Swedish University of Agricultural Sciences
Faculty of Veterinary Medicine and Animal Science

Bioinformatics and molecular genetic studies of domestic and wild buffalo species: Focus on evolutionary relationship of the *DGAT1* gene

Farzad Hakim Javadi

Examensarbete / Swedish University of Agricultural Sciences
Department of Animal Breeding and Genetics
427
Uppsala 2013

Master Thesis, 30 hp

Master's Programme
– Animal Science

Swedish University of Agricultural Sciences
Faculty of Veterinary Medicine and Animal Science
Department of Animal Breeding and Genetics

Bioinformatics and molecular genetic studies of domestic and wild buffalo species: Focus on evolutionary relationship of the *DGAT1* gene

Farzad Hakim Javadi

Supervisors:
Göran Andersson, SLU, Department of Animal Breeding and Genetics
Erik Bongcam-Rudloff, SLU, Department of Animal Breeding and Genetics

Examiner:
Gabriella Lindgren, SLU, Department of Animal Breeding and Genetics

Credits: 30 hp
Course title: Degree project in Animal Science
Course code: EX0556
Programme: Master's Programme - Animal Science
Level: Advanced, A2E

Place of publication: Uppsala
Year of publication: 2013
Name of series: Examensarbete / Swedish University of Agricultural Sciences, Department of Animal Breeding and Genetics, 427
On-line publicering: http://epsilon.slu.se

Key words: Buffalo, *DGAT1*, promoter, Swamp, River

Abstract

Improving milk yield and production is the main purpose of the selection in dairy industry. Genotyping tools give the opportunity to increase the frequency of favorable traits in the herd. Efficient breeding programs used in dairy industry have resulted in extensive genetic progress for many economically important traits. Progeny testing helps to choose the best young males based on their genetic merit and therefore their potential for being used in breeding. During the last few years statistical tools have used the massive genetic information available for selecting the young animals with high breeding values in the early ages which is referred to genomic selection. One such trait that has been selected for is milk yield and composition that is controlled by a Quantitative trait locus (QTL) in the centromeric region of cattle (*Bos taurus*) chromosome 14. This SNP is not conserved between different cattle breeds and buffalo. A previously reported amino acid substitution in bovine diacylglycerol acyl transferase 1 (*DGAT1*) gene, K232A (lysine to alanine substitution) has been correlated with milk quality traits (Grisart et al., 2004) (Figure 16). This polymorphism is localized in exon VIII of the *DGAT1* gene. DGAT is involved in the triglyceride biosynthesis and catalyzes the conversion of complex carbohydrates to triglycerides and therefore controls the lipid content in milk. Restriction fragment length polymorphism (RFLP) identified three *DGAT1* genotypes (AA, AG, GG). Genotype AA is associated with higher milk fat content. Basically, dairy industry was more interested to improve milk yield and protein content of the milk, therefore cow herd should have higher proportion of G allele. Water Buffalo (*Bubalus bubalis)* populations have not been subjected to that selection pressure since the aim of buffalo production is more milk fat content.

The aim of this project was to examine the potential association of *DGAT1* genotypes in different buffalo populations. Comparison was performed between Asian subtypes (Swamp and River), and River buffaloes that were imported to Sweden from Italy. Obtained sequences were compared to those present in the publically available cattle genome sequence. The identification of the promoter region of the *DGAT1* gene was done by bioinformatics approach to search for any significant differences within and between populations. Our result confirms that same SNP is available in the Swamp Buffalo population, but because of the lack of recording we have not been able to correlate genotypes with significant differences in milk fat content. Different length of promoter was observed when comparing cattle genome with human. A highly conserved region located near the DGAT1 translation start site was found using the 29 mammals' database, which suggest that this region is important for proper translation of *DGAT1* mRNA. The result of this experiment identified one SNP positioned at 857 bp 5' of the translation initiation codon (ATG) in a potential transcriptional regulatory region. The only difference is that our River buffalo population had different SNP on that position T instead of G in the cattle reference sequence (Figure 14 and 15). The other region up or down stream of this region might be the potential region regulating transcription, which increase the importance of the SNP position to function as an regulatory region for *DGAT1* transcription or the genes that surrounds the *DGAT1* gene. Further functional studies can confirm our hypothesis about transcription binding site and length of promoter in the genera bovina and bubalina.

