes, Crushed ice, Micropipette

3.7.5.3 Preparation of Competent Cells

1. The host cells were revived by adding 0.1 ml of the LB media in to the lyophilized vial and a loopful of suspension was streaked on to LB plates and incubated at 37 °C overnight.

2. Inoculated a single colony into 5 ml of LB medium and incubated in a shaker at 37 °C overnight.

3. Inoculated 1 ml of overnight culture into 100 ml LB medium in a 1 L conical flask and incubated at 37 °C in a shaker for about three hours until the OD A_{600} reaches 0.23-0.26.

4. Chilled the culture flask on ice for 10-20 min and transferred the culture aseptically into sterile centrifuge tubes and spun down at 6000 rpm for 8 min in a refrigerated centrifuge at 4 °C.

5. Discarded the supernatant and to the cell pellet added 15 ml of cold 0.1 *M* $CaCl_2$ solution under aseptic conditions. The cell pellet was resuspended gently in the solution using a pre-chilled pipette, keeping the tubes in ice.

6. The tube was incubated in ice for 30 min and centrifuged at 6000 rpm for 8 min at 4 °C.

7. The supernatant was discarded and the cell pellet was resuspended gently in 0.6 ml of cold 0.1*M* $CaCl_2$ solution.

8. The competent cells were aseptically split in to 100 µl aliquots into six prechilled vials, taking care not to remove the centrifuge tubes from ice.

9. The competent cells prepared were used the transformation experiment immediately, as the efficiency of transformation drops on storage at temperature higher than -70 °C.

3.7.5.4 Transformation Protocol

1. The three different ligation mixtures, containing the PCR product, the control insert and the self ligation control, were added to three aliquots of 100 µl of competent cells, gently tapped and incubated on ice for 20 min.
2. Heat shocked the cells by placing the vials in 42 °C water bath for 2 min and then returned the vials to ice to chill for 5 min.
3. Added 1 ml of LB broth aseptically to the vials and incubated at 37 °C in a shaker for an hour so as to allow the bacteria to recover and express the antibiotic resistance.
4. Labeled three LB-Amp plates with X-gal and IPTG as a, b, c. and pipetted 100 µl of LB broth onto each plate and added 25, 50 and 100 µl of the transformed cells to plates a, b and c, respectively. Mixed well and spread using a spreader.
5. Repeated the step 4 for other two aliquots of competent cells transformed.
6. Another LB-Amp plate was plated with 100 µl of competent cells that has not been transformed, to check for any cell contamination and labeled as control plate.
7. Incubated the plates at 37 °C overnight.

3.7.5.5 Screening of Transformants

The T-vector used in the experiment carries ampicillin resistance

marker that enables only transformed cells to grow on LB-Amp plates. The ampicillin-sensitive non-transformants, do not produce colonies on the selective medium.

3.7.5.6 Screening of Recombinants

The recombinant clones among the transformed cells were identified by the insertional-inactivation of the β-*galactosidase* (*LacZ*) gene in the multicloning site of the T-vector, referred to as the *Blue–White Screening*. The clones carrying non-recombinant plasmids were blue in colour in the presence of X-gal and IPTG whereas the recombinant clones appeared white in colour.

3.8 ISOLATION OF RECOMBINANT PLASMID FROM POSITIVE CLONES

The recombinant plasmid carrying the α-LA gene insert was isolated from the white clones by a modified SDS-alkaline lysis method using the AxyPrep Plasmid Miniprep Kit (Axygen Biosciences, USA) as per the manufacturer's instructions as follows.

3.8.1 Materials Provided in the Kit

Materials	Description
AxyPrep Column	Special DNA binding columns
2 ml microfuge tube	-
1.5 ml microfuge tube	-
RNase A	-
Buffer S1	Resuspension buffer, to which RNase A is added and stored at 4 °C
Buffer S2	Lysis buffer. Storage: RT
Buffer S3	Neutralisation buffer. Storage: RT
Buffer W2 concentrate	Desalting buffer, to which absolute ethanol is added and stored at RT.
Eluent	2.5 mM Tris-Cl. Storage: RT.

3.8.2 Protocol for Plasmid Isolation

1. Inoculated 3.5 ml of LB media with white clones and incubated at 37 °C in a shaker overnight. The culture was centrifuged at 12,000 X *g* for 1 min to pellet the bacteria and decanted the supernatant.

2. The bacterial pellet was resuspended in 250 µl of Buffer S1 by vortexing, ensuring complete resuspension of the bacteria.

3. To the cell suspension, 250 µl of Buffer S2 was added and mixed gently by inverting the tube for 4-6 times.

4. To the above suspension, 350 µl of Buffer S3 was added and mixed gently by inverting 6-8 times followed by centrifugation at 12,000 X *g* for 10 min to clarify the lysate.

5. Placed the AxyPrep column into an uncapped 2 ml microfuge tube and transferred the clarified supernatant from step 4 to the AxyPrep column. It was then spun at 12,000 X g for 1 min.

6 Pipetted 700 µl of Buffer W2 into the AxyPrep column and centrifuged at 12,000 X *g* for 1 min.

7. The filtrate from the microfuge tube was discarded and AxyPrep column was put back into the microfuge tube and the step 6 was repeated.

8. The filtrate was discarded the microfuge tube and the AxyPrep column was put back into the microfuge tube and centrifuged at 12,000 X *g* for 1 min.

9. The AxyPrep column was transferred to a clean 1.5 ml microfuge tube and the purified plasmid DNA was eluted from the column by adding 70 µl of eluent pre-warmed to 60 °C, followed by centrifugation at 12,000 X *g* for 1 min, after keeping at room temperature for 1 min.

10. The filtrate containing pure plasmid DNA was stored at -20 °C till further use.

3.8.3 Agarose Gel Electrophoresis

The recombinant plasmid DNA was checked in one per cent agarose gel using 1 X TAE buffer in a horizontal submarine electrophoresis unit as explained earlier.

3.9 DNA SEQUENCING.

Sequencing was carried out commercially (Bangalore Genei, Pvt. Ltd.) by the dideoxynucleotide sequencing method using an automated DNA sequencer (Applied Biosystems, USA).

Primer walking was the method of choice as the DNA of interest is of 1.7 kb that cannot be covered with a single read. The initial sequencing was performed from each end using T7 and SP6 primers. New sequencing primers were custom designed from the initial reads and sequencing reactions were continued to obtain further sequence of the insert. This process was repeated until the entire length of one strand of the insert was covered. From the T7 end the primers designed for sequencing were 67 BG and 11 BG and from the SP6 end the only one primer designed for primer walking was 12 BG. The primers used for sequencing the insert completely are shown below.

1. T7
2. 67BG - (5'-CAG TGG TCA AAG GGC GTG-3')
3. 11BG - (5'-GTG AAA AAT AGA GAG GTG AC-3')
4. 12 BG – (5'-CAA GAT ACT CTG TGG TCT G-3')
5. SP6

The direction of primer walking and the sequence of the different fragments were in the following order.

BG52ar – T7 -----> BG47c – 67BG -----> BG10b – 11BG -----> BG11b – 12BG -----> BG53a – S P6. The complete sequence of the PCR product was arrived at by combining together the sequences of BG52ar, BG47c and BG10b with the reverse complement of BG11b and BG53a, followed by deletion of the overlapping sequences.

3.10 SEQUENCE ANALYSIS

3.10.1 DNA Sequence Analysis

The 1756 bp sequence obtained through sequencing was analysed by BLAST search at NCBI site for homology using BLASTn programme (http://www.ncbi.nlm.nih.gov/BLAST). The raw sequence was annotated based on the published *Bos taurus* sequence (Accession #X06366; Vilotte *et al.*, 1987). The intron, exon and other structural and functional regions of the sequence were established. For interspecies nucleotide sequence homology between exonic regions of the α-LA gene in Vechur with human and other livestock species, individual exons were compared by BLAST2 programme. The BLASTn programme compares a nucleotide query sequence against a nucleotide sequence database. The 1756 nucleotide sequence of the Vechur α-LA gene was submitted as query sequence and selected the *MegaBLAST* programme for searching highly similar sequences from the DNA database.

3.10.2 Multiple sequence analysis of DNA

The multiple sequence alignment was done by the widely used computer programme EBI tool ClustalW (http://align.genome.jp/). This programme is available from European Bioinformatics Institute (EBI) ftp server. The input sequences were loaded in FASTA format. Using this programme the Vechur sequence was compared simultaneously with sequences of human and important

livestock species to highlight the areas of homology and divergence. The results were presented as a phylogram.

3.10.3 Submission of the Sequence to GenBank

The sequence of α-LA gene of Vechur cattle was submitted to the GenBank using BankIt option (http://www.ncbi.nlm.nih.gov/BankIt/) at the NCBI site and obtained the unique accession number for the new sequence.

3.10.4 Protein Sequence Analysis and Structure Prediction

The exonic regions were compiled together and using the programme ExPASy *Translate tool* (http://www.expasy.ch/tools/dna.html) the amino acid sequence of the signal peptide and the 123 amino acid residues of the polypeptide chain was predicted. The protein sequence was analysed by using the BLASTp programme at the NCBI site. This programme compares an amino acid query sequence against a protein sequence database. The predicted amino acid sequence was entered as query sequence and selected the PROTEIN database for the search.

A multiple sequence alignment of the predicted protein sequence was carried out by ClustalW programme with other mammalian species. A phylogram showing the evolutionary relationship was also constructed from the multiple alignment results.

The secondary structure of the Vechur α-LA was predicted using the PAPIA system (http://mbs.cbrc.jp/papia/papia.html) of Computational Biology Research Center, Japan. The 123 amino acid sequence in one letter code was submitted for the structure prediction.

The three dimensional structure of the sequenced protein was predicted using the automatic homology or comparative modeling servers ESyPred3D (http://www.fundp.ac.be/urbm/bioinfo/esypred/) and the Swiss Model Server

(http://swissmodel.expasy.org/). These servers build three-dimensional models for proteins based on homologues of known structure. The 123 amino acid sequence in one letter code was made as input. The predicted structure was viewed with *RasMol*, which is a molecular visualization software for proteins, DNA and macromolecules.

Results

4. RESULTS

4.1 ISOLATION OF GENOMIC DNA

DNA samples were isolated from whole blood obtained from four Vechur cattle showing typical breed characters.

4.1.1 Yield and Quality of DNA

The concentration and ratio of optical densities at 260 and 280 nm of DNA obtained from the whole blood sample selected for the experiment was 205 µg/ml and 1.78 respectively. The concentration and purity of DNA obtained from 5 ml of whole blood from the four samples are given in table 4.1. The samples, checked for quality by agarose gel electrophoresis, appeared as single bands without sheared fragments indicating good quality of DNA.

4.2 PCR AMPLIFICATION OF α-LACTALBUMIN (α-LA) GENE

The primers for amplification of the α-LA gene were designed based on published bovine sequence and were custom synthesized for PCR amplification of the samples. PCR amplification was performed on all the samples using the primer pair and the products were checked by agarose gel electrophoresis.

4.2.1 Optimization of PCR

The composition of the PCR reactions, temperature and time of the cycles were optimized for the efficient amplification of the PCR product. The reaction components optimized for the PCR assay are presented in Table 4.2. The annealing temperature was optimized by the gradient option in the thermal cycler and 57.4 °C was found to be optimum for the assay. The standardised temperature and cycling parameters are presented in Table 4.3.

Table 4.1 Concentration and purity of DNA obtained from blood samples

Sample No.	OD_{260}	OD_{280}	OD_{260} /OD_{280}	Dilution factor (DF)	Conc.(µg/ml) A_{260} x DF x 50
1	0.037	0.024	1.54	100	185
2	0.040	0.024	1.67	100	200
3	0.038	0.024	1.58	100	190
4	0.041	0.023	1.78	100	205

Table 4.2 Standardized conditions for the PCR assay

Sl. No.	Parameter	Value
1.	Template DNA	50 ng
2.	$MgCl_2$	1.5 m*M*
3.	10X Reaction Buffer µl	5 µl
4.	dNTPs	200 µ*M*
5.	Forward Primer	30 p*M*
6.	Reverse Primer	30 p*M*
7.	Taq DNA Polymerase	1 U
8.	Reaction Volume	50 µl

Table 4.3 Standardized temperature and cycling conditions for the PCR

Sl. No.	Parameter	Temperature/Time
1.	Denaturation	94 °C for 1 min
2.	Annealing	57.4 °C for 1.5 min
3.	Extension	72 °C for 1min
4.	No. of Cycles	35
5.	Final extension (°C/10 min)	72

The $MgCl_2$ concentration was found to be an important parameter for the efficient amplification of the gene. A concentration of 1.5 m*M* $MgCl_2$ in the reaction mixture resulted in robust amplification of the product without any non-specific products (Plate-1).

4.2.2 Agarose Gel Electrophoresis of PCR products

The PCR products were checked for amplification by electrophoresis on one per cent agarose gels using one kb marker as the size standard. The amplified products were found to be approximately 1.7 kb as expected from all the four DNA samples (Plate-2). There was no amplification in the negative control.

4.3 PURIFICATION OF DNA FROM AGAROSE GEL

The PCR products were purified from the gel slice using the Ultra Clean™ GelSpin™ DNA purification Kit (Catalog # 12400-50, MO BIO Laboratories, Inc. USA). The gel purification resulted in more than 50 per cent recovery of the PCR product (Plate-3). The purified PCR products, free of salts, buffers, enzymes, ethidium bromide and dNTPs, were collected in 50 µl elution buffer and stored at -20 °C for the cloning assay.

4.4 TA CLONING

The gel purified PCR product was cloned using the GeNei™ Instant cloning kit. The ligation reactions were performed in three different tubes, one each for the PCR product, the positive control insert and to check the self ligation of the linearised vector. After ligation, the recombinant vectors containing the PCR product and the control insert were checked by electrophoresis. The sizes were approximately 4.7 and 3.5 kb, respectively indicating successful ligation reaction.

4.4.1 Transformation

Competent cells for transformation of cloned vector were prepared by Genei™ Transformation kit (Catalog # KT 07, Bangalore Genei Pvt. Ltd.) as per the manufacturer's protocol. Transformation of the competent cells was carried out with three different ligation mixtures, namely the one with PCR product, the positive control insert and the self ligation reaction. In each case, the transformants were plated on to three LB plates with 25, 50 and 100 µl of transformed cells. In the case of PCR product, although clones were grown in all the three plates, the one with 100 µl showed sufficient number uniformly spread colonies with more than 70 per cent of positive colonies (Plate-4A). The control insert plate with 100 µl showed more than hundred colonies of which more than 80 per cent were white and the remaining being blue colonies (Plate-4B). The number of colonies was less in the other control plates in proportion to the volume of the transformed cells added. The plates with self ligation yielded only a few blue colonies.

4.5 ISOLATION OF PLASMID FROM POSITIVE CLONES

The recombinant plasmid carrying the α-LA gene insert was isolated from the genuine white colonies after propagating in LB-ampicillin broth for 12 hours. The isolated plasmid DNA was checked by agarose gel run to ensure that the recombinant plasmid contain the target gene insert. The plasmid band was found to be having a molecular weight of approximately 4.7 kb ensuring that the recombinant vector carries the DNA of interest (Plate-5). The purified plasmid DNA yield was found to be sufficient for subsequent sequencing steps.

4.6 DNA SEQUENCING

The T-vector cloned amplified product was sequenced by the dideoxy termination method. The sequencing results confirmed the expected 1756 bp

fragment of the α-LA gene, containing an initiation codon 'ATG' for translation. For complete sequencing, primer walk sequencing method was performed using three additional primers in addition to the T7 and SP6 primers. The complete nucleotide sequence of the α-LA gene is presented in Fig. 4.1.

4.7 SEQUENCE ANALYSIS

4.7.1 DNA sequence Analysis

The nucleotide sequence of the α-LA gene of Vechur cattle was compared with corresponding sequences from bovine and other species through BLAST analysis. The sequence was found to be having 99 per cent homology with α-LA gene of *Bos taurus* (Accession # X06366.1), 98 and 95 per cent homology with α-LA gene of yak (Accession #AF194372.1) and sheep (Accession # AB052168.1), respectively (Table 4.4 & Fig 4.2). The annotation of the Vechur α-LA gene sequence based on bovine sequence revealed the presence of four exons of 133, 159, 76 and 58 bp separated by three introns of 321, 473 and 504 bp (Fig. 4.3). The length of exon 1 and exon 3 was found to be 133 and 76 nucleotides in all species compared including human. A comparison of the lengths of exons and introns of the gene in different species is presented in table 4.5. The homology of the exonic regions of Vechur with human and livestock species ranges from 77 to 100 per cent as shown in Fig. 4.4. The results of ClustalW multiple nucleotide sequence alignment showed a very high homology with *Bos taurus*, buffalo and goat with a score of 99, 97 and 95, respectively and with human the score was found to be 63. The complete alignment of the corresponding sequences of Vechur, other livestock species and human is shown in Annexure-IV.

4.7.1.1 GenBank Accession Number

The nucleotide sequence data of Vechur α-LA gene has been submitted to the GenBank DNA database and obtained the unique Accession Number: EU 200932 (Annexure V).

4.7.2 Protein Sequence Analysis

The predicted amino acid sequence revealed the presence of 19 amino acid residues of the signal peptide and 123 amino acid residues of the mature polypeptide chain. The single letter code representing the amino acids and their triplet codons are given in the Annexure -VI. The 19 amino acid signal peptide was found to be 100 per cent homologous with the *Bos taurus* counterpart while seven variants were found with the human counterpart. A comparison of the deduced amino acid sequence of the Vechur α-LA with the amino acid sequence of *Bos taurus* and human α-LA is presented in Fig. 4.5. Similar to *Bos taurus* α-LA the Vechur α-LA also contains 123 amino acid residues with NH_2-terminal glutamic acid and COOH-terminal leucine. The two proteins differ only at position 122 where *Arg* (AGG) was substituted with *Lys* (AAG) in *Bos taurus* while with human protein, variations were found at 31 positions. The protein sequence analysis by BLASTp also revealed 99 per cent identity with the corresponding protein of *Bos taurus* (Accession # P00711) and 98 and 73 per cent identity with that of Yak (Accession # Q9TSR4) and Human (Accession # P00709), respectively (Table 4.6). The ClustalW results of amino acid sequences also produced a high alignment score with *Bos taurus*, buffalo, goat and human with scores of 98, 99, 94 and 72, respectively. The ClustalW protein alignment score for α-LA of Vechur and *Bos taurus* with that of human was found to be equal. The ClustalW multiple protein sequence alignment of Vechur α-LA and that of other mammalian species depicting the homology and divergence is presented in Annexure-VII and the constructed phylogram is shown in the Fig. 4.6.

The distance phylogram of mammals placed Vechur along with bovine species while human and other primate species were in distant group.

4.7.3 Protein Structure Prediction

The secondary structure of the Vechur α-LA was predicted using the PAPIA system (http://mbs.cbrc.jp/papia/papia.html) given in Plate 6. The predicted secondary structure and tertiary structure showed that the larger α-helical lobe is formed by the amino- and carboxyl-terminal sections of the polypeptide chain (residues 1-16 and 86-123) while the smaller lobe, which encompasses a small three stranded antiparallel beta-sheet, and small irregular structure, is formed by the central section of the polypeptide chain (residues 17-85).

The three dimensional structure was predicted using automatic comparative modeling servers Swiss Model Server and ESyPred3D. The 3D model was built using the 3D structure 1F6R chain 'A' as template (Crystal structure of apo- form of bovine α-LA- PDB structure Id=1F6R) which has 99 per cent sequence identity. The predicted 3D protein structures obtained from the above servers and the 3D structure of human α-LA (PDB structure Id=1A4V) from www.pdb.org are presented in Plate 7. It reveals that the α-helices, β-sheets and random coils in the predicted Vechur α-LA are similar to that of human α-LA structures.