1 Contents

2 Introduction

2.1 World water buffalo population

The global water buffalo population, which was estimated to be 177.25 million heads in 2007 are distributed in 42 different countries. From 1961 (88.2 million) to 2007 buffalo populations shows 100% increased (FAO, 2009). 96.96 % of buffalo population (171.86) is in Asia and the rest are spread in Africa, predominantly in Egypt (3.98), South America (1.14 million), Europe (0.27 million) and around 100,000 in Australia (Table 1). (Pasha T.N & Hayat. Z, 2012).

Table1. World water Buffalo distribution 1961-2007 (million heads)

Region	1961	1971	1981	1991	2001	2007
Asia	86.02	105.86	120.74	145.21	161.47	171.86
Africa	1.50	2.06	2.37	2.99	3.53	3.98
South America	0.070	0.14	0.55	1.44	1.12	1.14
Europe	0.73	0.66	0.55	0.57	0.23	0.27

Data from Australia is not available.

(FAO.org, 2009)

(Pasha T.N & Hayat. Z, 2012)

2.1.1 Asia

Asian buffalo or water buffalo is classified into the genus *Bubalus*. Species is *bubalis*. The *Bubalus bubalis* belongs to the Mammalian class, Subclass *Ungulata*, *Artiodactyla* order, *Ruminantia* suborder, *Bovine* family, *Bovinae* subfamily. *Bovini* subfamily includes the genera: Bovina (Cattle), *Bubalina* and *Syncerina*. *Syncerina* just has one species which is African buffalo *Syncerus caffer*. Asian buffalo (*Bubalina*) consists of: *Bubalus depressicornis* which are in Indonesia and also called Anoa. *Bubalus mindorensis* lives in the Philippines and *Bubalus bubalis* which is Indian wild buffalo. Indian buffalo results from *Bubalus arnee* domestication (Figure 1). Compared to *Bos taurus* and *Bos indicus* (10,000 years ago), *Bubalus bubalis* was domesticated approximately 5,000 years ago (Di Berardino and Iannuzzi, 1981).

River and Swamp buffalo are two subspecies of Asian buffalo (*Bubalus bubalis*). They are both genetically and morphologically different. The River buffalo has 50 chromosomes which include 5 pairs of submetacentric and 20 pairs are acrocentric (Di Berardino and Iannuzzi, 1981). The Swamp buffalo has 48 chromosomes of which 19 pairs are metacentric. Swamp buffalo chromosome 1, which is the largest metacentric chromosome originated from a tandem fusion translocation between the River buffalo chromosome 9 (centromere) and 4 (telomeres of p-arm). This caused the nucleolus organizer regions (NORs) which are located on chromosome 4p in the River buffalo to be lost and the inactivation of the centromere in chromosome 9 (Di Berardino and Iannuzzi, 1981). Both subspecies are inter-fertile and their progeny has 49 chromosomes. If female progenies are backcrossed they will have longer calving intervals. Male crossbred sometimes shows fertility problem (Di Berardino and Iannuzzi, 1981).

Morphologically, Swamp buffalo has smaller body size compared with River buffalo. Male Swamp buffaloes weigh between 325 and 450 Kg compared to River buffalo, which has body weights ranging between 450 and 1000 Kg. Even though Swamp type buffaloes are reared for draught purpose still they produce considerable amount of milk (600 Kg/ year). River type is mainly found in India and Pakistan

which is characterized by curled horn and bigger body size. They produce both high quality and high quantity milk (2200–2400 kg/year). The name River buffalo came from the fact that these animals prefer to enter in the clean water like rivers. They are mainly used for milk and meat production. Swamp buffaloes are mainly used for draught in swamp and haulage fields. Each subspecies consist of many breeds based on the location they come from (Sethi, 2003).

A variety of different production patterns were seen in terms of milk and meat production in different Asian countries in the last decades. Water buffalo populations in India, Pakistan and China have all shown decrease in milk yield per animal by 2.44 percent, 1 percent and 1.55 percent, respectively. Bangladesh, Myanmar, Nepal and Vietnam had no change or insignificant change in milk yield per animal. South and south-east Asian countries were shown negative growth in this concern. Pakistan shows increase in total buffalo milk production by 1.43 percent at the same time there was no changes in the other countries. In total, the average increase in buffalo milk production in Asian countries is 2.26 percent (Dhanda, 2004).

Buffalo can convert low-quality roughage into milk and meat. Mudgal 1988, showed that buffaloes compared to high-productive cows have a 5 percent higher crude fiber digestibility. They utilized metabolic energy for milk production by 4-5 percent more efficient compare to cows.