The ALIGN program from the EsyPred3D server showed that the templates 1F6R chain A and 1A4V chain A shared 85.9 per cent and 64.8 per cent identity with Vechur sequence, respectively.

Table 4.4 Results of BLASTn of the Vechur α-LA gene nucleotide sequence showing the percentage of identity with other species

Accession	Description	Maximum score	Total score	Query coverage	E value	Maximum identity
X06366.1	*Bos taurus* gene for α-LA	3166	3166	100%	0.0	99%
AF194372.1	Yak α-LA gene, complete CDS	3107	3107	100%	0.0	98%
AF194373.1	Buffalo α-LA gene, complete CDS	2968	2968	100%	0.0	97%
AB052168.1	Sheep α-LA gene, complete CDS	2843	2843	100%	0.0	95%
M63868.1	Goat α-LA gene, exons 1-4	2772	2772	100%	0.0	95%
AY122023.1	*Bos taurus* α-LA gene, intron 2 and partial CDS	883	883	27%	0.0	100%
EF419240.1	Buffalo α-LA (LALBA) gene, intron 3	843	843	28%	0.0	96%
X05153.1	Human α-LA gene	542	542	83%	2e-150	74%

Table 4.5 Comparison of the length of exons and introns of the α-LA gene in different species

Species	No. of nucleotide residues						
	Exon 1	**Intron 1**	**Exon 2**	**Intron 2**	**Exon 3**	**Intron 3**	**Exon 4**
Vechur	133	321	159	473	76	504	330
Bos taurus	133	321	159	473	76	504	330
Buffalo	133	326	159	469	76	505	61
Yak	133	321	159	471	76	503	61
Goat	133	327	159	474	76	503	58
Human	133	648	159	489	76	499	333
Rat	133	341	159	429	76	1016	328

Table 4.6 Protein sequence analyses by BLASTp showing percentage of identity with Vechur α-LA gene

Accession	Protein Name	Length	Organism	Length	Identity (max)	Score (max)	E-Value (min)
P00711	α-LA precursor	142	*Bos taurus*	142	99	302	0
Q9TSR4	α-LA precursor	142	Yak	142	98	300	0
Q9TSN6	α-LA precursor	142	Water buffalo	142	98	299	0
P09462	α-LA precursor	142	Sheep	142	96	294	0
P00712	α-LA precursor	142	Goat	142	94	289	0
Q9N2G9	α-LA precursor	142	Dog	142	74	235	0
P18137	α-LA precursor	141	Pig	141	73	230	0
P00709	α-LA precursor	142	Human	142	73	228	0
P00713	α-LA precursor	142	Guinea pig	142	69	207	0

```
   1 GGGGTCACCAAAATGATGTCCTTTGTTTCTCTGCTCCTGGTAGGCATCCTATTCCATGCC 60
  61 ACCCAGGCTGAACAGTTAACAAAATGTGAGGTGTTCCGGGAGCTGAAAGACTTGAAGGGC 120
 121 TACGGAGGTGTCAGTTTGCCTGAATGTGAGTTCCCTGCTATTTTGCTTTGTCCCATAATT 180
 181 CATCCTCTTCACTCTTTCCCTCCATTCTCTTCATCCTCTTTTTCCCCTCTACTTTTAATT 240
 241 ATCAAACAATTCTCTTATTTGTTTACTCTTTTATTACATTTATTTATCTGCCTCTCCTTT 300
 301 TTCCCATTGTCTGATCCTTTGGAACTCTTTTCACCTTAACAAGATACTCTGTGGTCTGCC 360
 361 ATATTTGGAGATTGGTTGGAGAGCCTTTTTCGGTCTGGGAATACAGGTCCTCATTTATGC 420
 421 TATACATGAACATCCTTGTGAAATCTCTTTTTCATCTTTCTTTCAGGGGTCTGTACCACG 480
 481 TTTCATACCAGTGGTTATGACACACAAGCCATAGTACAAAACAATGACAGCACAGAATAT 540
 541 GGACTCTTCCAGATAAATAATAAAATTTGGTGCAAAGACGACCAGAACCCTCACTCAAGC 600
 601 AACATCTGTAACATCTCCTGTGACAGTGAGTAACTTCTTTTTACTCTGTTCCTGTGTTTT 660
 661 TCTGAAACCTACTCCTGGGATAACCTCCTTTTTTTTGGTGTGAAGCACACCTCTGGCTTC 720
 721 ACTGCCTTGGACTCCAAATTAACTGTGGGACTTGATAATACCGAGTAAGAGGCTCTTAGA 780
 781 ATTTTTCATTAACACTAAATCCCCAGACAGTTTCTTAAAGTTCCTGGGTAGGTGACCTGA 840
 841 GCTGTTTGGGGATCTTGATGTATAATACCCTGTATTTTCAGACTAAGTTGGTTGATGAAG 900
 901 TTGATAATTCCTAAGGAGCTGCCCCAGAGAAGAGAAGGGAGTCCTTACCTAGGGATAGGC 960
 961 ATTACTGTATTAAATTTCTCACCCAGAAGGCAACAGGCATAAGCCTCTAGTTCAGAGAAA 1020
1021 ACCAGAGAAGAGGGAAATTCATTATCCTTCTGGGTAATACTTAGCTCTCTCATTTTTTCC 1080
1081 ACCTATAACTCCTGCCAGAGTTCCTGGATGATGATCTTACTGATGACATTATGTGTGTCA 1140
1141 AGAAGATTCTGGATAAAGTAGGAATTAACTACTGGTGAGTCACCTCTCTATTTTTCACTT 1200
1201 AATCTTTCCTCTCTTTCTTCTCAGTCCTTTCGTCCCAGCACTATACTCCTTTCTCTCTAT 1260
1261 TTCCTGGTCTTTTAAGCTAGAATGTAATCTTAAAAACAAAAATCATCAAGCAGACTCCGG 1320
1321 TTTCCAATTTTGAAGCTTCACTTACTTCACTCCCGTTAGCAATTTTCCTACCTAAGGTTC 1380
1381 CCTAATAGAGGGCTGAGATCCAGGATTTCCTTCACCAGGACTTGAACATCTAATTCTACT 1440
1441 TGTTCAGTCCTACATCCTAAGGCACGCCCTTTGACCACTGCCCCGCAACTTTCTTGGAGT 1500
1501 TTTAAAAAATGGACCTTACTCCACTAAGTGGCTCAGTGTCTCTAGCCATGTGGCTAGGAA 1560
1561 AGTCTGTCTGTAATTTTAACCCACAGTCTTCGACCTCAGCCTTCCTGGGGATAAAGCTAG 1620
1621 ATGTAAATCTAACCAAGATCCTGTCAGTAATTTGCCTTGTCTCCTTCTTCATGATCAGGT 1680
1681 TGGCCCATAAAGCACTCTGTTCTGAGAAGCTGGATCAGTGGCTCTGTGAGAGGTTGTGAA 1740
1741 CACCTGCTGTCTTTGC                                              1756
```

Fig. 4.1 The coding strand of 1756 nucleotides long fragment of Vechur α-LA gene

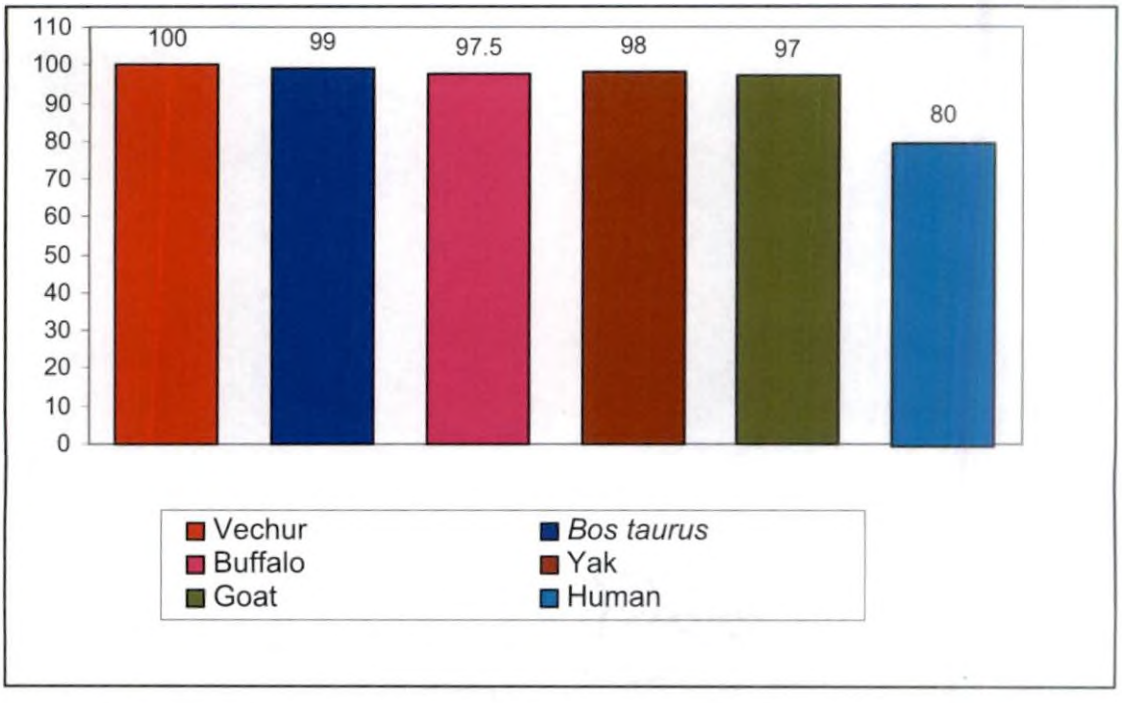


Fig. 4.2 Interspecies nucleotide sequence homology of Vechur α-LA gene with Human and livestock species.

CDS	join (13..145, 467..625, 1099..1174, 1679..1736)
	/codon_start=1
	/product="alpha-lactalbumin"
sig_peptide	13..69
	/note="signal peptide (AA -19 to -1)"
exon	13..145
	/number=1
misc_feature	17..22
	/note="pot. glucocorticoid site"
intron	146..466
	/number=1
misc_feature	146..152
	/note="5' splicing sequence"
misc_feature	438..448
	/note="put. internal splicing sequence"
misc_feature	464..466
	/note="3' splicing sequence"
exon	467..625
	/number=2
intron	626..1098
	/number=2
misc_feature	626..631
	/note="5' splicing sequence"
misc_feature	1060..1070
	/note="put. internal splicing sequence"
misc_feature	1096..1098
	/note="3' splicing sequence"
exon	1099..1174
	/number=3
intron	1175..1678
	/number=3
misc_feature	1175..1180
	/note="5' splicing site"
misc_feature	1635..1644
	/note="put. internal splicing site"
misc_feature	1676..1678
	/note="3' splicing sequence"
exon	1679..1736
	/number=4
misc_feature	1698..1703
	/note="pot. glucocorticoid site"

Fig. 4.3 Features of α-lactalbumin gene of Vechur Cattle

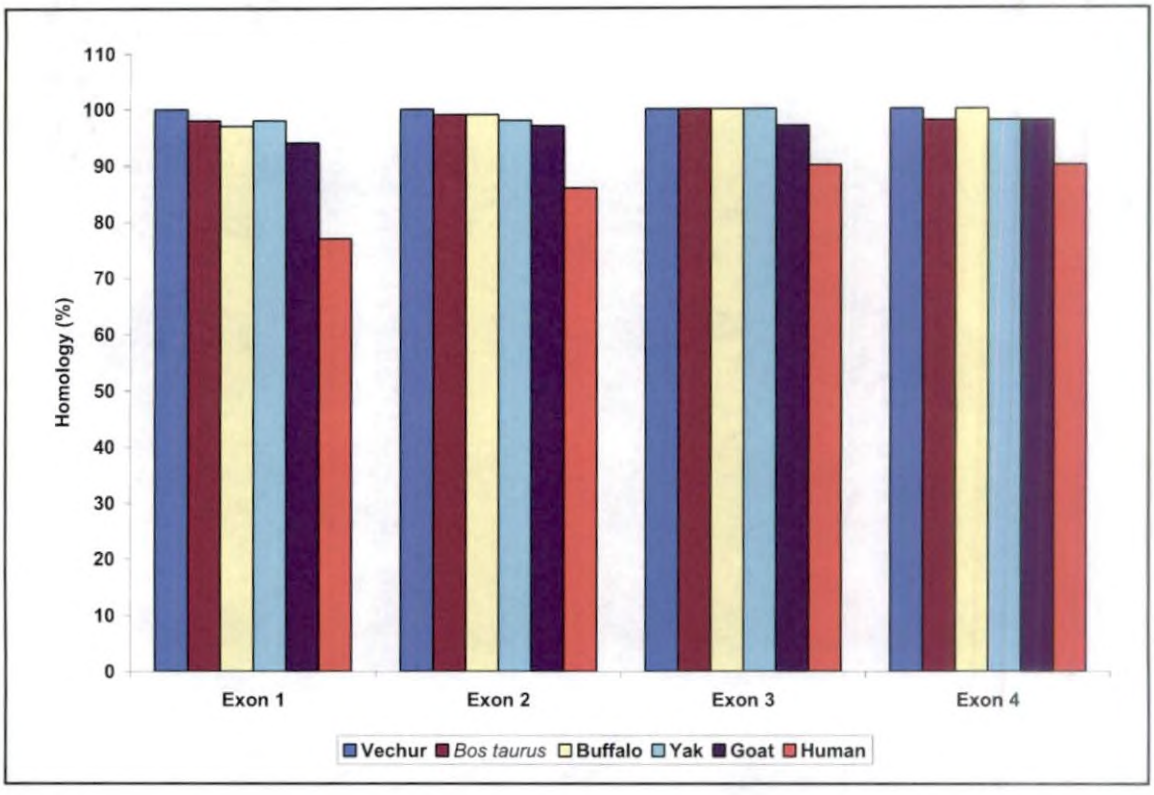


Fig. 4.4 Interspecies nucleotide sequence homology of the exonic regions of α-LA gene based on Vechur sequence as a reference

Fig. 4.5 The deduced amino acid sequence of Vechur α-LA and the amino acid sequence of *Bos taurus* and human α-LA. The variants amino acids are shown in bold red letters.

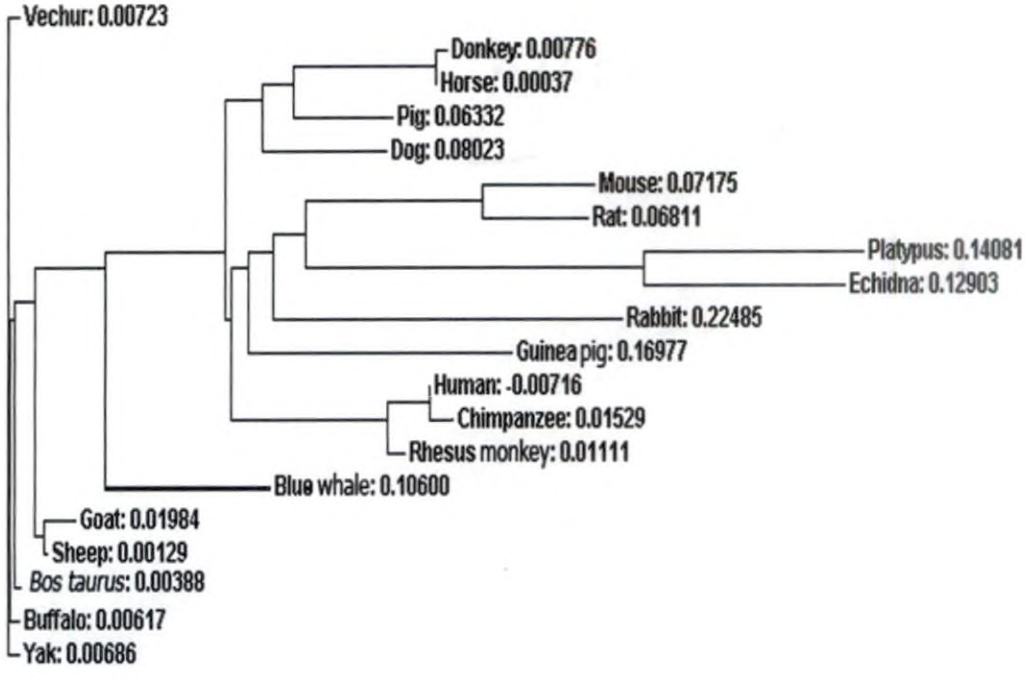


Fig. 4.6. Phylogram constructed by ClustalW depicting the relationships between species and rate of evolution

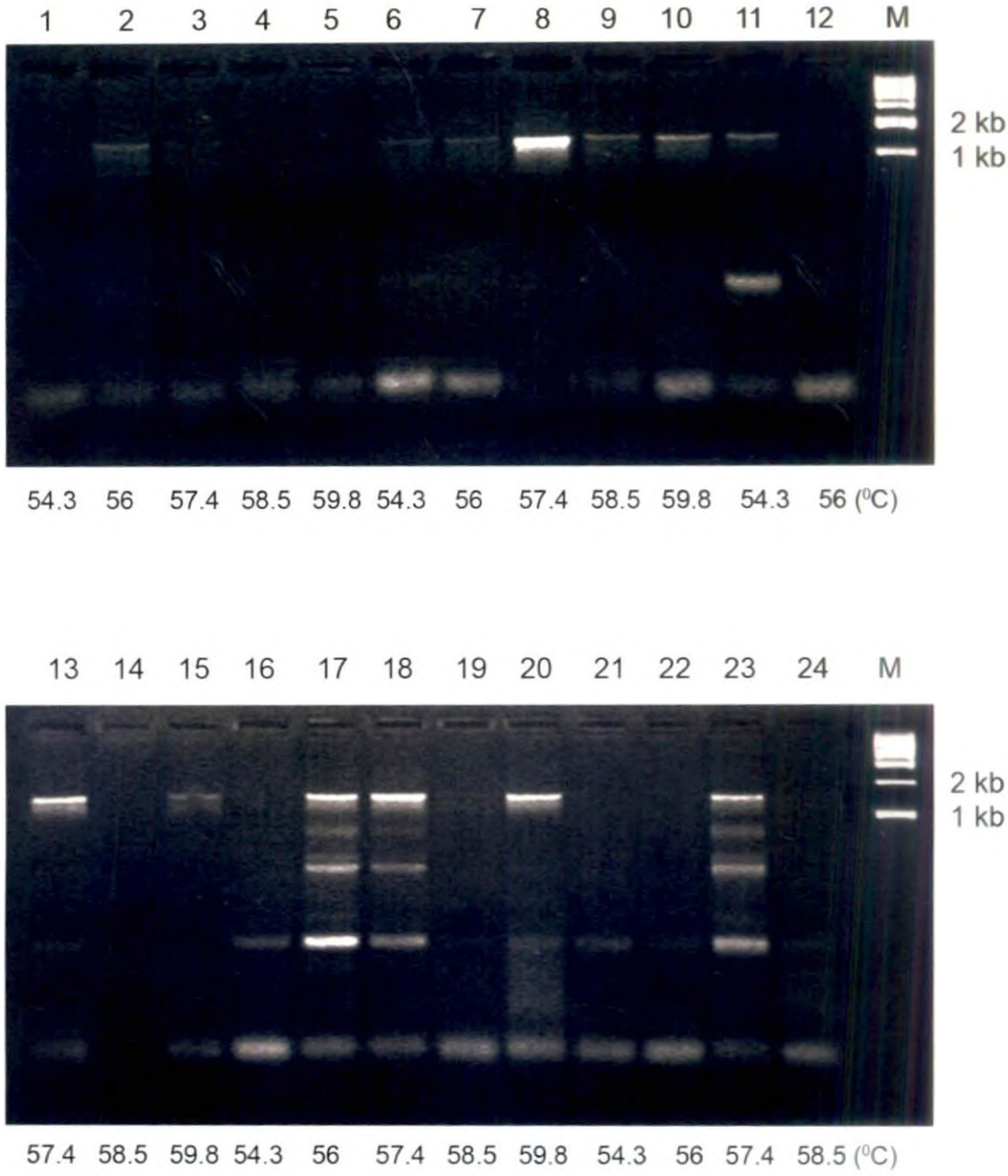


Plate 1. Optimization of $MgCl_2$ concentration and annealing temperature of PCR

Lanes 1-5: 1 m*M*, 6-10: 1.5m*M*, 11-15: 2m*M*, 16-20: 2.5m*M* and 21-24: 3m*M*. Lane M: 1kb DNA Ladder. The anneling temperature is shown in the respective lanes

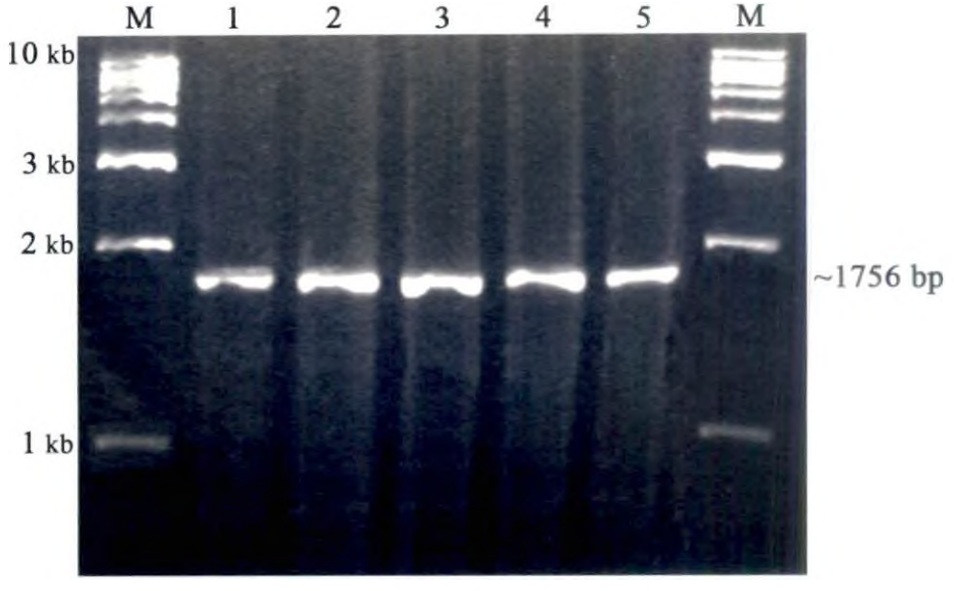


Plate 2. Gel showing ~1756 bp amplified fragment of α-LA gene
Lane M. 1kb DNA ladder, Lane 1 to 5. PCR products

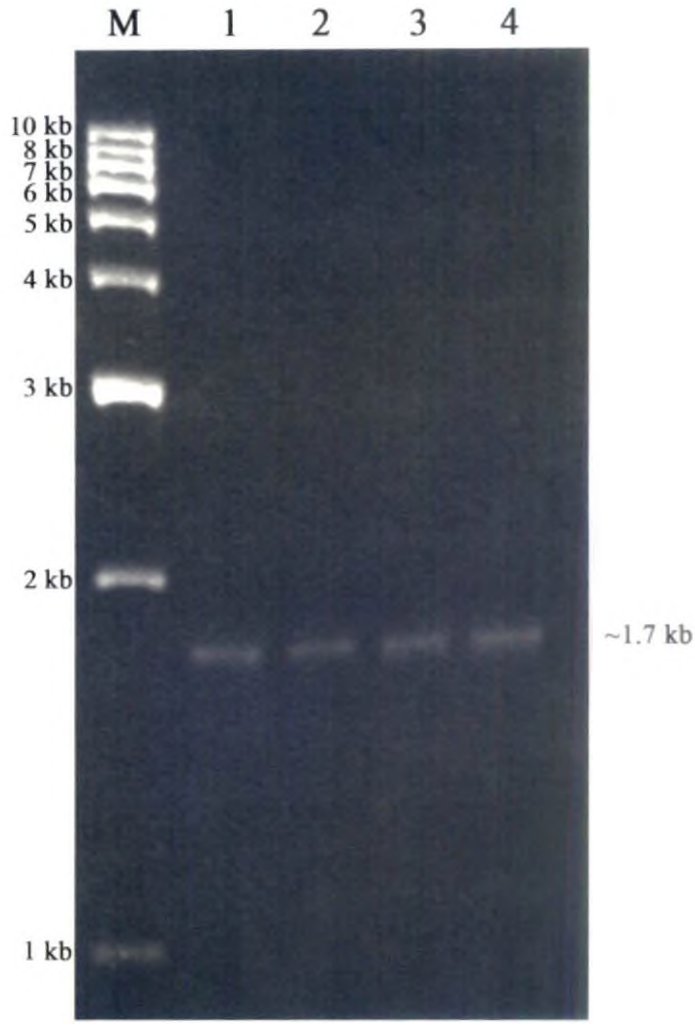
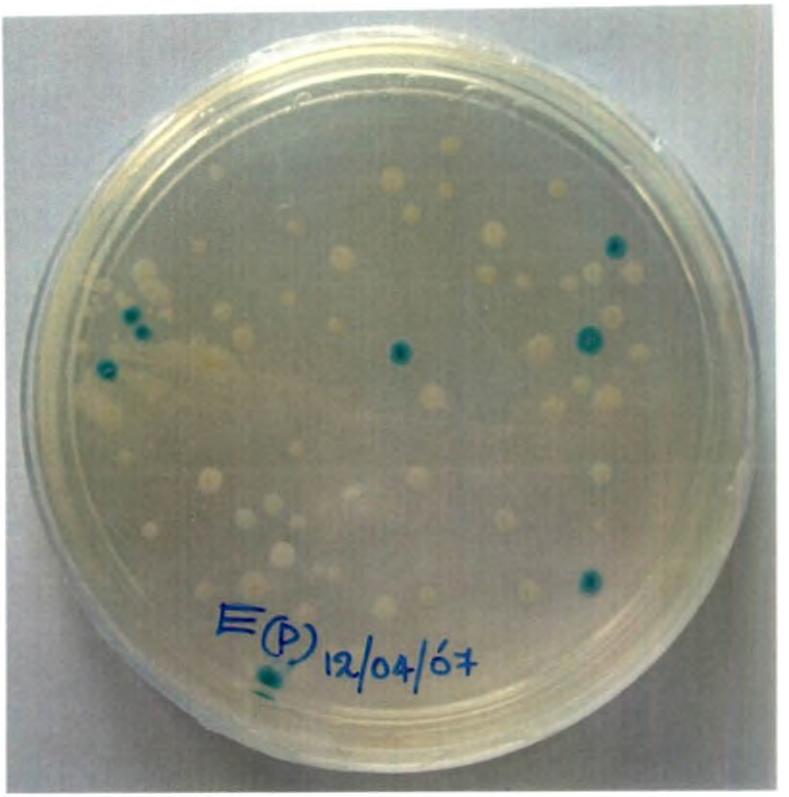
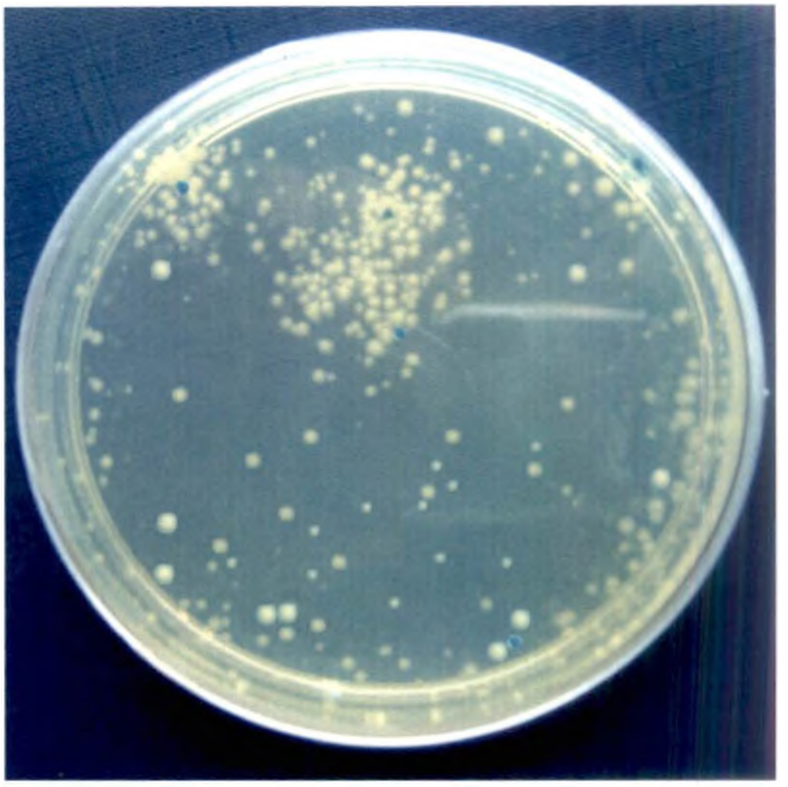


Plate 3. Gel showing purified PCR products
Lane M. 1 kb DNA ladder
Lane 1 to 4. Gel purified PCR products ~1.7 kb



A



B

Plate 4. Results of transformation

(A) Transformation of recombinant plasmid carrying α- LA gene

(B) Transformation of recombinant plasmid carrying control DNA

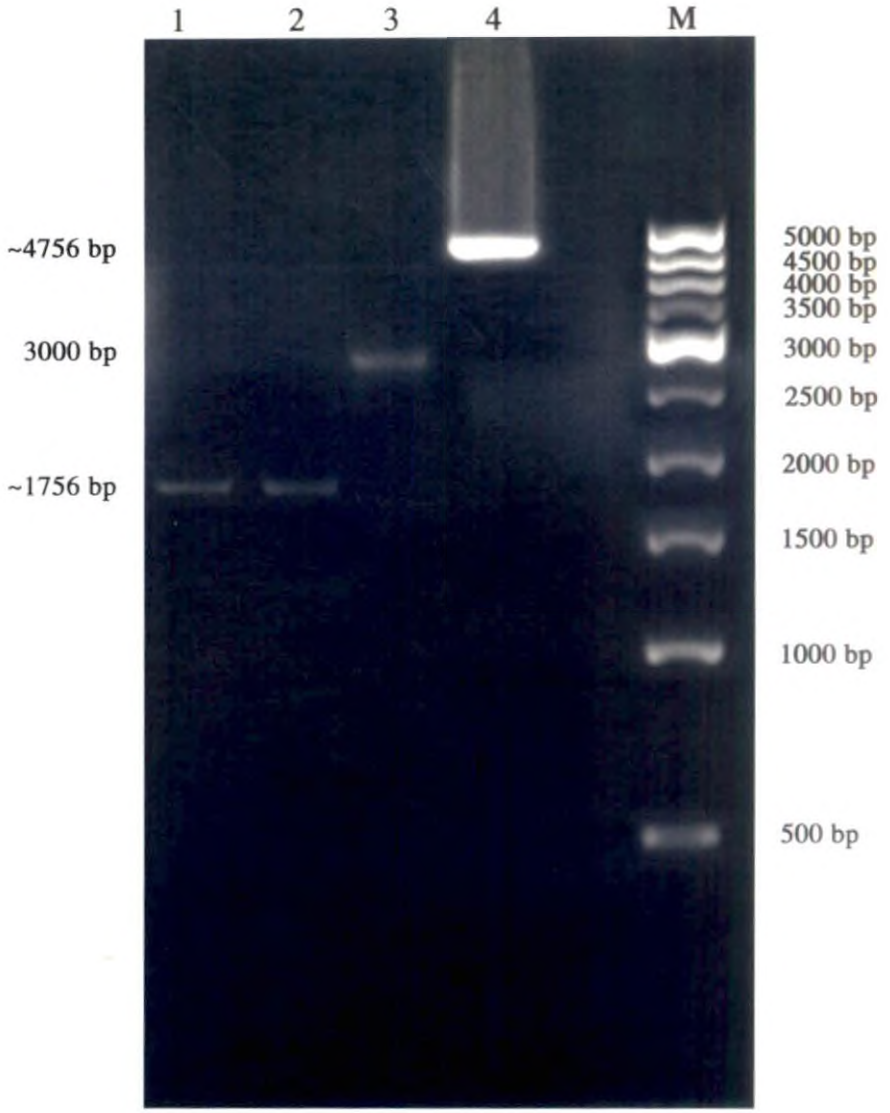


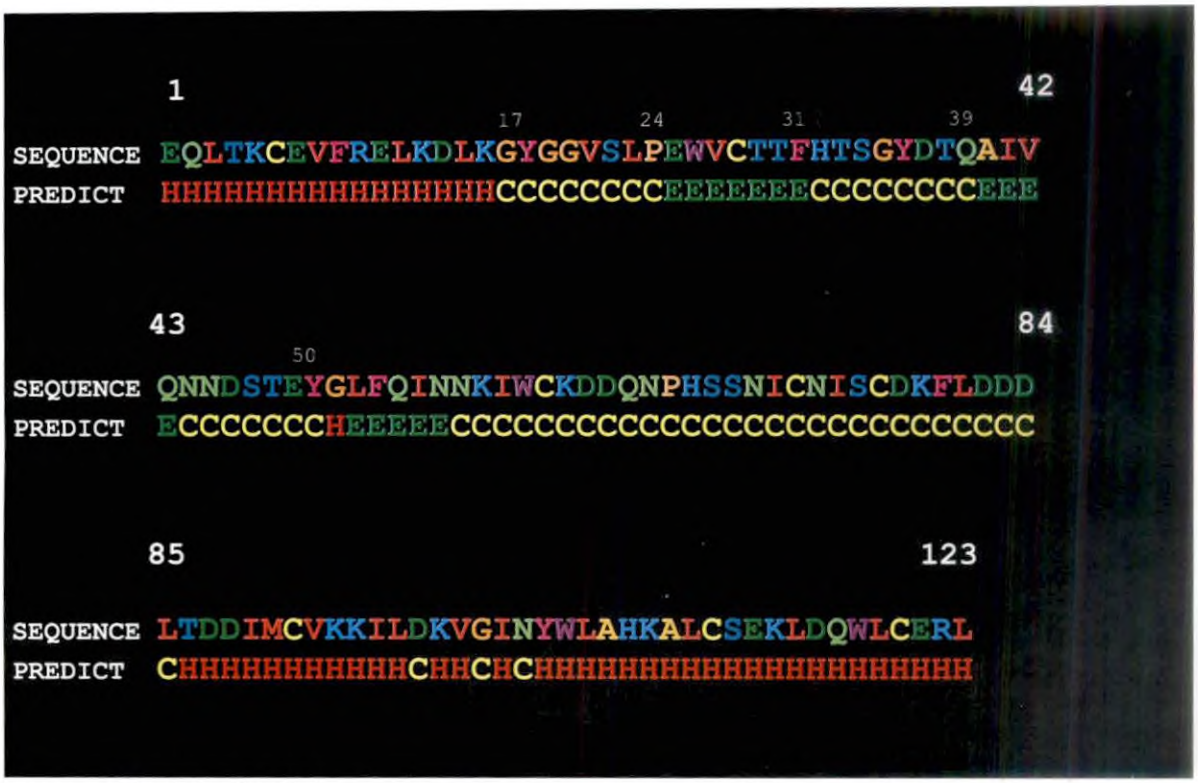
Plate 5. Gel showing recombinant vector

Lane 1- 2. Gel purified PCR product (~1756 bp)

Lane 3. T vector (3 kb)

Lane 4. Recombinant vector (~4.7 kb)

Lane M. 500 bp DNA ladder



H : alpha helix
E : beta sheet
C : coil

Plate 6. Predicted secondary structure of Vechur α-LA showing sites of alpha helices, beta sheets and coils

Plate 7. Three dimensional structure of α-lactalbumin
(A) Vechur α-lactalbumin predicted using Swiss Model
(B) Vechur α-lactalbumin predicted using ESyPred3D
(C) Human α-lactalbumin (PDB code 1A4V)

Discussion

5. DISCUSSION

One of the major differences of human milk over milk from other mammals, notably cow, goat, sheep and buffalo is the presence of α-LA as the major whey protein. The concentration of α-LA in the milk of these species is relatively low and instead the major whey protein is β-lactoglobulin. The gene sequence encoding for the *Bos taurus* α-LA and for the human α-LA proteins have been elucidated and the sequence information published by Vilotte *et al.* (1987) and Hall *et al.* (1987) respectively. The present study of characterization of α-LA in Vechur cattle, one of the indigenous breeds was aimed at revealing the homology of this protein in sequence and structure with the human and *Bos taurus* counterparts. The α-LA protein being a major ingredient in infant food, information on the degree of homology of Vechur α-LA with human α-LA would help to determine its proportion of incorporation in infant food. Isolation of α-LA cDNA and gene, and the knowledge of their primary structures were a pre requisite to further long range studies.

5.1 ISOLATION OF GENOMIC DNA

DNA isolation from whole blood was carried out using the standard phenol chloroform extraction procedure (Sambrook *et al.,* 1989), with the modification of overnight incubation of the WBC suspension with SDS and proteinase-K. Phenol chloroform extraction procedure is a common and efficient technique for DNA isolation from blood (Tantia *et al.,* 2004; Araujo *et al.*, 2006; Hirbo *et al.*, 2006).

5.2 PCR ANALYSIS

The polymerase chain reaction is an *in vitro* method of nucleic acid synthesis that enables the specific replication of a targeted segment of DNA, providing a rapid, highly sensitive and specific means of nucleic acid detection and isolation. A

standard reaction mixture contains the sample DNA, two oligonucleotide primers, thermostable *Taq* DNA polymerase and four deoxyribonucleotide triphosphates (dNTPs) in a buffered solution. Through repeated cycles of heat denaturation of the DNA, annealing of the primers to their complimentary sequences, and extension of annealed primers, an exponential accumulation of the specific target fragment is achieved (Saiki *et al.*, 1988).

5.2.1 Optimization of PCR

PCR conditions were optimized for the primers selected to obtain specific products and avoiding multiple bands on the gel. If the PCR conditions are well optimized, most of the added primers and dNTPs will be used up during the PCR amplifications, thus there will be only very few interfering factors during cloning and sequencing. For cloning the PCR fragment into a vector it is necessary to ensure that there is only one major PCR band visible on the gel.

Magnesium chloride concentration and annealing temperature of the primer were critical in obtaining specific amplifications. The annealing temperature for the primer pair was optimized using gradient thermal cycler in the range of 55 to 60 °C. An annealing temperature of 57.4 °C for 1.5 min was found optimal for the reaction. According to Don *et al.* (1991) the occurrence of spurious bands can be reduced by increasing the annealing temperature of PCR. Specific product was obtained at 1.5 m*M* $MgCl_2$ concentration.

Other conditions affecting the specificity of the PCR include concentration of enzyme and primers as well as the annealing time, extension time and number of cycles (Cha and Thilly, 1993). These conditions were kept constant for all the primers used. A concentration of 1U/reaction of *Taq* DNA polymerase and 30 p*M* of primer per reaction gave satisfactory results. Time for denaturation and extension was 1 min and for annealing was 1.5 min with a final extension time of 10 min.

5.3 GEL PURIFICATION

Purification of PCR products after amplification is usually recommended to obtain accurate sequencing data. Agarose contains various impurities which may inhibit downstream reactions if not efficiently removed from the DNA. Modern spin-column kits are very good at removing impurities. When cutting out the DNA band from the gel, we must visualise the band, in an ethidium bromide stained gel, in a dark-room on a UV trans-illuminator. The trans-illuminator was set to a long-wavelength UV and minimized the amount of time the DNA is exposed. This is because the UV light causes mutation in DNA at a measurable rate.