2.1.1.1 India

98.70 million Heads of buffaloes are present in India and this is the country with the largest amount of buffalo milk produced by 63 million tons per year (FAO, 2009) (Table 2).

Table2. Buffalo milk production worldwide (10^3 kg)

Country Name	2009	2008	2007	2006	2005	2004	2003
India	62,860,000	60,900,000	55,913,000	54,382,000	52,070,000	50,178,000	47,979,000
Pakistan	21,622,000	20,971,000	20,372,000	19,779,000	19,884,000	19,240,000	18,617,000
China	3,000,000	2,950,000	2,900,000	2,850,000	2,800,000	3,523,200	2,750,000
Egypt	2,702,960	2,640,640	2,609,820	1,840,000	2,300,000	2,266,770	2,549,580
Nepal	1,031,500	987,780	958,603	926,850	894,591	863,322	834,376
Iran	279,054	270,403	252,033	245,000	232,432	217,560	208,428
Myanmar	240,000	238,704	220,462	198,025	170,600	162,792	150,459
Italy	210,000	216,780	234,999	221,069	215,228	183,881	171,870
Sri Lanka	41,600	30,105	27,260	26,730	26,120	25,840	25,560
Bangladesh	35,088	34,000	32,000	31,000	30,000	28,000	27,000
Viet Nam	33,198	32,000	33,000	39,001	32,000	32,174	31,558
Turkey	32,443	31,422	30,375	36,358	38,058	39,279	48,778
Iraq	23,000	24,572	23,695	40,000	65,000	95,060	96,030
Malaysia	10,659	9,041	8,350	7,500	8,101	7,935	7,475
Bulgaria	7,022	7,173	7,052	7,132	6,989	6,229	5,276
Georgia	5,400	6,300	6,100	5,600	NA	NA	NA
Syria	3,600	3,517	3,098	4,000	5,000	2,000	1,520
Bhutan	389	369	339	330	345	334	323
Greece	158	160	156	123	90	131	55

Brunei Darussalam	65	55	51	46	53	52	51
Albania	10	10	6	7	7	7	7
Total	92140155	89356039	83634406	80641777	5,495	2,524	1,956

NA: data Not Available.
FAO Statistics Division 2011
http://chartsbin.com

India is a leading country in Asia for developing scientific and technological improvement in buffalo nutrition, production, biotechnologies and genetic improvement. Concerning nutrition India has different programs such as "green revolution" (for increasing crop production), the "White revolution" (for increasing milk productivity and improve protein supply for human) and "Red revolution" (increasing red meat production) with focus on buffalo meat production (Borghese & Mazzi, 2005).
The best River buffalo breeds which are Murrah, Nili-Ravi, Surti and Jaffarabadi originate from the north-western part of the India. They have a good potential for milk and fat production at the same time they could be used for meat production and draught power. The Murrah breed, is distributed almost all over the world from Asia to Bulgaria and South America mainly because it is highly productive (Figure 2). Table 3 represents production traits which are used in breeding program for Indian buffalo.

Table3. Production traits of Indian breeds (ICAR, 1997)

Name of the Breed	Average Lactation length (days)	Average Lactation yield (kg)	Milk Composition
Badhawari	272	780	Av.Fat 8.6%
Jaffarabadi	319	2151	Av.Fat 7.86%
Marathwada	302	900	NA
Mehsana	305	1893	Av.Fat 7%
Murrah	305	1675	Av.Fat 7.3%
Nagpuri	286	1055	Av.Fat 7.7%
Nili-Ravi	294	1820	Av.Fat 6.8%
Pandharpuri	305	1142	Av.Fat 7.0%
Surti	305	1289	Av. Fat 7.9%
Toda	200	500	Protein 4.45%

ICAR: Indian Council of Agricultural Research

Buffalo is playing an important role in the farmer economy in Asian countries such as: India, Pakistan, China, Indonesia and Vietnam. Buffaloes are highly adapted to a hot and humid climate and they are used as a draught power for small farms. They produce high quality milk and meat, and are consequently important as financial backup when the risk of crop failure in case of natural calamities (Dhanda, 2004).
Information about Swamp buffalo is limited. Faruque 2003, reported that Swamp type mostly used for milk production and commonly bulls are used as a power source in paddy field or any other agricultural operation.