The unused primers, if contaminate the PCR product, can act as extension primers during sequencing, resulting in the generation of an additional set of dye-labeled sequencing fragments which can make sequencing data interpretation difficult. Post-purified PCR products were subjected to an agarose gel run along with an appropriate DNA ladder to examine the quality of the purified product. If smearing or multiple bands are present, we will not usually obtain a high quality sequence data. The PCR product purified using the Ultra Clean™ GelSpin™ DNA purification Kit (Catalog # 12400-50, MO BIO Laboratories, Inc. USA) was found to be of high quality suitable for sequencing.

5.4 CLONING

The cloning of genes, gene fragments and other DNA sequences is a fundamental part of molecular biology. This enables us to manipulate and study the function of a particular sequence. The PCR products generated using a nonproofreading DNA polymerase, which lacks, 3′→5′ exonuclease activity, have a single template-independent nucleotide at the 3′ end of the DNA strand (Clark, 1988; Newton and Graham, 1994). This single-nucleotide, which is most commonly an

'dA' residue, allows hybridization with and cloning into T-vectors, which have a complementary 3′ single dT overhang.

5.4.1 T-Cloning Vectors

T-vectors are a specific type of cloning vector that get their name from the dT overhangs added to a linearized plasmid. These vectors take advantage of the dA overhangs on PCR products after amplification with *Taq* DNA polymerase by providing compatible ends for ligation (Mezei and Storts, 1994; Robles and Doers, 1994). The single 3′-dT overhangs at the insertion site greatly improve the efficiency of ligation of a PCR product into the plasmids by preventing recircularization of the vector and providing a compatible overhang for PCR products generated by certain thermostable polymerases. Inserts of several kilobases have been successfully cloned into the T-Vectors (D'Avino *et al.,* 2004). However, as the insert gets larger, the ratio of vector to insert may need to be optimized further to maximize ligation efficiency.

5.4.2 Ligation

The amount of insert DNA for ligation was estimated by comparing the staining intensity with that of DNA molecular weight standards of known quantity on an ethidium bromide-stained agarose gel. In most cases, either a 1:1 or a 1:3 molar ratio of vector:insert works well. In the present study a 1:1 ratio gave excellent results as evidenced by large proportion of positive clones following transformation.

5.4.3 Transformation of Competent cells

5.4.3.1 Competent Cells

Transforming a newly constructed plasmid into competent *E. coli* cells is the primary method to propagate and select for the clone or clones of interest.

Competent bacterial cells are receptive to importing foreign DNA and replicating it. High-quality competent *E. coli* is an integral part of a successful cloning protocol. In the present study JM109 competent cells were prepared according to the manufacturer's instructions. These competent cells when transformed with the ligated plasmid DNA via the heat-shock method resulted in sufficient numbers of transformant clones. The cells are reported to be an ideal host for many molecular biology applications (Yanisch-Perron *et al.,* 1985) and can be used for α-complementation of β-galactosidase for blue/white screening.

5.4.3.2 Transformation and Screening

The ligated product was transformed into competent cells of *E. coli* JM109 cells prepared by $CaCl_2$ treatment as per the manufacture's instructions. Recombinants were selected through blue-white screening on LB-amp plates. Successful cloning of an insert into the T-Vector disrupts the coding sequence of the β-galactosidase α-peptide leads to failure of α-complementation resulting in formation of white colonies on X-gal/IPTG indicator plates following transformation of competent cells where as simple vector reclosure without insert results in blue colonies.

5.5 DNA SEQUENCING

Presence of the insert in single white colonies was confirmed by PCR with the same primer combination and purified plasmids from positive clones were sequenced at the DNA Sequencing Facility, Bangalore Genei Pvt. Ltd. Nucleotide sequences were determined by the dideoxy-chain termination method with the Applied Biosystems model 373A DNA sequencing system (Applied Biosystems, Foster City, Calif).

5.5.1 Structure of the Vechur α-LA Gene

The sequence data of 1756 bp obtained from sequencing of the cloned PCR product provide the complete structural organization of the transcription unit of the Vechur α-LA gene. The transcription unit with a size of 1724 bp begins at position 13 as evidence by the presence of the start codon 'ATG'. It is composed of four exons and three introns and sequences that respond to hormones and regulatory elements. The hexanucleotide sequence TGT(C/T)CT at positions 17 and 1698 might be possible sites for glucocorticoids as suggested by Scheidereit *et al.* (1986). The three intervening sequences begin and end with the consensus sequences (GTGAGT...) and (..CAG), respectively and these sequences involve in the splicing process in which the sequences corresponding to the introns are removed from the primary transcript of the gene as reported by Breathnach and Chambon (1981). The sequences upstream from the intron 3′ end such as the consensus sequence (CTGATCTCTCT) identified in the human globin gene (Rautmann and Breathnach, 1985) seem to be necessary for the splicing process to occur (Ruskin and Green, 1985). Sequences similar to the above mentioned occur at positions -29 (73 per cent homology), -39 (82 per cent) and -43 (73 percent) of the 3′ end of introns 1, 2 and 3, respectively.

5.5.1.1 Comparison with Bos taurus and Human α-LA

The sequences obtained for Vechur α-LA gene was found to be 99 per cent homologous to that of *Bos taurus* and this high homology was as expected. The 123 amino acid composition of the Vechur α-LA determined from the sequence was found to be in good agreement with the corresponding sequences of *Bos taurus* (Vilotte *et al.*, 1987) with a single variant amino acid of Arginine (AGG) at position 122 in Vechur against Lysine (AAG) in *Bos taurus*. The derived amino acid sequence contained the 19 amino acid residues of the signal peptide which was also

found to be identical to that reported for *Bos taurus* α-LA and ovine α-LA (Gaye *et al.*, 1987).

Comparison of the human α-LA gene and Vechur α-LA gene showed identical structural organization and identifies extensive homology within the transcription unit. The exon-intron boundaries in the Human α-LA gene occur at the codons of the amino acid residues, Leu-26 (intron 1), Lys-79 (intron 2) and Trp-104 (intron 3). The latter two residues are conserved in both species and also in *Bos taurus*, whereas Leu-26 is replaced by Trp-26 in the Vechur and *Bos taurus* proteins. This strict homology in the sites of insertion of introns suggests that the exon-intron organization of these genes was established before the divergence of these species. The positions of the exon-intron boundaries are also conserved as evidenced from similar sizes of the exons. The introns sizes are also comparable except in the case of intron 1, which is much larger in the human gene as a consequence of the insertion of a *Alu* family repeat sequence as reported by Deininger *et al.* (1981).

Examination of the nucleotide sequences for conserved regions thought to be involved in gene transcription and RNA processing reveals a number of features in common with other mammalian single-copy genes. The dinucleotides GT and AG occurring at the 5′ and 3′ ends of the introns, respectively is a feature in common with other eukaryotic genes (Mount, 1982).

Characterization of sequence homology between the human and Vechur α-LA genes by using BLAST2 sequence align programme demonstrated a high degree of homology within the exons. Evidence for some homology was also seen in intron 2. In contrast, intron 1 has diverged considerably both in length and in sequence and intron 3 in sequence. Part of the difference in length found in intron 1 can be accounted for by presence of *Alu* repetitive sequence in the human gene (Hall *et al.*, 1987). The di- and tri nucleotide repeats of (TG), (TCC) and (TAT) present in intron

1 and 3 of Vechur α-LA gene are not present in the human gene. Alignment of the Vechur and Human α-LA genes, allowing gaps for best fit showed homology between exons 1, 2, 3 and 4 to be 77, 86, 90, and 90 per cent, respectively.

A molecule of Vechur α-LA comprises 123 amino acid residues with only single residues of methionine and two residues each of arginine and proline. The alignment of the 123 amino acid sequences of the Vechur and Human α-LA showed a high degree of homology (74 %) with 31 variant amino acid residues in human protein. A relatively higher percentage of valine in the protein of Vechur and *Bos taurus* was found in comparison to human alpha lactalbumin. The percentage of tryptophan was also found to be higher in the α-LA of bovine species than the human counterpart. The percentage of essential amino acids in α-LA of human milk and Vechur milk was found to be 52 and 51, respectively.

The predicted secondary structure of Vechur α-LA showed that the larger α-helical lobe is formed by the amino- and carboxyl-terminal sections of the polypeptide chain (residues 1-16 and 86-123) while the smaller lobe, which encompasses a small three stranded antiparallel beta-sheet, and small irregular structure, is formed by the central section of the polypeptide chain (residues 17-85). The *Bos taurus* alpha lactalbumin consists of four α-helices, five 3_{10} helices, two β-structures, and an appreciable amount of random coil, representing 30, 20, 6 and 44 per cent, respectively (Acharya *et al.,* 1989). The predicted tertiary structure of Vechur α-LA also showed high homology with the *Bos taurus* and human α-LA structures.

Overall, the structures of Vechur α-LA was found to be very similar to that of *Bos taurus* and human reflecting their high degree of amino acid sequence identity (74-99 %). Similar results were also reported by Brew and Grobler (1992) with respect to human and *Bos taurus* proteins. The present study did not reveal any

higher degree of structural or functional similarity between Vechur and human α-LA proteins as compared to that of *Bos taurus*. The higher percentage of fat and total solids and lower fat globule size than other breeds of cattle is considered to be some reasons for the medicinal value of Vechur milk. The superiority of human milk and its high suitability to infants could be due to the higher content of α-LA and might not be attributed to any structural variations of the protein. This view is supported by the fact that the ratio of whey proteins to casein in human and bovine milks are 60: 40 and 20:80, respectively as reported by Lien (2003). Since the higher content of α-LA in human milk could be due to the high expression of this gene, further studies may be carried out to find out sequence variations, if any, occur in the regulatory sequences upstream of the gene. Another possibility could be the presence of *enhancer* sequences or variation in the induction of expression of this protein in human. Gene expression studies are suggestive as α-LA locus can also be used as a genetic marker to increase milk production in Vechur cattle, as this marker may be directly responsible for increased milk production.

Summary

6. SUMMARY

The awareness of the value of native cattle of Kerala, especially of Vechur cattle was the stimulus for the study. To verify the uniqueness of this breed, it is essential to undertake the characterization works at molecular level. The study helps in exploring the basis of popular beliefs about Vechur cow milk. It is believed to be easily digestible and good for infants, convalescents and old people. The studies on fat globules and iodine value were suggestive of a scientific basis for this. The Vechur cow milk is used as an important ingredient in some Ayurvedic medicines and medicated oils, as a soothening and healing agent in case of irritation or injuries of eye.

Genomic DNA was isolated from Vechur cattle using the phenol-chloroform extraction procedure. A 1756 bp fragment of alpha lactalbumin (α-LA) gene which covers the complete gene was amplified by PCR using specific primers designed on the basis of bovine sequence available in the database. The PCR assay was performed in 200 µl reaction tubes with a total volume of 50 µl, containing 100 ng of DNA template, 1 X PCR buffer, 1.5 m*M* $MgCl_2$, 30 p*M* of each primer, 200 µ*M* of each dNTP and one unit *Taq* DNA polymerase. The PCR product was checked in one per cent agarose gel in 1 X TAE buffer in a horizontal submarine electrophoresis unit and kept at -20 °C in the freezer till use.

The PCR product was purified by a preparative gel electrophoresis using a gel elution kit (Mo Bio laboratories, USA) following manufacturer's instructions. The purified DNA was analyzed on agarose gel, checked for the presence of any smearing, primer-dimer or additional bands, and quantified before using it for cloning.

The gel purified PCR product was cloned in to TA cloning vector prior to sequencing. The ligation reaction was carried out in 10 µl volume and contained 3µl

of PCR product, 5 µl of 2X instant ligation buffer, 1 µl T- Vector and 1 µl Instant T4 DNA ligase. The ligated product was examined in one per cent agarose gel in 1 X TAE buffer in a horizontal submarine electrophoresis unit and kept at -20 °C till transformation.

Transformation was induced by treating host *E. coli* cells with ice-cold calcium chloride followed by a brief exposure to high temperature using the Genie Transformation Kit. The recombinant plasmid and the competent bacterial cells were mixed and incubated at 42 °C in a water bath for 2 min and then returned to ice for 5 min. The mixture was then mixed with 1 ml of LB broth and incubated in a shaker at 37 °C for one hour. These transformed cells at different volumes were added on to LB-Amp plates with X-gal and IPTG and incubated for 37 °C overnight. The cells carrying recombinant plasmid produced white colonies.

The recombinant plasmid carrying the α-LA gene insert was isolated from the white clones by a modified SDS-alkaline lysis method using the AxyPrep Plasmid Miniprep Kit. The purified plasmid gene insert was examined by a one per cent agarose gel run and the band showed a molecular size of approximately 4.7 kb. The purified recombinant plasmid was sequenced by the dideoxynucleotide sequencing method using an automated DNA sequencer (Applied Biosystems, USA) at Bangalore Genei, Pvt. Ltd.

The 1756 bp nucleotide sequence of the α-LA gene and the predicted protein of Vechur cattle was compared with corresponding sequences from bovine and other species through BLAST analysis. The nucleotide sequence showed 99 and 74 per cent homology with *Bos taurus* and human α-LA genes respectively. Multiple sequence analysis was done with the corresponding nucleotide and amino acid sequences of livestock species and human and a phylogram was constructed using the ClustalW program. The results of ClustalW multiple nucleotide sequence

alignment showed a very high homology with *Bos taurus*, buffalo and goat with a score of 99, 97 and 95 respectively and with human the score was found to be 63. The ClustalW results of amino acid sequences also produced high alignment scores with *Bos taurus*, buffalo, goat and human with scores of 99, 98, 94 and 72 respectively.

The Vechur α-LA contains 123 amino acid residues with amino-terminal glutamic acid and carboxyl-terminal leucine and alignment of the Vechur α-LA amino acid sequence with bovine α-LA showed high degree of homology with a single variant amino acid of arginine (AGG) at position 122 in Vechur against lysine (AAG) in *Bos taurus* and with human α-LA, 74 percent identity was observed.

The secondary structure of the Vechur α-LA was predicted using the PAPIA system (http://mbs.cbrc.jp/papia/papia.html). The predicted secondary structure showed that the larger α- helical lobe is formed by the amino- and carboxyl-terminal sections similar to bovine alpha lactalabumin. The three dimensional structure of the protein was predicted using automatic comparative modeling servers Swiss Model Server and ESyPred3D. The servers selected bovine α-LA (PDB structure Id=1F6RA) which has 99 per cent sequence identity as template for predicting the structure. The ALIGN program from the EsyPred3D server showed that the templates 1F6R chain A and human α-LA (PDB structure Id= 1A4V) chain A shared 85.9 per cent and 64.8 per cent homology with Vechur sequence, respectively.

The Vechur α-LA showed very high homology to that of *Bos taurus* and human reflecting their high degree of amino acid sequence identity (74-99 %). The study did not reveal any higher structural or functional similarity between α-LA of Vechur and human. The superiority of human milk and its high suitability to infants could be due to the higher content of α-LA and might not be attributed to any structural variations of the protein. Since the higher content of α-LA in human milk

could be due to the high expression of this gene, further studies may be carried out to find out sequence variations, if any, occur in the regulatory sequences upstream of the gene. The presence of *enhancer* sequences or variation in the induction of expression of this protein in human may also be responsible. Gene expression studies are suggestive as α-LA locus can also be used as a genetic marker to increase milk production in Vechur cattle, as this marker may be directly responsible for increased milk production.

References

REFERENCES

Abed, Y., Bollet, C. and De Micco, P. 1995. Comparative study of three methods for cloning PCR products. *World J. Microbiol. Biotechnol.* 11: 478-480

Acharya, K.R., Stuart, D.I., Walker, N.P.C., Lewis, H. and Philips, D.C. 1989. Refined structure of baboon alpha-lactalbumin at 1.7 Angstroms resolution. Comparison with C-type lysozyme. *J. Mol. Biol.* 208: 99-127

Acharya, K.R., Ren, J.S., Stuart, D.I., Phillips, D.C. and Fenna, R.E. 1991. Crystal structure of human alpha-lactalbumin and the lactose synthetase reaction. *J. Mol. Biol.* 221: 571-581

Ahmad, N.N., Cu-Unjieng, A.B. and Donoso, L.A. 1995. Modification of standard proteinase K/phenol method for DNA isolation to improve yield and purity from frozen blood. *J. Med. Genet.* 32: 129-130

Alexander, L.J., Stewart, A.F., Mackinlay, A.G., Kapelinskaya, T.V., Tkach, T.M. and Gorodetsky, S.I. 1988. Isolation and characterization of the bovine κ-casein gene. *Eur. J. Biochem.* 178: 395-401

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol.* 215: 403-410

Andrews, P. 1969. Lactose synthetase A protein from human milk. *Biochem. J.* 111: 14-15

Apparao, K.B.C., Bhat, P.P. and Bhat, P.N. 1994. A rapid and simple procedure for isolation of genomic DNA from blood for RFLP studies in livestock animals. *Indian J. Anim. Sci.* 64: 460-462

Araujo, A.M., Guimaraes, S.E.F., Machado, T.M.M., Lopes, P.S., Pereira, C.S., Silva, F.L.R., Rodrigues, M.T., Columbiano, V.S. and Fonseca, C.G. 2006. Genetic diversity between herds of Alpine and Saanen dairy goats and the naturalized Brazilian Moxoto breed. *Genet. Mol. Biol.* 29: 67-74

Aravindakshan, T.V., Nainar, A.M. and Nachimuthu, K. 1998. Extraction of DNA from cattle blood cells. *Cheiron* 27: 5-6

Aravindakshan, T.V., Venkatachalapathy, R.T., James, S.P., Kurien, E. and Iype, S. 2003. Polymorphism of the prolactin gene locus in Vechur and other dwarf (Kassargode) cattle of Kerala detected by PCR-RFLP. *Indian J. Dairy Sci.* 57: 246-249

Aravindakshan, T.V., Venkatachalapathy, R.T., Iype, S. and James, S.P. 2004. Polymorphism at the the β-lactoglobulin locus in Vechur and Dwarf cattle of Kerala using PCR-RFLP. *Indian J. Dairy Sci.* 56: 163-165

Aravindakshan, T.V., James, S.P and Raghunandanan, K.V. 2006. Screening for bovine leukocyte adhesion deficiency (BLAD) in Vechur cattle. *Indian J. Anim. Sci.* 76(9): 712-715

Arnold, K., Bordoli, L., Kopp. J. and Schwede, T. 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling . *Bioinformatics* 22(2): 195-201

Arranz, J.J., Bayon, Y. and San Primitivo, F. 2001. Genetic variation at microsatellite loci in Spanish sheep. *Small Rum. Res.* 39: 3-10

Banerjee, M.R. 1976. Responses of mammary cells to hormones. *Int. Rev. Cytol.* 47: 1-97

Barker, J.S.F., Tan, S.G., Moore, S.S., Mukherjee, T.K., Matheson, J.L. and Selvaraj, O.S. 2001. Genetic variation within and relationships among population of Asian goats (*Capra hircus*). *J. Anim.Breed. Genet.* 118: 213-233

Bhure, S. and Sharma, B. 2007. Cloning and characterization of ovine αS1-casein gene promoter: A transfection study in rat mammary gland cell line. *DNA Sequence* 18: 39 - 46

Bleck, G.T. and Bremel, R.D. 1993. Correlation of the alpha lactalbumin (+15) polymorphism to milk composition of Holsteins. *J. Dairy Sci.* 76: 2292-2298

Booij, L., Merens, W., Markus, C.R. and Van der Does, V.J. 2006. Diet rich in α-lactalbumin improves memory in unmedicated recovered depressed patients and matched controls. *J. Psychopharmacol.* 20: 526-535

Boston, W.S., Bleck, G.T., Conroy, J.C., Wheeler, M.B. and Miller, D.J. 2001. The effects of increased expression of α-lactalbumin in transgenic mice on milk yield and pup growth. *J. Dairy Sci.* 84: 620-622

Brawerman, G., Mendecki, J. and Lee, S.Y. 1972. A procedure for the isolation of mammalian messenger ribonucleic acid. *Biochemistry* 11: 637–641

Breathnach, R. and Chambon, P. 1981. Organization and expression of eukaryotic split genes coding for proteins. *Annu. Rev. Biochem.* 50: 349–383

Brew, K., Vanaman, T.C. and Hill, R.L. 1968. The role of α-lactalbumin and the A protein in lactose synthetase: a unique mechanism for the control of a biological reaction. *Proc. Natl. Acad. Sci. USA.* 59: 491-497

Brew, K. and Grobler, J.A. 1992. α-lactalbumin. *Adv. Dairy Chem.* 1: 191–229

Brodbeck, U., Denton, W.L., Tanahashi, N. and Ebner, K.E. 1967. The isolation and identification of the B protein of lactose synthetase as alpha-lactalbumin. *J. Biol. Chem.* 242: 1391–1397

Cha, R.S. and Thilly, W.G. 1993. Specificity, efficiency, and fidelity of PCR. *Genome Res.* 3: 18-29

Chan, E.Y. 2005. Advances in sequencing technology. *Mutat. Res.* 573: 13-40

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G. and Thompson, J.D. 2003. Multiple sequence alignment with the Clustal series of programs. *Nucl. Acids Res.* 31: 3497-3500

Clark, J.M. 1988. Novel non-templated nucleotide addition reactions catalyzed by prokaryotic and eucaryotic DNA polymerases. *Nucl. Acids Res.* 16: 9677-9686

Cohen, S.N., Chang, A.C.Y., Boyer, H.W. and Helling, R.B. 1973. Construction of biologically functional bacterial plasmid *in vitro. Proc. Natl. Acad. Sci. USA* 70: 3240-3244

D'Avino, P.P., Savoian, M.S. and Glover, D.M. 2004. Mutations in *sticky* lead to defective organization of the contractile ring during cytokinesis and are enhanced by *Rho* and suppressed by *Rac. J. Cell Biol.* 166: 61–71

Davies, M.S., West, L.F., Davis, M.B., Povey, S. and Craig, R.K. 1987. The gene for human α-lactalbumin is assigned to chromosome 12q13. *Ann. Hum. Genet.* 51: 183-188

Deininger, P.L., Jolly, D.J., Rubin, C.M., Friedmann, T. and Schmid, C.W. 1981. Base sequence studies of 300 nucleotide renatured repeated human DNA clones. *J. Mol. Biol.* 151: 17–33

Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K. and Mattick, J.S. 1991. 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucl. Acids Res.* 19: 4008

Dong, H. and Kurland, C.G. 1996. Bacterial growth inhibition by overproduction of proteins. *Mol. Microbiol.* 21: 1-4

Findlay, J.B.C. and Brew, K. 1972. The complete amino acid sequence of human alpha-lactalbumin. *Eur. J. Biochem.* 27: 65-86

Fitzgerald, D.K., Brodbeck, U., Kiyosawa, I., Mawal, R., Colvin, B. and Ebner, K.E. 1970. Alpha-lactalbumin and the lactose synthetase reaction. *J. Biol. Chem.* 245: 2103-2108

Forsum, E. 1973. Nutritional evaluation of whey protein concentrates and their fractions. *J. Dairy Sci.* 57: 665-670

Gaye, P., Hue-Delahaie, D., Mercier, J.C., Soulier, S., Vilotte., J.L. and Furet, J.P. 1987. Complete nucleotide sequence of ovine alpha-lactalbumin mRNA. *Biochimie.* 69: 601-608

Girija, C.R. 1995. Characterization and evaluation of dwarf cattle of Kerala. PhD thesis, Kerala Agricultural University, Thrissur. 118p.

Gordon, K., Lee, E., Vitale, J.A., Smith, A.E., Westphat, H. and Hennighausen, L. 1980. Production of human tissue plasminogen activator in transgenic mouse milk. *Bio/Technology* 5: 1183-1187

Guo, B. and Bi, Y. 2002. Cloning PCR products. An overview. *Methods Mol. Biol.* 192: 1111-1119

Hakansson, A., Zhivotovsky, B., Orrenius, S., Sabharwal, H. and Svanborg, C. 1995. Apoptosis induced by a human milk protein. *Proc. Natl. Acad. Sci. USA*. 92: 8064-8068

Hakansson, A., Svensson, M., Mossberg, A.K., Sabharwal, H., Linse, S., Lazou, I., Lonnerdal, B. and Svanborg, C. 2000. A folding variant of alpha lactalbumin with bactericidal activity against *Streptococcus pneumoniae. Mol. Micriobiol.* 35: 589-600

Hall, L., Emery, C.D., Davies, M.S., Parker, D. and Craig, R.K. 1987. Organization and sequence of the human α-lactalbumin gene. *Biochem. J.* 242: 735-742

Heine, W.E., Klein, P.D. and Reeds, P.J. 1991. The importance of α-lactalbumin in infant nutrition. *J. Nutr.* 121: 277–283

Heine, W.E., Radke, M. and Wutzke, K.D. 1995. The significance of tryptophan in human nutrition. *Amino Acids* 9: 91-205

Hill, R.L. and Brew, K. 1975. Lactose synthetase. *Adv. Enzymol. Relat. Areas Mol. Biol.* 72: 105-157

Hiraoka, Y., Segawa, T., Kuwajima, K., Sugai, S. and Murai, N. 1980. Alpha-lactalbumin: a calcium metallo-protein. *Biochem. Biophys. Res. Commun.* 95: 1098-1104

Hirbo, J., Muigai, A., Naqvi, A.N., Rege, E.D. and Hanotte, O. 2006. The genetic diversity of trans-Caucasian native sheep breeds. *Asian-Aust. J. Anim. Sci.* 19: 943-952

Hurley, W.L. and Schuler, L.A. 1987. Molecular cloning and nucleotide sequence of a bovine alpha lactalbumin cDNA. *Gene* 61: 119-122

Icho, T., Sparrow, C.P. and Raetz, C.R.H. 1985. Molecular Cloning and Sequencing of the Gene for CDP-diglyceride Synthetase of *Escherichia coli. J. Biol. Chem.* 260: 12078-12083

Innis, M.A. and Gelfand, D.H. 1990. Optimization of PCRs. *PCR Protocols.* (eds. Innis, M.A., Gelfand, D.H., Sninsky, J.J. and White, T.J.). Academic Press, New York, pp. 3-12

Iype, S. and Venkatachalapathy, R.T. 1997. Vechur Cattle of Kerala. *Animal Genetic Resource Information.* FAO (UNEP), Rome. 18: 62

Iype, S. and Venkatachalapathy, R.T. 2001. *Vechur Cattle of Kerala*, Kerala Agricultural University, Mannuthy, Kerala. 32p.

Jackson, J.G., Janszen, D.B., Lonnerdal, B., Lien, E.L., Pramuk, K.P. and Kuhlman, C.F. 2004. A multinational study of α-lactalbumin concentrations in human milk. *J. Nutr. Biochem.* 15(9): 517-521

Kaczmarczyk, E., Malolepszy, J., Nosowicz, B.B. and Bongarc, E. 2005. Polymorphism of the alpha lactalbumin gene and populations and sub populations of peripheral blood lymphocytes in heifers naturally infected with Bovine Leukemia Virus. *Bull Vet. Inst. Pulawy.* 49: 279-286

Knapp, S. 2003. Phylogenesis and the tree of life. *Phylogenesis foa's ark.* Foreign office architects. ICA, London, pp. 638-647

Kuhn, N.J., Carrick, D.T. and Wilde, C.J. 1980. Lactose synthesis: possibilities of regulation. *J. Dairy Sci.* 63: 328-336

Laird. J.E., Jack, L., Hall, L., Boulton, A.P., Parker, D. and Craig, R.K. 1998. Structure and expression of guinea-pig α-lactalbumin gene. *Biochem. J.* 254: 85-94

Lambert, C., Leonard, N., De Bolle, X. and Depiereux, E. 2002. ESyPred3D: Prediction of proteins 3D structures. *Bioinformatics*. 18(9): 1250-1256

Larson, B.L. 1985. Biosynthesis and cellular secretion of milk. *Lactation*. (ed. Larson, B.L.). The Iowa State University Press, Ames. 163 p.

Lee, S.Y. 1996. High cell-density culture of *Escherichia coli*. *Trends Biotechnol.* 14: 98-105

Lee, H., Kim, K.N. and Chae, Y.K. 2007. Reevaluating the capability of *Taq* DNA polymerase: Long PCR Amplification. *Protein and Peptide Letters*. 14: 321-323

Lexa, M., Horak, J., and Brzobohaty, B. 2001. Virtual PCR. *Bioinformatics*. 17: 192-193

Lien, E.L. 2003. Infant formulas with increased concentrations of α-lactalbumin. *Am. J. Clin. Nutr.* 77: 1555S-1558S

Liu, M.Q., Shen, X., Yin, W.L. and Lu, C.F. 2007. Changes of 5' terminal nucleotides of PCR primers causing variable T-A cloning efficiency. *J. Integrative Plant Biol.* 49(3): 382–385

Lonnerdal, B. and Lien, E.L. 2003. Nutritional and physiologic significance of α-lactalbumin in infants. *Nutr. Rev.* 61: 295-305

Maga, E.A. and Murray, J.D. 1995. Mammary gland expression of transgenes and the potential for altering the properties of milk. *Bio-Technology* 13: 1452-1457

Marchuk, D., Drumm, M., Saulino, A. and Collins, S. 1990. Construction of T-vectors, a rapid and general system for direct cloning of unmodified PCR products. *Nucl. Acids Res.* 19: 1154-1155

Markus, C.R., Olivier, B., Panhuysen, G.E.M., Van der Gugten, J., Alles, M.S., Tuiten, A., Westenberg, G.M., Fekkes, D., Koppeschaar, H.F. and Haan, E.E.H.F. 2000. The bovine protein α-lactalbumin increases the plasma ratio of tryptophan to the other large neutral amino acids, and in vulnerable subjects raises brain serotonin activity, reduces cortisol concentration, and improves mood under stress. *Am. J. Clin. Nutr.* 71: 1536-1544

Markus, C.R., Olivier, B. and Hann, E.E.H.F. 2002. Whey Protein rich in alpha lactalbumin increases the ratio of plasma tryptophan to the sum of the other large neutral amino acids and improves cognitive performance in stress vulnerable subjects. *Am. J. Clin. Nutr.* 75: 1051-1056

Matsumoto, H., Shimokawa, Y., Ushida, Y. and Toida, T. 2001. New biological function of bovine alpha-lactalbumin: protective effect against ethanol-and stress- induced gastric mucosal injury in rats. *Biosc. Biotechnol. Biochem.* 65: 1104-1111

McGinnis, S. and Madden, T.L. 2004. BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucl. Acids Res.* 32: 20-25

Mead, D.A., Pey, N.K., Hermstadt. C., Marcil, R.A. and Smith, L.M. 1991. A universal method for the direct cloning of PCR amplified nucleic acid. *Bio/Technology* 9: 657-663

Mercier, J.C., Haze, G., Gaye, P., Petrissant, G., Hue, D. and Boisnard, M. 1978. Amino terminal sequence of the precursor of ovine alpha-lactalbumin. *Biochem. Biophys. Res. Commun.* 85: 662-670

Mercier, J.C. and Vilotte, J.L. 1993. Structure and function of milk protein genes. *J. Dairy Sci.* 76: 3079-3098

Mezei, L.M. and Storts, D.R. 1994. Purification of PCR products. *PCR Technology: Current Innovations*. (eds. Griffin, H.G. and Griffin, A.M.). CRC Press, Boca Raton. pp. 21-27

Miller, H. 1987. Practical aspects of preparing phage and plasmid DNA: Growth, maintenance, and storage of bacteria and bacteriophage. *Methods Enzymol.* 152: 145-170

Mincheva-Nilsson, L., Hammarstrom, M.L., Juto, P. and Hammarstrom, S. 1990. Human milk contains proteins that stimulate and suppress T lymphocyte proliferation. *Clin. Exp. Immunol.* 79: 463-469

Morrison, D.A. 1977. Transformation in *Escherichia coli*: cryogenic preservation of competent cells. *J. Bacteriol.* 132: 349-351

Mount, S.M. 1982. A catalogue of splice junction sequences. *Nucl. Acids Res.* 10: 459-472

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. and Erlich, H.A. 1986. Specific enzymatic amplification of DNA invitro: the polymerase chain reaction. *Cold Spring Harbor Symposia on Quantitative Biology.* 51: 263-273

Newton, C.R. and Graham, A. 1994. *PCR.* BIOS Scientific Publishers, Ltd., Oxford, UK. 13p.

Novy, R.E., Yaeger, K.E. and Kolb, K.M. 1996. Perfectly blunt cloning: A superior method for cloning PCR products or any DNA. *InNovations.* 6: 7-11

Okamoto, A., Kosugi, A., Koizumi, Y., Yanagida, F. and Udaka, S. 1997. High efficiency transformation of *Bacillus brevis* by electroporation. *Biosci. Biotechnol. Biochem.* 61: 202-203

Oliner, J.D., Kinzler, K.W. and Vogelstein, B. 1993. In vivo cloning of PCR products in *E. coli. Nucl. Acids Res.* 21: 5192-5197

Orosco, M., Rouch, C., Beslot, F., Feurte, S., Regnault, A. and Dauge, V. 2004. Alpha-lactalbumin enriched diets enhance serotonin release and induce anxiolytic and rewarding effects in the rat. *Behav. Brain Res.* 148: 1-10

Permyakov, E.A. and Berliner, L.J. 2000. α-Lactalbumin: structure and function. *FEBS Lett.* 473: 269-274

Pettersson, J., Mossberg, A.K. and Svanborg, C. 2006. Alpha-lactalbumin species variation, HAMLET formation, and tumor cell death. *Biochem. Biophys. Res. Commun.* 23: 260-270

Pillai, T.K.V. 1940. *Travancore state Mannual.* Vol. III. Published by Government of Travancore. 383p.

Qasba, P.K. and Safaya, S.K. 1984. Similarity of nucleotide sequences of rat α-lactalbumin and chicken lysozyme genes. *Nature* 608: 377-378

Rao, K.R. and Brew, K. 1989. Calcium regulates folding and disulfide-bond formation in alpha-lactalbumin. *Biochem. Biophys. Res. Commun.* 163: 1390–1396

Rautmann, G. and Breathnach, R. 1985. A role for branchpoints in splicing *in vivo*. *Nature* 315: 430 - 432

Robles, J. and Doers, M. 1994 pGEM®-T Vector Systems troubleshooting guide. *Promega Notes* 45: 19-20

Ruskin, B. and Green, M.R. 1985. Role of the 3' splice site consensus sequence in mammalian pre-mRNA splicing. *Nature* 317: 732-734

Rychlik, W., Spencer, W.J. and Rhoads, R.E. 1990. Optimization of the annealing temperature for DNA amplification *in vitro*. *Nucl. Acids Res.* 18: 6409-6412

Rychlik,W. 1995. Selection of primers for polymerase chain reaction. *Mol. Biotechnol.* 3: 129-134

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B. and Erlich, H.A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239: 487-491

Sambrook, J., Fritsch, E.F. and Maniatis, T. 1989. *Molecular Cloning: A Laboratory Manual.* Second edition. Cold Spring Harbor Laboratory Press. New York, p. 1886

Sanger, F. and Coulson, A. R. 1975. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* 94(3): 441-448

Sanger, F., Nicklen, S. and Coulson, A.R. 1977. DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74: 5463-5467

Scheidereit, C., Westphal, H. M., Carlson, C., Bosshard, H. and Beato, M. 1986. Molecular model of the interaction between the glucocorticoid receptor and the regulatory elements of inducible genes. *DNA* 5: 383-391

Suprabha, P., Anilkumar, K., Aravindakshan, T.V. and Raghunandanan, K.V. 2005. Assessment of bandsharing values in RAPD-PCR analysis of dwarf cattle of Kerala. *Asian-Aust. J. Anim. Sci.* 18(9): 1217-1220

Svensson, M., Sabharwal, H., Hakansson, A., Mossberg, A.K., Lipniunas, P., Leffler, H., Svanborg, C. and Linse, S. 1999. Molecular characterization of alpha-lactalbumin folding variants that induce apoptosis in tumor cells. *J. Biol. Chem.* 274: 6388-6396

Svensson, M., Hakansson, A., Mossberg, A.K., Linse, S. and Svanborg, C. 2000. Conversion of alpha-lactalbumin to a protein inducing apoptosis. *Proc. Natl. Acad. Sci. USA.* 97: 4221-4226

Tantia, M.S., Behl, R., Sheoran, N., Singh, R. and Vijh, R.K. 2004. Microsatellite data analysis for conservation of two goat breeds. *Indian J. Anim. Sci.* 74: 761-767

Tatusova, T. and Madden, T. L. 1999. Blast 2 sequences. A new tool for comparing protein and nucleotide sequences. *FEMS Microbiol. Lett.* 12: 247 -250

Topcu, Z. 2000. An optimized recipe for cloning of the polymerase chain reaction-amplified DNA inserts into plasmid vectors. *Acta Biochimica Polonica.* 47(3): 841-846

Tsen, S.D., Fang, S.S., Chen, M.J., Chien, J.Y., Lee, C.C. and Tsen, D.H. 2002. Natural plasmid transformation in *Escherichia coli. J Biomed Sci.* 9: 246-252

Tucker, H.A. 1981. Physiological control of mammary growth, lactogenesis, and lactation. *J. Dairy Sci.* 64: 1403-1421

Venkatachalapathy, R.T. 1996. Milk protein polymorphism and milk composition in Vechur cattle. MVSc thesis, Kerala Agricultural University, Thrissur. 97p.

Venkatachalapathy, R.T. and Iype, S. 1997. Fat globule size and distribution in milk and iodine value of Vechur cattle of Kerala. *Indian J. Anim. Sci.* 67(10): 910-912

Vilotte, J.L., Soulier, S., Mercier, J.C., Gaye, P., Hue-Delahaie, D. and Furet, J.P. 1987. Complete nucleotide sequence of bovine alpha lactalbumin gene: comparison with its rat counterpart. *Biochimie.* 69: 609-620

Vilotte, J.L., Soulier, S., Printz, C. and Mercier, J.C. 1991. Sequence of the goat α-lactalbumin-encoding gene: comparison with the bovine gene and evidence of related sequences in the goat genome. *Gene* 98: 271-276

Voelker, G.R., Bleck, G.T. and Wheeler, M.B. 1997. Single-base polymorphisms within the 5′ flanking region of the bovine α-lactalbumin gene. *J. Dairy Sci.* 80: 194-197

Wall, R.J., Kerr, D.E. and Bondioli, K.R. 1997. Transgenic dairy cattle: Genetic engineering on a large Scale. *J. Dairy Sci.* 80: 2213–2224

White, B.A. 2005. Polymerase Chain Reaction (PCR): Design and optimization of reactions. *Encyclopedia of Life Sciences.* 60: 25-30

Wong, C.W., Seow, H.F., Husband, A.J., Regester, G.O. and Watson, D.L. 1997. Effects of purified bovine whey factors on cellular immune functions in ruminants. *Vet. Immunol. Immunopathol.* 56: 85-96

Wu, D.Y., Ugozzoli, L., Pal, B.K., Qian, J. and Wallace, R.B. 1991. The effect of temperature and oligonucleotide primer length on the specificity and efficiency of amplification by the polymerase chain reaction. *DNA Cell Biol.* 10: 233-238

Yanisch-Perron, C., Vieira, J. and Messing, J. 1985. Improved M13 phage cloning vectors and host strains: Nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* 33: 103–119

Yap, E.P.H. and McGee, J.O.D. 1991. Short PCR product yields improved by lower denaturation temperatures. *Nucl. Acids Res.* 19: 1713

Ye, J., McGinnis, S. and Madden, T.L. 2006. BLAST: improvements for better sequence analysis. *Nucl. Acids Res.* 34: 6-9

Zhou, M.Y., Clark, S.E. and Sanchez, G.C.E. 1995. Universal Cloning Methods by TA Strategy. *Bio Techniques* 19: 34-35

Zuelke, K.A. 1998. Transgenic modification of cows milk for value-added processing. *Reprod. Fertil. Dev.* 10: 671-676

Annexures

ANNEXURE – 1

COMPOSITION OF REAGENTS AND BUFFERS USED IN THE STUDY

Agarose (1 %)

Weighed 0.6 g of agarose powder and mixed with 60 ml of 1 X TAE buffer in a conical flask. Solution was heated in a microwave oven until boiling and cooled slowly.