2.1.1.2 China

China has the most variety of the buffalo population. Mostly they are of the Swamp buffalo type with long domestication history and they provide different variety of products.

Buffalo in China is spread all over the country in environments with different climate and altitudes. From mountain to the lowland, different breeds can be found. Lowland breeds are raised on paddy field where they are mostly used for agricultural activities. Breeds included are: Binhu breed in Hunan province (461,000 heads), the Fuan in Fujian province (70,000 head), Xinyang in Henan province (290,000 head), Yanjin in Yunnan (45,000 head), Xinglong in Hainan and Wenzhou (24,000 head), Wenzhou in Zhejiang (10,000 head), Shanghai breed in Shanghai (36,000). Mountain breed is such as Fuling (415,000 head), Dehong (390,000 head) and Diandong (220,000 head) (Zhang Chunxi and Li Zhongquan, 2001).

Buffalo can tolerate high range of temperature, from 0°C up to 30°C or even higher. Swamp buffalo morphologically has long horns, grey coat color, varying from deep grey to black or brown. Commonly they have white spot in the breast or ring shape spot on the neck.

Swamp buffalo in China is like in other parts of the world used for draught. The exception is the Wenzhou breed which is used mainly for milk production and they produce 1020 Kg milk in 278 days. Jianghan with 800 kg milk in 8-12 months has the second place for milk production. A Buffalo lactation period is highly variable from 150 to 300 days (Zhang Chunxi and Li Zhongquan, 2001).

Two exotic dairy buffalos imported to China, one was Indian Murrah breed in the 1950s and the other one was Nili-Ravi from Pakistan in 1970s. Extensive experimentation such as: Artificial Insemination, feeding observation and cross breeding have been performed on these breeds (Liang Xian-wei et al., 2004).

Crossbreeding system significantly improve milk production in Chinese Swamp buffalo. Milk production upgrading in crossbred Murrah F1, F2 was 1240.5 kg and 1423.3 kg, respectively, which is higher than 1092.8 kg from selected Swamp herd. Normal local Swamp buffaloes produce between 500-800 kg per lactation. Crossbred Nili-Ravi F1, F2 were produced 2041.2 kg and 2325.6 kg per lactation. The data for Triple-crossbred offspring are also shown milk yield improvement which was 2294.6 kg and 1994.9 kg respectively (Yang Bingzhuang et al., 2003) shown in Table 4.

Table4. Milk performance comparison between different crossbreed

(Yang Bingzhuang et al., 2003)

Breed	Lactations (n)	Lactation length (days)	Milk yield (kg)	Average milk yield per day (kg)	Highest daily milk yield (kg)
L	70	280.4±20.2	1092.8±207.4	3.79	6.6
M	237	324.7±73.9	2132.9±78.3	6.57	17.4
N	164	316.8±83.6	2262±663.9	7.14	18.4
MLF1	157	313.7±96.7	1240.5±479.8	3.95	7.57
MLF2	118	313.9±90.1	1423.3±534.5	4.53	8.3
NLF1	45	326.7±96.4	2041.2±540.9	6.25	16.65
NLF2	55	321.4±118	2325.6±994.4	7.22	19.35
N.MLF2	168	317.6±78.4	2294.6±772.1	7.22	18.8
N.MLG1	70	329.1±89.8	1994.9±635.0	6.06	18,50

L: local, M: Murrah, N: Nili-Ravi, G: Santa Gertrudis

2.1.1.3 Indonesia

The buffalo population in Indonesia has been divided into Swamp and River buffaloes. Most of buffalos are of the Swamp type with different subtypes and varieties of breeds. Variation is in different locations period, size, weight, color, marketing and horn dimension. Buffaloes are mostly used for working power, but also they have capacity for milk production. Buffaloes in Java lowland and Sumatra uplands area are used for beef production. A special subtype of the swamp buffalo is called Spotted buffalo because of its unique coat color patterns. These buffaloes are considered holy and have higher price and they are consumed on special ceremonies like marriage and funerals. Most Buffalo populations, which are held by small farmers has 1000 kg milk yield per lactation. Fresh milk production has not increased over the recent years and production could not meet the consumer demand. Smallholder dairy farms are producing over 90 percent of the fresh milk. Farmers have basic technology, insufficient human resource, management problem and limited animal health care. Beef cattle are feeding by the crop residue without sufficient supplementary feed. They have late first calving age and calving interval is quite long between 18 to 24 months. These low performances compared to cow and buffalo means that Indonesian buffalo do not meet the demand for milk and beef supply (Borghese & Mazzi, 2005).