EDTA (0.5*M*, pH 8.3)

Dissolved 18.61 g of EDTA (disodium, dihydrate) in 80 ml of distilled water by bringing the pH to 8.3 with NaOH solution and volume made upto 100 ml. Stored at room temperature after filtration and autoclaving.

Ethidium Bromide (10 mg/ml)

Dissolved 100 mg of Ethidium bromide in 10 ml of distilled water. Solution stored at 4 °C in a dark coloured bottle.

Formamide dye / Stop buffer

Deionised formamide	98 %
Xylene cyanol	0.025 %
Bromophenol blue	0.025 %
0.5 *M* EDTA	10m*M*

Gel loading buffer

Bromophenol blue	0.25 %	50 mg
Xylene cyanol	0.25 %	50 mg

Sucrose	40 %	8 g

Components were mixed and stirred well in 20 ml distilled water and stored at 4 °C.

LB agar/broth

Ingredients	Grams/Litre
Tryptone	10.0
Yeast extract	5.0
Sodium chloride	5.0
Agar	10.0

Final pH 7.2 at 37 °C. Sterilized by autoclaving. Stored prepared media below 8 °C, protected from direct light. Stored dehydrated powder, in a dry place, in tightly-sealed containers at 2-25 °C.
For LB agar, added 1.5 % agar and autoclaved.

LB Ampicillin media:

Dissolved 100 mg of ampicillin in 1 ml sterile water to get Ampicillin stock concentration of 100 mg/ml. Stored at 4 °C.

Added ampicillin to LB broth or agar at a final concentration of 100 μg/ml, when the temperature of the media is around 40-45 °C.

LB-Amp. Plates with X-Gal. and IPTG:

Added 40 μl each of X-Gal and IPTG for every 20 ml of LB-Amp agar at 40-45 °C. Mixed well and poured media into required number of plates.

Phenol (Saturated, pH 7.8)

Commercially available crystalline phenol was melted at 65 °C in a water bath. Hydroxyquinolone was added to a final concentration of 0.1 percent. To the molten phenol, 0.5 *M* Tris HCl (pH 8.0) was added in equal volume and stirred for 30 min in a magnetic stirrer. The contents were transferred into a separating funnel. Lower phenolic phase was collected and mixed with equal volume of 0.1 *M* Tris HCl (pH 8.0) and stirred again for 30 min. The phenolic phase was collected and extraction repeated with 0.1*M* Tris HCl (pH 8.0) until the pH of phenolic phase was more than 7.8. Finally 0.1 volume of 0.01 *M* Tris HCl (pH 8.0) added and stored in a dark bottle at 4 °C.

RBC lysis buffer

Ammonium chloride	150 mM	8.0235 g
Potassium chloride	10 mM	0.7455 g
EDTA	0.1 mM	0.0372 g

Dissolved the contents in distilled water and volume made up to 1000 ml. Stored at 4 °C after filtration and autoclaving.

Sodium acetate

Dissolved 40.824 g of Sodium acetate in 70 ml of distilled water and pH adjusted to 5.5 with glacial acetic acid. Volume was made up to 100 ml, autoclaved and stored at room temperature.

Sodium chloride (5 M)

Dissolved 29.22 g of sodium chloride in 80 ml distilled water and volume made up to 100 ml. Solution filtered and stored at room temperature.

Sodium chloride- EDTA (SE) buffer (pH, 8.0)

Sodium chloride	75 mM	4.383 g
EDTA	35 mM	9.306 g

Dissolved the contents in 900 ml distilled water and pH adjusted to 8.0. Made up the volume to 1000 ml, filtered, autoclaved and stored at 4 °C.

Tris Acetate EDTA (TAE) buffer (50X)

Tris base	48.4 g
Glacial acetic acid	11.42 ml
0.5 M EDTA (pH 8.0)	20 ml
Distilled water up to	1000 ml

Autoclaved and stored at room temperature.

Tris-Borate (TBE) buffer pH 8.3) 10X

Tris base	108.0 g
Boric acid	55.0 g
EDTA	9.3 g

Dissolved in 700 ml of distilled water and pH adjusted to 8.3. Volume was made up to 1000 ml, autoclaved and stored at room temperature.

Tris Buffered Saline (TBS) pH 7.4

Sodium chloride	140 mM	8.18 g
Potassium chloride	0.5 mM	0.0373 g
Tris base	0.25 mM	0.0303 g

Dissolved in 900 ml of distilled water and pH adjusted to 7.4. Made up the volume to 1000 ml, filtered, autoclaved and stored at 4 °C.

Tris EDTA (TE) buffer (pH 8.0)

Tris base	10 mM	1.2114 g
EDTA	0.1 mM	0.3722 g

Dissolved in 900 ml of distilled water and adjusted the pH to 8.0. Made up the volume to 1000 ml, filtered, autoclaved and stored at 4 °C.

Tris 1*M* (pH 8.0)

Tris base	121.14 g

Added distilled water up to 1000 ml, pH adjusted to 8.0, filtered and stored at room temperature.

ANNEXURE – II

SOURCES OF IMPORTANT CHEMICALS AND ENZYMES USED IN THIS STUDY

(A) CHEMICALS

Agar	- Himedia Laboratories,. Mumbai
Agarose (Low EED)	- Bangalore Genei Pvt. Ltd.
Ammonium chloride	- SRL, Bombay
Boric acid	- SRL, Bombay
Chloroform	- Merck
Crystalline phenol	- Merck
Di-sodium hydrogen orthophosphate	- SRL, Bombay
dNTPs.	- Finn Enzymes
EDTA	- SRL, Bombay
Ethanol	- Merck
Ethidum bromide	- BDH lab, England
6 X gel loading buffer	- Bangalore Genei Pvt. Ltd.
Glacial acetic acid	- BDH-E, Merck (India) Ltd.
Isoamyl alcohol	- Merck
Isopropyl alcohol	- SRL, Bombay
Potassium chloride	- SRL, Bombay
Sodium acetate	- SRL, Bombay
Sodium chloride	- SRL, Bombay

Tris base	- SRL, Bombay
Tryptone	- Himedia Laboratories, Mumbai
Yeast extract	- Himedia Laboratories, Mumbai

(B) PRIMERS

Integrated DNA Technologies, Inc.

1710 Commercial Park, Coralville, IA (www.idtdna.com)

(C) MOLECULAR MARKERS

500 bp DNA Ladder	- Bangalore Genei Pvt. Ltd.
1 kb DNA Ladder	- Bangalore Genei Pvt. Ltd.

(D) ENZYMES

Taq DNA polymerase	- Bangalore Genei Pvt. Ltd.
Proteinase-K	- Bangalore Genei Pvt. Ltd.
PNK	- Bangalore Genei Pvt. Ltd.

(E) KITS

DNA purification Kit	- MO BIO Laboratories, Inc. USA
Transformation kit	- Bangalore Genei Pvt. Ltd.
Cloning kit	- Bangalore Genei Pvt. Ltd.
Plasmid Miniprep kit	- Axygen Biosciences, USA

ANNEXURE – III

ABBREVIATIONS

A	Absorbance
BLAST	Basic Local Alignment Search Tool
BLASTn	Nucleotide BLAST
BLASTp	Protein BLAST
bp	base pair
$CaCl_2$	Calcium Chloride
cDNA	complementary Deoxyribo Nucleic acid
CDS	Coding Sequence
cm	centimeter
DNA	Deoxyribo Nucleic Acid
dNTP.	Deoxyribo Nucleotide Triphosphate
E. coli	*Escherichia coli*
EDTA	Ethylene Diamine Tetraacetic Acid
E-value	Expect value
ExPASy	Expert Protein Analysis System
FAO	Food and Agricultural Organization
g	gram
g	Relative centrifugal force
i.e.	that is
IPTG	Isopropyl β-D-1-thiogalactopyranoside.
K_m	Michaelis constant
kb	Kilo basepair
kDa	Kilo Dalton
kg	kilogram
L	litre
LB-Amp	Luria Bertani-Ampicillin
M	moles per litre
MCS	Multi Cloning Site

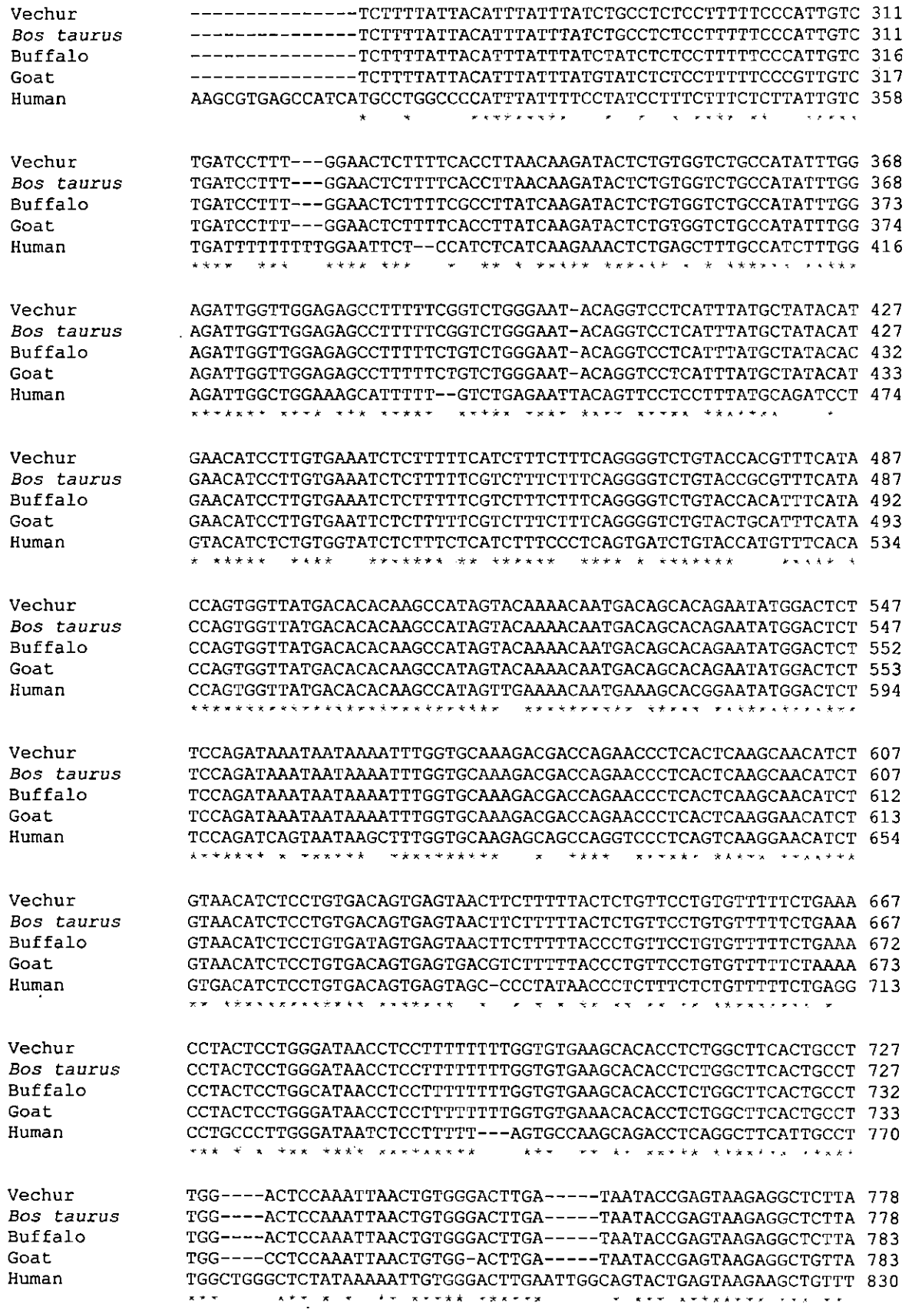
mg	milligram
$MgCl_2$	Magnesium chloride
m*M.*	millimolar
mRNA	messenger Ribo Nucleic Acid
NaCl	Sodium Chloride
NaOH	Sodium hydroxide
NCBI	National Center for Biotechnology Information
NH_4Cl	Ammonium Chloride
ng	nano gram
nm	nanometer
OD	Optical Density
PAPIA	Parallel Protein Information Analysis
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
p*M*	Pico moles
RFLP	Restriction Fragment Length Polymorphism
rpm	Revolutions per minute
RT	Room temperature
s	second
SDS	Sodium Dodecyl Sulphate
ss	single stranded
UV	ultra violet
WBC	white blood cell
X-gal	5-bromo-4-chloro-3-indolyl- beta-D- galactopyranoside
α-LA	alpha lactalbumin
κ-CN	kappa Casein
µg	microgram
µl	microlitre
µm	micrometer
°C	degree Celsius
3D	Three Dimensional

ANNEXURE - IV

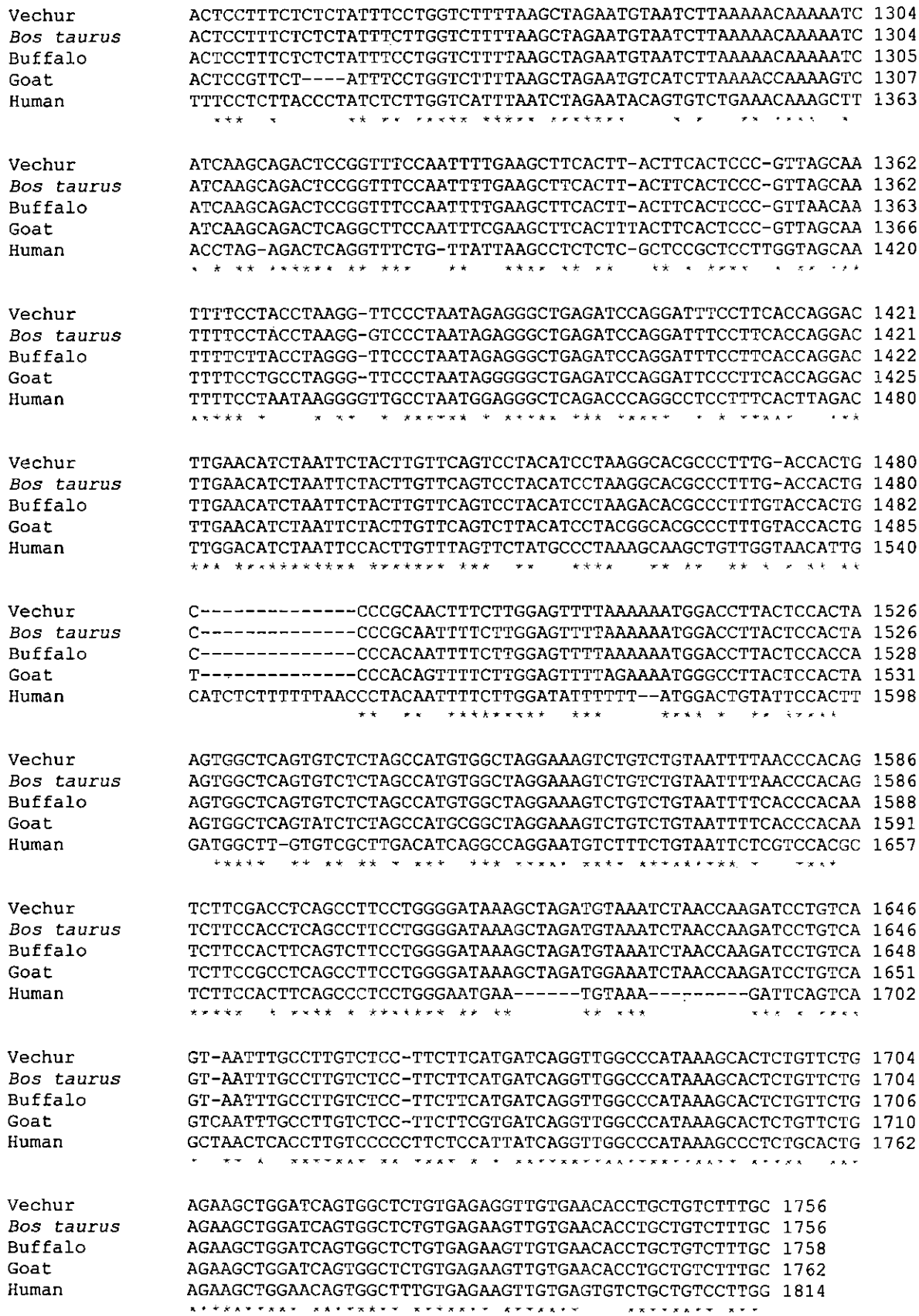
CLUSTAL W MULTIPLE NUCLEOTIDE SEQUENCE ALIGNMENT OF VECHUR ALPHA-LACTALBUMIN GENE WITH CORRESPONDING REGION OF ALPHA-LACTALBUMIN GENE OF LIVESTOCK SPECIES AND HUMAN

SeqA	Name	Len(nt)	SeqB	Name	Len(nt)	Score
1	Vechur	1756	2	*Bos taurus*	1756	99
1	Vechur	1756	3	Buffalo	1758	97
1	Vechur	1756	4	Goat	1762	95
1	Vechur	1756	5	Human	1814	63
2	*Bos taurus*	1756	3	Buffalo	1758	97
2	*Bos taurus*	1756	4	Goat	1762	95
2	*Bos taurus*	1756	5	Human	1814	63
3	Buffalo	1758	4	Goat	1762	96
3	Buffalo	1758	5	Human	1814	63
4	Goat	1762	5	Human	1814	62