Borghese and Mazzi 2005 suggested a plan with the aim to increase of animal protein (Milk and meat) in the human nutrition. This will provide more attention to this business from stakeholders to improve the genetic capacity of Indonesian buffalo. Cross-breeding between Swamp and Mediterranean Italian River buffalo and production of F1 individuals has been used to increase the performance of Swamp buffalo at the same time improve pure local Swamp type.

2.1.2 Africa

Bubalis buffaloes do not exist in Africa except in Egypt. The species known as African buffalo (*Syncerus caffer*) belongs to another classification of "*Bovini*". *Bovine* has tree members: Cattle, Asian buffalo and African buffalo (Figure 1). *Syncerus caffer* belongs to another genus (*Syncerus*) as well. Swamp and River buffalo both belong to *bubalus* genus with 50 and 48 chromosomes, respectively. Number of chromosomes in African buffalo is 52, which totally separate these species from each other (Borghese 2005).

Three million African buffaloes are distributed in different biotopes such as forest, savannah region and South of the Sahara (Ethiopia, Sudan, Zaire, Congo, Chad and South Africa). They can tolerate the harsh environment in that area and they have immunity to the tsetse fly and development of trypanosomiasis. This high tolerance to harsh environment make hybrid between African and Asian buffalos very interesting, but it reality this experiment failed to succeed (Borghese and Moioli, 2000).

Domestication of African buffalo has been done is some countries and they proved that it would be possible to use this wild animal for draught and agriculture uses in *e.g.* Zimbabwe. Introducing some River buffalo traits to the African buffalo was not successful due to lack of immunity to trypanosome, harsh climate, lack of food and water (Alexiev, 1998). Buffalo is more sensitive to direct solar radiation which they usually wallowing in the river for regulating body temperature, so it make it impossible for them to be adapted to the African countries where they have low water resource (Borghese and Moioli, 2000).

2.1.2.1 Egypt

Buffalo population in Egypt was around 3.98 million heads in 2007 and Egypt has highest water buffalo population in Africa. Forty two percent of this population are used in dairy production, 6 percent bulls, 32 percent are heifers with an age of less than two years and 20 percent male calves. Growth rate of buffalo during the period of 2001 to 2007 was 2.12 percent, but still it count as 1 percent of the cattle population (Pasha T.N & Hayat. Z, 2012). Buffalo produces 81 percent of total milk production in Egypt (Borghese & Mazzi, 2005).

Buffalo milk production cost less than importing powdered milk to Egypt. Buffalo milk generated by one ton concentrate of feed mix makes commercial buffalo industry to become more thriving in this area (Soliman and Sadek, 2004). Buffalo spread in Delta region which is located beside the river Nile and Fayum Oasis. Productivity of Egyptian buffalo is about 210-280 days/lactation. They have in average seven lactation periods with milk yield around 1600 kg/year. At the age of 34-40 months they have first calving (Fikri El-Kirabi, 1995).
Table 5 is showing traditional and progressive system in Egypt, two systems which they have different production standards.

Table 5. Different production standards in Egypt (Soliman, 2009)

milk yield for lactation	**Traditional mixed Small Farm System**	**Progressive Farm System**
Total milk/head/season 1	1650	1850
Total milk/head/season 2	1790	2000
Total milk/head/season 3	1950	2400
Total milk/head/season 4	1870	2350
Total milk/head/season 5	1800	2200
Grand Average Milk Yield per Milking Head	1850	2250
Average lactation length (Days)	180	250

2.1.3 Europe

2.1.3.1 Italy

The Water Buffalo population in Italy is smaller compared to the large population sizes found in many East Asian countries, but Italy takes first place in having high quality production system. Quality products are considered produced from animals with high genetic capacity, high technology for producing buffalo's products, pathologies monitoring and hygienically during food production. Buffalo is economically important for worker, occupation and the animal used as producer of mozzarella cheese which represents a typical Italian product of global importance (Borghese & Mazzi, 2005).