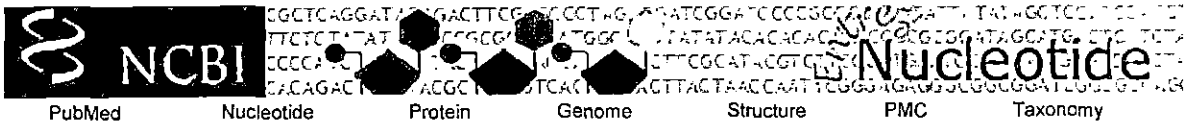
```
Vechur          GGGGTCACCAAAATG-ATGTCCTT-TGTTTCTCTGCTCCTGGTAGGCATCCTATTCCATG 58
Bos taurus      GGGGTCACCAAAATG-ATGTCCTT-TGTCTCTCTGCTCCTAGTAGGCATCCTATTCCATG 58
Buffalo         GGGGTAACCAAAATG-ATGTCCTT-TGTCTCTCTGCTCCTGGTAGGCATCCTATTCCATG 58
Goat            GGGGTAACCAAAATG-ATGTCCTT-TGTCTCTCTGCTCCTGGTAGGCATCCTATTCCACG 58
Human           TATTTCATGTATGTATATATGTATGTATTTATTTATTTTTGAGGTGGAGTTTCGCTCTTG 60

Vechur          CCACCCAGGCTGAACAGT--TAACAAAATGTGAGGTGTTCCGGGAGCTGAAAGAC----- 111
Bos taurus      CCACCCAGGCTGAACAGT--TAACAAAATGTGAGGTGTTCCGGGAGCTGAAAGAC----- 111
Buffalo         CCACCCAGGCTGAACAAT--TAACAAAATGTGAGGTGTTCCGGGAGCTGAAAGAC----- 111
Goat            CCACCCAGGCTGAACAAT--TAACAAAATGTGAGGTGTTCCAGAAGCTGAAGGAC----- 111
Human           TTGCCCAGACTGGAGTGCAATGGTGTAATCTC-GGCTCACTGCAACCTCCGCCTCCTCGG 119

Vechur          TTGAAGGGCTACGGAGGTGTCAGTTT--------GCCTGAATGTGAGTTCCCTGCTATTT 163
Bos taurus      TTGAAGGGCTACGGAGGTGTCAGTTT--------GCCTGAATGTGAGTTCCCTGCTATTT 163
Buffalo         TTGAAGGACTACGGAGGTGTCAGTTT--------GCCTGAATGTGAGTTCCCTGCTATTT 163
Goat            TTGAAGGACTACGGAGGTGTCAGTTT--------GCCTGAATGTGAGTTCCCTGCTATTT 163
Human           TTCAAGTGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGAATTACAGGCACCCACCACCA 179

Vechur          TGCTTTGTCCCATAATTCATCCTCTTCA----CTCTTTCCCTCCATTCTCTTCATCCTCT 219
Bos taurus      TGCTTTGTCCCATAATTCATCCTCTTCA----CTCTTTCCCTCCATTCTCTTCATCCTCT 219
Buffalo         TGCTTTGTCCCATAATTCATCCTCTTCA----CTCTTTCCCTCCATTCTCTTCATCCTCT 219
Goat            CGCTTTGTCCCATAATTCATCCTCTTCA----CTCTTTCCCTCCATTCTCTTCATCCTCT 219
Human           TGCCT-GGCTAATTTTGTATTTTTTGTAGAGACAGGGTTTCACCATGTTGGCCAGGCTGG 238

Vechur          TTTT-CCCCTCTACTTTTAATT-----ATCA--AACAATTCTCTTATTTGTT-----TAC 266
Bos taurus      TTTC-CCCCTCTACTTTTAATT-----ATCA--AACAATTCTCTTATTTGTT-----TAC 266
Buffalo         TTTT-CCCCTCTACTTTTAATTTAATTATCG--AACAATTCTCTTATTTGTT-----TAC 271
Goat            TTTTTCCCCTCTACTTTTAATTTAATTATCG--CACAATTCTCTTATTTGTT-----TAC 272
Human           TCTCAAACTTCTGACCTCAGGTGATCCGCCCTCCTCAGCCTCCCAAAGTGTTGGGATTAC 298
```


Vechur ----------TCTTTTATTACATTTATTTATCTGCCTCTCCTTTTTCCCATTGTC 311
Bos taurus ----------TCTTTTATTACATTTATTTATCTGCCTCTCCTTTTTCCCATTGTC 311
Buffalo ----------TCTTTTATTACATTTATTTATCTATCTCTCCTTTTTCCCATTGTC 316
Goat ----------TCTTTTATTACATTTATTTATGTATCTCTCCTTTTTCCCGTTGTC 317
Human AAGCGTGAGCCATCATGCCTGGCCCCATTTATTTTCCTATCCTTTCTTTCTCTTATTGTC 358
Vechur TGATCCTTT---GGAACTCTTTTCACCTTAACAAGATACTCTGTGGTCTGCCATATTTGG 368
Bos taurus TGATCCTTT---GGAACTCTTTTCACCTTAACAAGATACTCTGTGGTCTGCCATATTTGG 368
Buffalo TGATCCTTT---GGAACTCTTTTCGCCTTATCAAGATACTCTGTGGTCTGCCATATTTGG 373
Goat TGATCCTTT---GGAACTCTTTTCACCTTATCAAGATACTCTGTGGTCTGCCATATTTGG 374
Human TGATTTTTTTTTGGAATTCT--CCATCTCATCAAGAAACTCTGAGCTTTGCCATCTTTGG 416
Vechur AGATTGGTTGGAGAGCCTTTTTCGGTCTGGGAAT-ACAGGTCCTCATTTATGCTATACAT 427
Bos taurus AGATTGGTTGGAGAGCCTTTTTCGGTCTGGGAAT-ACAGGTCCTCATTTATGCTATACAT 427
Buffalo AGATTGGTTGGAGAGCCTTTTTCTGTCTGGGAAT-ACAGGTCCTCATTTATGCTATACAC 432
Goat AGATTGGTTGGAGAGCCTTTTTCTGTCTGGGAAT-ACAGGTCCTCATTTATGCTATACAT 433
Human AGATTGGCTGGAAAGCATTTTT--GTCTGAGAATTACAGTTCCTCCTTTATGCAGATCCT 474
Vechur GAACATCCTTGTGAAATCTCTTTTTCATCTTTCTTTCAGGGGTCTGTACCACGTTTCATA 487
Bos taurus GAACATCCTTGTGAAATCTCTTTTTCGTCTTTCTTTCAGGGGTCTGTACCGCGTTTCATA 487
Buffalo GAACATCCTTGTGAAATCTCTTTTTCGTCTTTCTTTCAGGGGTCTGTACCACATTTCATA 492
Goat GAACATCCTTGTGAATTCTCTTTTTCGTCTTTCTTTCAGGGGTCTGTACTGCATTTCATA 493
Human GTACATCTCTGTGGTATCTCTTTCTCATCTTTCCCTCAGTGATCTGTACCATGTTTCACA 534
Vechur CCAGTGGTTATGACACACAAGCCATAGTACAAAACAATGACAGCACAGAATATGGACTCT 547
Bos taurus CCAGTGGTTATGACACACAAGCCATAGTACAAAACAATGACAGCACAGAATATGGACTCT 547
Buffalo CCAGTGGTTATGACACACAAGCCATAGTACAAAACAATGACAGCACAGAATATGGACTCT 552
Goat CCAGTGGTTATGACACACAAGCCATAGTACAAAACAATGACAGCACAGAATATGGACTCT 553
Human CCAGTGGTTATGACACACAAGCCATAGTTGAAAACAATGAAAGCACGGAATATGGACTCT 594
Vechur TCCAGATAAATAATAAAATTTGGTGCAAAGACGACCAGAACCCTCACTCAAGCAACATCT 607
Bos taurus TCCAGATAAATAATAAAATTTGGTGCAAAGACGACCAGAACCCTCACTCAAGCAACATCT 607
Buffalo TCCAGATAAATAATAAAATTTGGTGCAAAGACGACCAGAACCCTCACTCAAGCAACATCT 612
Goat TCCAGATAAATAATAAAATTTGGTGCAAAGACGACCAGAACCCTCACTCAAGGAACATCT 613
Human TCCAGATCAGTAATAAGCTTTGGTGCAAGAGCAGCCAGGTCCCTCAGTCAAGGAACATCT 654
Vechur GTAACATCTCCTGTGACAGTGAGTAACTTCTTTTTACTCTGTTCCTGTGTTTTTCTGAAA 667
Bos taurus GTAACATCTCCTGTGACAGTGAGTAACTTCTTTTTACTCTGTTCCTGTGTTTTTCTGAAA 667
Buffalo GTAACATCTCCTGTGATAGTGAGTAACTTCTTTTTACCCTGTTCCTGTGTTTTTCTGAAA 672
Goat GTAACATCTCCTGTGACAGTGAGTGACGTCTTTTTACCCTGTTCCTGTGTTTTTCTAAAA 673
Human GTGACATCTCCTGTGACAGTGAGTAGC-CCCTATAACCCTCTTTCTCTGTTTTTCTGAGG 713
Vechur CCTACTCCTGGGATAACCTCCTTTTTTTTTGGTGTGAAGCACACCTCTGGCTTCACTGCCT 727
Bos taurus CCTACTCCTGGGATAACCTCCTTTTTTTTTGGTGTGAAGCACACCTCTGGCTTCACTGCCT 727
Buffalo CCTACTCCTGGCATAACCTCCTTTTTTTTTGGTGTGAAGCACACCTCTGGCTTCACTGCCT 732
Goat CCTACTCCTGGGATAACCTCCTTTTTTTTTGGTGTGAAACACACCTCTGGCTTCACTGCCT 733
Human CCTGCCCTTGGGATAATCTCCTTTTT---AGTGCCAAGCAGACCTCAGGCTTCATTGCCT 770
Vechur TGG----ACTCCAAATTAACTGTGGGACTTGA-----TAATACCGAGTAAGAGGCTCTTA 778
Bos taurus TGG----ACTCCAAATTAACTGTGGGACTTGA-----TAATACCGAGTAAGAGGCTCTTA 778
Buffalo TGG----ACTCCAAATTAACTGTGGGACTTGA-----TAATACCGAGTAAGAGGCTCTTA 783
Goat TGG----CCTCCAAATTAACTGTGG-ACTTGA-----TAATACCGAGTAAGAGGCTGTTA 783
Human TGGCTGGGCTCTATAAAAATTGTGGGACTTGAATTGGCAGTACTGAGTAAGAAGCTGTTT 830

```
Vechur          GAATTTTTCATTAACACTAAATCCCCAGACAGTTTCTTAAAGTTCCTGGGTAGGTGACCT 838
Bos taurus      GAATTTTTCATTAACACTAAATCCCCAGACAGTTTCTTAAAGTTCCTGGGTAGGTGACCT 838
Buffalo         GAATTTTTCATTAACACCAAATCCCCA----GTTTCTTAAAGTTCATGGGTAGGTGACCT 839
Goat            GAATTTTTCATTAACACCAAATCCCCAGACAGTTTCTTAAAGTTCATGGGTAGGTGACCT 843
Human           GGATTTTTCATGGTCATCAAATCCCCAGACAGTTCCTTGAGGTTCAGTGGTAGACAATCG 890
                * ********** **  *-*******     *** *** * **** *****   * *

Vechur          GAGCTGTTTGGGGATCTTGATGTATAATACCCTGTATTTTCAGACTAAGTTGGTTGATGA 898
Bos taurus      GAGCTGTTTGGGGATCTTGATGTATAATACCCTGTATTTTCAGACTAAGTTGGTTGATGA 898
Buffalo         GAGCTGTTTGGGGATCTTGATGTATAATACCCTGTATTTTCAGACTAAGTTGGTTGCTGA 899
Goat            GAGCTGTTTGGGGATCTTGATGTATAATACCCTGTATTTTCAGACTAAGTTGGTTGATGA 903
Human           GAGCTGTCTGAGAGTCTTGGAATCTGATTGTCTGCATTTTCAGGGTAAGTCAGTTGATGA 950
                ******* ** *  *****   * * **   *** ********  *****  **** ***

Vechur          AGTTGATAATTCCTAAGGAGCTGCCCCAGAGAAGAGAAGGGAGTCCTTACCTAGGGATAG 958
Bos taurus      AGTTGATAATTCCTAAGGAGCTGCCCCAGAGAAGAGAAGGGAGTCCTTACCTAGGGATAG 958
Buffalo         AGTTGATAATTCCTAAGGAGCTGCCCAAGAGAAGAGAAGGGAGTCCTTACCTAGGGATAG 959
Goat            AGTTGATAATTCCTAAGGAGCTGCCCCAGAGAAGAGAAGGGAGTCCTTACCTAGGGATAG 963
Human           AGCTGATGATTCCTCCAGAGATATCCCAGGGAAATGAAGGAAGTCCCTACCCAGGGTTAG 1010
                ** **** ******   *** *  ** ** ***  ***** ***** **** **** ***

Vechur          GCATTACTGTATTAAATTTCTCACCCAGAAGG-CAACAGGCATAAGCCTCTAGTTCAGAG 1017
Bos taurus      GCATTACTGTATTAAATTTCTCACCCAGAAGG-CAACAGGCATAAGCCTCTAGTTCAGAG 1017
Buffalo         GCATTACTATATTAAATTTCTCACCCAGAAGG-CAACAGGCATAAGCCTCTAGTTTAGAG 1018
Goat            GCATTACTGTATTAAATTTCTCATCTAAAAGG-CAACAGGTATAAGCCTCTAGTTCAGAG 1022
Human           ACATTACCACATTGGTCCTTTCATATAGAAAGACAACAGGCACAAGCCTTGAGTTTAGAG 1070
                *******   ***     * ***   * ** * ******** * ******  **** ****

Vechur          AA---------AACCAGAGAAGAGGGAAATTCATTATCCTTCTGGGTAATACTT---AGCT 1066
Bos taurus      AA---------AACCAGAGAAGAGGGAAATTCATTATCCTTCTGGGTAATACTT---AGCT 1066
Buffalo         AA---------ATCCAGAGAAGAGGGAAATTCATTACCCTTCTGGGTAATACTT---AGCT 1067
Goat            AA---------AACCAGAGAAGAGGGAAATTCATTACCCTTCTGGGTAATACTT---AGCT 1071
Human           AACCCACTGGATCCAGGGGTTAGGGGAACTCAGTGCCTTTCTGGGTAATACTTGTCAGCT 1130
                **         * **** *   **** ** *** *  * ********************   ****

Vechur          CTCTCATTTTTT--CCACCTATAACTCCTGCCAGAGTTCCTGGATGATGATCTTACTGAT 1124
Bos taurus      CTCTCATTTTTT--CCACCAGAGGCTCCTGCCAGAGTTCCTGGATGATGATCTTACTGAT 1124
Buffalo         CTCTCATTTTTT--CCACCTGTAACTCCTGCCAGAGTTCCTGGATGATGATCTTACTGAT 1125
Goat            CTCTCATTTTTTTCCCACCTGTAACTCCTGCCAGAGTTCCTGGATGATGATCTTACTGAT 1131
Human           GTCTCAATCCTT---TCCCTGTAACTCCTGCCAGAGTTCCTGGATGATGACATTACTGAT 1187
                *****  *   **      **      ********************************  ********

Vechur          GACATTATGTGTGTCAAGAAGATTCTGGATAAAGTAGGAATTAACTACTGGTGAGTCACC 1184
Bos taurus      GACATTATGTGTGTCAAGAAGATTCTGGATAAAGTAGGAATTAACTACTGGTGAGTCACC 1184
Buffalo         GACATTATGTGTGTCAAGAAGATTCTGGATAAAGTAGGAATTAACTACTGGTGAGTCACC 1185
Goat            GACATTGTGTGTGCCAAGAAGATTCTGGATAAAGTAGGAATTAACTACTGGTGAGTCACC 1191
Human           GACATAATGTGTGCCAAGAAGATCCTGGATATTAAAGGAATTGACTACTGGTGAATCCTT 1247
                *****  ****** ********** ********    ******** ************ **

Vechur          TCTCTATTTTTCACTTAATCTTTCCTCTCTTTCTTCTCAGTCCTTTCGTCCCAGCACTAT 1244
Bos taurus      TCTCTATTTTTCACTTAATCTTTCCTCTCTTTCTTCTCAGTCCTTTCGTCCCAGCACTAT 1244
Buffalo         TCTCTATTTTTTACTTAATCTTTCCCCTCTTTCTTCTCAGTCCTTTCGTCCCAGCACTAT 1245
Goat            TCTCTATTTTTCACTTAATCTTTCCCCTCTTTCTTCTCAGTCCTTTCGCCCCAGCACTAT 1251
Human           ATTCTATTTTCTATT---TCCCCATCCTCCTTCTCCTTACCCCATTAG-CCCAGCACCCC 1303
                ********  * *   **       *** **** ** *  ** ** * ********
```


Vechur ACTCCTTTCTCTCTATTTCCTGGTCTTTTAAGCTAGAATGTAATCTTAAAAACAAAAATC 1304
Bos taurus ACTCCTTTCTCTCTATTTCTTGGTCTTTTAAGCTAGAATGTAATCTTAAAAACAAAAATC 1304
Buffalo ACTCCTTTCTCTCTATTTCCTGGTCTTTTAAGCTAGAATGTAATCTTAAAAACAAAAATC 1305
Goat ACTCCGTTCT----ATTTCCTGGTCTTTTAAGCTAGAATGTCATCTTAAAACCAAAAGTC 1307
Human TTTCCTCTTACCCTATCTCTTGGTCATTTAATCTAGAATACAGTGTCTGAAACAAAGCTT 1363
Vechur ATCAAGCAGACTCCGGTTTCCAATTTTGAAGCTTCACTT-ACTTCACTCCC-GTTAGCAA 1362
Bos taurus ATCAAGCAGACTCCGGTTTCCAATTTTGAAGCTTCACTT-ACTTCACTCCC-GTTAGCAA 1362
Buffalo ATCAAGCAGACTCCGGTTTCCAATTTTGAAGCTTCACTT-ACTTCACTCCC-GTTAACAA 1363
Goat ATCAAGCAGACTCAGGCTTCCAATTTCGAAGCTTCACTTTACTTCACTCCC-GTTAGCAA 1366
Human ACCTAG-AGACTCAGGTTTCTG-TTATTAAGCCTCTCTC-GCTCCGCTCCTTGGTAGCAA 1420
Vechur TTTTCCTACCTAAGG-TTCCCTAATAGAGGGCTGAGATCCAGGATTTCCTTCACCAGGAC 1421
Bos taurus TTTTCCTACCTAAGG-GTCCCTAATAGAGGGCTGAGATCCAGGATTTCCTTCACCAGGAC 1421
Buffalo TTTTCTTACCTAGGG-TTCCCTAATAGAGGGCTGAGATCCAGGATTTCCTTCACCAGGAC 1422
Goat TTTTCCTGCCTAGGG-TTCCCTAATAGGGGGCTGAGATCCAGGATTCCCTTCACCAGGAC 1425
Human TTTTCCTAATAAGGGGTTGCCTAATGGAGGGCTCAGACCCAGGCCTCCTTTCACTTAGAC 1480
Vechur TTGAACATCTAATTCTACTTGTTCAGTCCTACATCCTAAGGCACGCCCTTTG-ACCACTG 1480
Bos taurus TTGAACATCTAATTCTACTTGTTCAGTCCTACATCCTAAGGCACGCCCTTTG-ACCACTG 1480
Buffalo TTGAACATCTAATTCTACTTGTTCAGTCCTACATCCTAAGACACGCCCTTTGTACCACTG 1482
Goat TTGAACATCTAATTCTACTTGTTCAGTCTTACATCCTACGGCACGCCCTTTGTACCACTG 1485
Human TTGGACATCTAATTCCACTTGTTTAGTTCTATGCCCTAAAGCAAGCTGTTGGTAACATTG 1540
Vechur C---------------CCCGCAACTTTCTTGGAGTTTTAAAAAATGGACCTTACTCCACTA 1526
Bos taurus C---------------CCCGCAATTTTCTTGGAGTTTTAAAAAATGGACCTTACTCCACTA 1526
Buffalo C---------------CCCACAATTTTCTTGGAGTTTTAAAAAATGGACCTTACTCCACCA 1528
Goat T---------------CCCACAGTTTTCTTGGAGTTTTAGAAAATGGGCCTTACTCCACTA 1531
Human CATCTCTTTTTTAACCCTACAATTTTCTTGGATATTTTTT--ATGGACTGTATTCCACTT 1598
Vechur AGTGGCTCAGTGTCTCTAGCCATGTGGCTAGGAAAGTCTGTCTGTAATTTTAACCCACAG 1586
Bos taurus AGTGGCTCAGTGTCTCTAGCCATGTGGCTAGGAAAGTCTGTCTGTAATTTTAACCCACAG 1586
Buffalo AGTGGCTCAGTGTCTCTAGCCATGTGGCTAGGAAAGTCTGTCTGTAATTTTCACCCACAA 1588
Goat AGTGGCTCAGTATCTCTAGCCATGCGGCTAGGAAAGTCTGTCTGTAATTTTCACCCACAA 1591
Human GATGGCTT-GTGTCGCTTGACATCAGGCCAGGAATGTCTTTCTGTAATTCTCGTCCACGC 1657
Vechur TCTTCGACCTCAGCCTTCCTGGGGATAAAGCTAGATGTAAATCTAACCAAGATCCTGTCA 1646
Bos taurus TCTTCCACCTCAGCCTTCCTGGGGATAAAGCTAGATGTAAATCTAACCAAGATCCTGTCA 1646
Buffalo TCTTCCACTTCAGTCTTCCTGGGGATAAAGCTAGATGTAAATCTAACCAAGATCCTGTCA 1648
Goat TCTTCCGCCTCAGCCTTCCTGGGGATAAAGCTAGATGGAAATCTAACCAAGATCCTGTCA 1651
Human TCTTCCACTTCAGCCCTCCTGGGAATGAA------TGTAAA---------GATTCAGTCA 1702
Vechur GT-AATTTGCCTTGTCTCC-TTCTTCATGATCAGGTTGGCCCATAAAGCACTCTGTTCTG 1704
Bos taurus GT-AATTTGCCTTGTCTCC-TTCTTCATGATCAGGTTGGCCCATAAAGCACTCTGTTCTG 1704
Buffalo GT-AATTTGCCTTGTCTCC-TTCTTCATGATCAGGTTGGCCCATAAAGCACTCTGTTCTG 1706
Goat GTCAATTTGCCTTGTCTCC-TTCTTCGTGATCAGGTTGGCCCATAAAGCACTCTGTTCTG 1710
Human GCTAACTCACCTTGTCCCCCTTCTCCATTATCAGGTTGGCCCATAAAGCCCTCTGCACTG 1762
Vechur AGAAGCTGGATCAGTGGCTCTGTGAGAGGTTGTGAACACCTGCTGTCTTTGC 1756
Bos taurus AGAAGCTGGATCAGTGGCTCTGTGAGAAGTTGTGAACACCTGCTGTCTTTGC 1756
Buffalo AGAAGCTGGATCAGTGGCTCTGTGAGAAGTTGTGAACACCTGCTGTCTTTGC 1758
Goat AGAAGCTGGATCAGTGGCTCTGTGAGAAGTTGTGAACACCTGCTGTCTTTGC 1762
Human AGAAGCTGGAACAGTGGCTTTGTGAGAAGTTGTGAGTGTCTGCTGTCCTTGG 1814

ANNEXURE – V



☐ **1:** EU200932. Reports ...[gi:158702916]

Features Sequence

```
LOCUS       EU200932                1756 bp    DNA     linear   MAM 30-OCT-2007
DEFINITION  Bos indicus breed Vechur alpha lactalbumin precursor, gene,
            complete cds.
ACCESSION   EU200932
VERSION     EU200932.1  GI:158702916
KEYWORDS    .
SOURCE      Bos indicus (Bos taurus indicus)
  ORGANISM  Bos indicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
            Pecora; Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 1756)
  AUTHORS   Rajeev,M. and Aravindakshan,T.V.
  TITLE     Molecular characterization of alpha lactalbumin gene in Vechur
            cattle
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1756)
  AUTHORS   Rajeev,M. and Aravindakshan,T.V.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-OCT-2007) Centre for Advanced Studies in Animal
            Genetics & Breeding, Kerala Agricultural University, Mannuthy,
            Thrissur, Kerala 680 651, India
FEATURES             Location/Qualifiers
     source          1..1756
                     /organism="Bos indicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9915"
                     /chromosome="5"
                     /map="5q21"
                     /country="India: Kerala"
                     /note="breed: Vechur"
     mRNA            join(<13..145,467..625,1099..1174,1679..>1739)
                     /product="alpha lactalbumin precursor"
     CDS             join(13..145,467..625,1099..1174,1679..1739)
                     /codon_start=1
                     /product="alpha lactalbumin precursor"
```

```
                     /protein_id="ABW77760.1"
                     /db_xref="GI:158702917"
                     /translation="MMSFVSLLLVGILFHATQAEQLTKCEVFRELKDLKGYGGVSLPE
                     WVCTTFHTSGYDTQAIVQNNDSTEYGLFQINNKIWCKDDQNPHSSNICNISCDKFLDD
                     DLTDDIMCVKKILDKVGINYWLAHKALCSEKLDQWLCERL"
     sig_peptide     13..69
     exon            <13..145
                     /number=1
     misc_feature    17..22
                     /note="potential glucocorticoid site"
     exon            467..625
                     /number=2
     exon            1099..1174
                     /number=3
     exon            1679..>1739
                     /number=4
     misc_feature    1698..1703
                     /note="potential glucocorticoid site"
ORIGIN
        1 ggggtcacca aaatgatgtc ctttgtttct ctgctcctgg taggcatcct attccatgcc
       61 acccaggctg aacagttaac aaaatgtgag gtgttccggg agctgaaaga cttgaagggc
      121 tacggaggtg tcagtttgcc tgaatgtgag ttccctgcta ttttgctttg tcccataatt
      181 catcctcttc actctttccc tccattctct tcatcctctt tttccctct actttaatt
      241 atcaaacaat tctcttattt gtttactctt ttattacatt tatttatctg cctctccttt
      301 ttcccattgt ctgatccttt ggaactcttt tcaccttaac aagatactct gtggtctgcc
      361 atatttggag attggttgga gagccttttt cggtctggga atacaggtcc tcatttatgc
      421 tatacatgaa catccttgtg aaatctcttt ttcatctttc tttcaggggt ctgtaccacg
      481 tttcatacca gtggttatga cacacaagcc atagtacaaa acaatgacag cacagaatat
      541 ggactcttcc agataaataa taaaatttgg tgcaaagacg accagaaccc tcactcaagc
      601 aacatctgta acatctcctg tgacagtgag taacttcttt ttactctgtt cctgtgtttt
      661 tctgaaacct actcctggga taacctcctt ttttttggtg tgaagcacac ctctggcttc
      721 actgccttgg actccaaatt aactgtggga cttgataata ccgagtaaga ggctcttaga
      781 atttttcatt aacactaaat ccccagacag tttcttaaag ttcctgggta ggtgacctga
      841 gctgtttggg gatcttgatg tataataccc tgtattttca gactaagttg gttgatgaag
      901 ttgataattc ctaaggagct gccccagaga agagaaggga gtccttacct agggataggc
      961 attactgtat taaatttctc acccagaagg caacaggcat aagcctctag ttcagagaaa
     1021 accagagaag agggaaattc attatccttc tgggtaatac ttagctctct catttttcc
     1081 acctataact cctgccagag ttcctggatg atgatcttac tgatgacatt atgtgtgtca
     1141 agaagattct ggataaagta ggaattaact actggtgagt cacctctcta ttttcactt
     1201 aatctttcct ctctttcttc tcagtccttt cgtcccagca ctatactcct ttctctctat
     1261 ttcctggtct tttaagctag aatgtaatct taaaaacaaa aatcatcaag cagactccgg
     1321 tttccaattt tgaagcttca cttacttcac tcccgttagc aattttccta cctaaggttc
     1381 cctaatagag ggctgagatc caggatttcc ttcaccagga cttgaacatc taattctact
     1441 tgttcagtcc tacatcctaa ggcacgccct ttgaccactg ccccgcaact ttcttggagt
     1501 tttaaaaaat ggaccttact ccactaagtg gctcagtgtc tctagccatg tggctaggaa
     1561 agtctgtctg taattttaac ccacagtctt cgacctcagc cttcctgggg ataaagctag
     1621 atgtaaatct aaccaagatc ctgtcagtaa tttgccttgt ctccttcttc atgatcaggt
     1681 tggcccataa agcactctgt tctgagaagc tggatcagtg gctctgtgag aggttgtgaa
     1741 cacctgctgt ctttgc
//
```

ANNEXURE -VI

ABBREVIATIONS AND TRIPLET CODONS OF AMINO ACIDS

Amino Acid	Abbreviation		Codons
	3 letter	Single Letter	
Isoleucine	Ile	I	ATT, ATC, ATA
Leucine	Leu	L	CTT, CTC, CTA, CTG, TTA, TTG
Valine	Val	V	GTT, GTC, GTA, GTG
Phenylalanine	Phe	F	TTT, TTC
Methionine	Met	M	ATG
Cysteine	Cys	C	TGT, TGC
Alanine	Ala	A	GCT, GCC, GCA, GCG
Glycine	Gly	G	GGT, GGC, GGA, GGG
Proline	Pro	P	CCT, CCC, CCA, CCG
Threonine	Thr	T	ACT, ACC, ACA, ACG
Serine	Ser	S	TCT, TCC, TCA, TCG, AGT, AGC
Tyrosine	Tyr	Y	TAT, TAC
Tryptophan	Trp	W	TGG
Glutamine	Gln	Q	CAA, CAG
Asparagine	Asn	N	AAT, AAC
Histidine	His	H	CAT, CAC
Glutamic acid	Glu	E	GAA, GAG
Aspartic acid	Asp	D	GAT, GAC
Lysine	Lys	K	AAA, AAG
Arginine	Arg	R	CGT, CGC, CGA, CGG, AGA, AGG
Termination	Ter	X	TAA, TAG, TGA

ANNEXURE - VII

CLUSTAL W MULTIPLE PROTEIN SEQUENCE ALIGNMENT OF VECHUR ALPHA LACTALBUMIN WITH OTHER SPECIES

```
Mouse          MMHFVPLFLVCILSLPAFQATELTKCKVSHAIKDIDGYQGISLLEWACVLFHTSGYDTQA 60
Rat            MMRFVPLFLACIS-LPAFQATEFTKCEVSHAIEDMDGYQGISLLEWTCVLFHTSGYDSQA 59
Donkey         --------------------KQFTKCELSQVLKSMDGYKGVTLPEWICTIFHSSGYDTQT 40
Horse          --------------------KQFTKCELSQVLKSMDGYKGVTLPEWICTIFHSSGYDTQT 40
Pig            MMSFVSLLVVGIL-FPAIQAKQFTKCELSQVLKDMDGYGDITLPEWICTIFHISGYDTKT 59
Dog            MMSFVSLLLVSIL-FPAIQAKQFTKCELPQVLKDMDGFGGIALPEWICTIFHTSGYDTQT 59
Bos            MMSFVSLLLVGIL-FHATQAEQLTKCEVFRELKDLKGYGGVSLPEWVCTTFHTSGYDTQA 59
Yak            MMSFVSLLLVGIL-FHATQAEQLTKCEVFRELKDLKGYGGVSLPEWVCTTFHTSGYDTQA 59
Vechur         MMSFVSLLLVGIL-FHATQAEQLTKCEVFRELKDLKGYGGVSLPEWVCTTFHTSGYDTQA 59
Buffalo        MMSFVSLLLVGIL-FHATQAEQLTKCEVFRELKDLKDYGGVSLPEWVCTTFHTSGYDTQA 59
Sheep          MMSFVSLLLVGIL-FHATQAEQLTKCEVFQELKDLKDYGGVSLPEWVCTAFHTSGYDTQA 59
Goat           MMSFVSLLLVGIL-FHATQAEQLTKCEVFQKLKDLKDYGGVSLPEWVCTAFHTSGYDTQA 59
Human          MRFFVPLFLVGIL-FPAILAKQFTKCELSQLLKDIDGYGGIALPELICTMFHTSGYDTQA 59
Chimpanzee     MRFFVPLFLVGIL-FPAILAKQFTKCELSQLLKDIDGYGGIALPELICTMFHTSGYDTQA 59
Rhesus         MRSFVPLFLVGIL-FPAIPAKQFTKCELSQLLKDIDGYGGIALPEFICTMFHTSGYDTQA 59
Guinea         MMSFFPLLLVGIL-FPAVQAKQLTKCALSHELNDLAGYRDITLPEWLCIIFHISGYDTQA 59
Rabbit         MMPLVPLLLVSIV-FPGIQATQLTRCELTEKLKELDGYRDISMSEWICTLFHTSGLDTKI 59
                                    ::*:* : . ::.: .: .::: *  *  ** ** *::

Mouse          VVNDNGSTEYGLFQISDRFWCKSSEFPESENICGISCDKLLDDELDDDIACAKKILAIKG 120
Rat            IVKNNGSTEYGLFQISNRNWCKSSEFPESENICDISCDKFLDDELADDIVCAKKIVAIKG 119
Donkey         IVKNNGKTEYGLFQINNKMWCRDNQILPSRNICGISCNKFLDDDLTDDVMCAKKILDSEG 100
Horse          IVKNNGKTEYGLFQINNKMWCRDNQILPSRNICGISCDKFLDDDLTDDVMCAKKILDSEG 100
Pig            IVHDNGSTEYGLFQINNKLWCRDNQIQ-SKNICGISCDKFLDDDLTDDMMCAKKILDNEG 118
Dog            IVNNNGGTDYGLFQISNKFWCKDDQNLQSRNICDISCDKFLDDDLTDDMICAKKILDKEG 119
Bos            IVQNNDSTEYGLFQINNKIWCKDDQNPHSSNICNISCDKFLDDDLTDDIMCVKKILDKVG 119
Yak            IVQNNDSTEYGLFQINNKIWCKDDQNPHSSDICNISCDKFLDDDLTDDIMCVKKILDKVG 119
Vechur         IVQNNDSTEYGLFQINNKIWCKDDQNPHSSNICNISCDKFLDDDLTDDIMCVKKILDKVG 119
Buffalo        IVQNNDSTEYGLFQINNKIWCKDDQNPHSSNICNISCDKFLDDDLTDDIMCVKKILDKVG 119
Sheep          IVQNNDSTEYGLFQINNKIWCKDDQNPHSRNICNISCDKFLDDDLTDDIMCVKKILDKVG 119
Goat           IVQNNDSTEYGLFQINNKIWCKDDQNPHSRNICNISCDKFLDDDLTDDIVCAKKILDKVG 119
Human          IVENNESTEYGLFQISNKLWCKSSQVPQSRNICDISCDKFLDDDITDDIMCAKKILDIKG 119
Chimpanzee     IVENNESTEYGLFQISNKLWCKSSQVPHSRNICDISCDKFLDDDITDDIMCAKKILDIKG 119
Rhesus         IVESNGSTEYGLFQISNKLWCKSSQVPQSRNICDISCDKFLDDDITDDIMCAKKILDIKG 119
Guinea         IVKNSDHKEYGLFQINDKDFCESSTTVQSRNICDISCDKLLDDDLTDDIMCVKKILDIKG 119
Rabbit         TVNNNGSTEYGIFQISDKLWCVSKQNPQSKNICDTPCENFLDDNLTDDVKCAMKILDKEG 119
                *...  .:**:***.:: :* ..    * :**. .*:::***:: **: *. **:   *

Mouse          IDYWKAYKPMCSEKLEQWRCEKP------------------ 143
Rat            IDYWKAHKPMCSEKLEQWRCEKPGAPALVVPALNSETPVP 159
Donkey         IDYWLAHKPLCSEKLEQWLCEEL------------------ 123
Horse          IDYWLAHKPLCSEKLEQWLCEEL------------------ 123
Pig            IDYWLAHKALCSEKLDQWLCEKM------------------ 141
Dog            IDYWLAHKPLCSEKLEQWRCEKL------------------ 142
Bos            INYWLAHKALCSEKLDQWLCEKL------------------ 142
Yak            INYWLAHKALCSEKLDQWLCEKL------------------ 142
Vechur         INYWLAHKALCSEKLDQWLCERL------------------ 142
Buffalo        INYWLAHKALCSEKLDQWLCEKL------------------ 142
Sheep          INYWLAHKALCSEKLDQWLCEKL------------------ 142
Goat           INYWLAHKALCSEKLDQWLCEKL------------------ 142
Human          IDYWLAHKALCTEKLEQWLCEKL------------------ 142
Chimpanzee     IDYW------------------------------------- 123
Rhesus         IDYWLAHKALCTEKLEQWLCEKL------------------ 142
Guinea         IDYWLAHKPLCSDKLEQWYCEAQ------------------ 142
Rabbit         IDHWLAHKPLCSENLEQWVCKK------------------- 141
               *::*
```

SeqA	Name	Len(aa)	SeqB	Name	Len(aa)	Score
1	Vechur	142	2	Buffalo	142	98
1	Vechur	142	3	*Bos taurus*	142	99
1	Vechur	142	4	Goat	142	94
1	Vechur	142	5	Human	142	72

MOLECULAR CLONING AND CHARACTERIZATION OF ALPHA LACTALBUMIN GENE IN VECHUR CATTLE

RAJEEV. M.

Abstract of the thesis submitted in partial fulfilment of the requirement for the degree of

Master of Veterinary Science

Faculty of Veterinary and Animal Sciences
Kerala Agricultural University, Thrissur

2007

Department of Animal Breeding and Genetics
COLLEGE OF VETERINARY AND ANIMAL SCIENCES
MANNUTHY, THRISSUR-680651
KERALA, INDIA

ABSTRACT

The study was undertaken with the objectives of cloning and characterization of the gene encoding the milk protein alpha-lactalbumin (α-LA) of Vechur cow of Kerala. The α-LA is a mammary gland specific protein found in high concentrations in milk of many species and has a role in regulating lactose synthase. Alpha-lactalbumin is homologous with the c-type lysozymes and provides an example of extreme functional divergence in homologous proteins with closely similar structures.

The genomic DNA was isolated from blood samples and a 1756 bp fragment of the entire transcriptional unit of the α-LA was amplified by PCR using synthetic oligonucleotide primer pair designed based on the bovine α-LA gene sequence. The gel purified PCR product was ligated in to a T-vector and was transformed by giving heat shock to competent *E. coli* cells prepared by $CaCl_2$ treatment. The recombinant clones among the transformed cells were identified by *Blue–White Screening* and the recombinant plasmid carrying the insert gene was isolated from the white clones by a modified SDS-alkaline lysis method. The 1.756 kb α-LA gene insert in the vector was sequenced by the dideoxynucleotide sequencing method with primer walking using an automated DNA sequencer.

The sequence was found to be having 99 per cent homology with that of *Bos taurus*, 98 per cent with that of Yak and 95 per cent with that of sheep α-LA gene. Comparison of the human and Vechur α-LA genes showed identical structural organization and identifies extensive homology within the transcription unit. The exon-intron boundaries in the human α-LA gene occur at the codons of the amino acid residues, Leu-26 (intron 1), Lys-79 (intron 2) and Trp-104 (intron 3). The latter two residues are conserved in both species and also in *Bos taurus*, whereas Leu-26 is replaced by Trp-26 in the Vechur and bovine proteins. This strict homology in the sites of insertion of introns suggests that the exon-intron organization of these genes was established before the divergence of these species. The positions of the exon-

intron boundaries are also conserved as evidenced from similar sizes of the exons. The introns sizes are also comparable except in the case of intron 1, which is much larger in the human gene as a consequence of the insertion of a *Alu* family repeat sequence.

The Vechur α-LA gene has an open reading frame of 426 nucleotides encoding a signal peptide of 19 amino acid residues and a mature protein of 123 amino acid residues with NH_2 terminal glutamic acid and COOH- terminal leucine. Alignment of this sequence with bovine counterpart showed that 122 amino acid residues are identical and with human α-LA sequence showed 73 per cent identity.

The predicted secondary structure of Vechur α-LA showed that the larger α-helical lobe is formed by the amino- and carboxyl-terminal sections of the polypeptide chain while the smaller lobe, which encompasses a small three stranded antiparallel beta-sheet, and a small irregular structure, is formed by the central section of the polypeptide chain. The predicted tertiary structure of Vechur α-LA also showed high homology with the bovine and human α-LA structures.

Overall, the structures of Vechur α-LA was found to be very similar to that of *Bos taurus* and human reflecting their high degree of amino acid sequence identity. The present study did not reveal any higher degree of structural or functional similarity between Vechur and human α-LA proteins as compared to that of *Bos taurus*. The superiority of human milk and its high suitability to infants could be due to the higher content of α-LA and might not be attributed to any structural variations of the protein. Since the higher content of α-LA in human milk could be due to the high expression of this gene, further studies may be carried out to find out sequence variations, if any, occur in the regulatory sequences upstream of the gene. Gene expression studies are suggestive as α-LA locus can also be used as a genetic marker to increase milk production in Vechur cattle, as this marker may be directly responsible for increased milk production.