ANASB (Italian Buffalo Breeders Association) is controlling the selection and genetic improvement of buffalo in Italy. The aim of this association is to recording all the buffalo population in Italy. In 2010 they recorded 26 percent of total dairy buffalo. They record production twice a day (Morning and evening)

(Table 6). Because of this well organized system, buffalo cows (Buffalo that inseminated by cattle sperm) can produce up to 5000 kg/lactation in 270 days. Many million doses of high quality bull semen with registered performance and progeny test are available in Italy and in the world for artificial insemination. There are several semen centers in Italy for producing semen with high quality such as CoFA (Cooperativa Fecondazione Artificiale) is in Cremona in the north of Italy. They call their breed Mediterranean Italian. They have cooperation with breeder organizations like Holding and Interstates for progeny test. The aim of these centers is to not only increase the quantity of the milk, but also improve mozzarella index (Borghese. 2010).

Mozzarella (kg) = Milk (kg) x (3.5 x % proteins + 1.23 x %fat – 0.88) / 100

The largest population is located in the province of Frosinine and Latina (Lazio region). This location is Denomination of Protected Origin (D.O.P.) area in Italy. Market is based on cheese, mostly mozzarella, and the area is considered to produce the best quality of this cheese. Quality mozzarella is made by hand from raw buffalo milk. This mozzarella is soft, juicy and tasty, includes live ferments and natural yeasts. It doesn't have any preservative and is kept in the mozzarella water, so it's shelf life is only 5 days. The industrial Mozzarella which is also produced in this area is machine made and has a higher shelf life (up to 2 weeks). This mozzarella is suitable for distribution around the country or for export (Borghese, 2010).

Price of the buffalo milk is 3 three times higher than cow milk in the Italian market, which makes the buffalo mozzarella quite expensive. Italian market is consuming 82 percent of mozzarella from D.O.P area and the rest is exported to Germany, France, USA and U.K, respectively (Borghese, 2010).

Buffalo meat has a good market in Italy. They produce high quality meat which obtained I.G.P (Indication Geographic Protected) "Carne di bufala campana". Bull at the age of 15-16 months can produce 400-440 kg carcass with 800-1000 g/d of daily gain. The buffalo meat has a good nutrition quality, less than 50mg cholesterol/100g, both saturated and unsaturated fatty acid less than 1 and iron more than 1.5mg/100g (Borghese et al., 1996).

Table 6. Italian Buffaloes (ANASB 2009)

N° Head	370 000
N° Dairy buffaloes	180 000
N° Recorded buffaloes	46 799
% Recorded Buffaloes	26
N° Recorded farms	290
N° Head/farm	161.3
kg milk production (in 270 d)	2 221
% Fat	8.24
% Protein	4.66

2.2 Buffalo breeds

River buffalo has several breeds of which the most popular are: Jafarabadi, Murrah, and Nili-Ravi. They have different morphological and performance traits. Table 7 takes a brief look at these breeds concerning the fact that data comes from before 2005.

2.3 Buffalo cheese and milk industry

More consumer interests for buffalo products causes an increase in the buffalo population size in those countries where buffalo products are currently being produced. Buffalo milk in Italy has three times higher price than cattle milk. Italy exports 14 different buffalo products (mostly mozzarella) to all over the world and countries such as Germany, France, UK, Switzerland and USA are among the countries that import high quantities of Italian-produced mozzarella (Borghese 2010).

Higher demand for Italian products is due to: D.O.P "Mozzarella di Bufala campana" which is registered in E.U (European Union), unique quality of Italian mozzarella (soft, milk rich and flavors) and Italian style cooking which spread all over the world. Small farmers have marketing problem compared to industrial scale. They can produce high quality mozzarella, but it has short shelf life and it is not suitable for export. Industrial products have lower quality (taste and flavors), but those are more suitable for export to other countries (Borghese 2010).

Cheese classified into three categories based on the coagulation: enzymatic coagulation (due to the rennet), acid coagulation (lactic bacteria or after natural acidification) and mixed of both acid and enzymatic. In the recent type, in some cases, acid coagulation prevails and in the others enzymatic prevails. Most Mediterranean cheeses are belonging to the category based on acid coagulation. Acid coagulation could happen because of milk left to be acidified and then they add rennet or long coagulation time. Temperature, climate and herd condition can affect the decision of which technology should be used for cheese making. Mozzarella is produced from the southern part of Italy where the weather is hot and animal feed on the pasture, and when the milk arrives to factory for cheese making process, the acidification has already started. Because of the same situation in other Mediterranean countries, most of their cheeses are classified as acid coagulation (Borghese et al., 2000).

Cheese producers in Italy started an organization in 1993 to differentiate the buffalo mozzarella from bovine or mixed milk mozzarella. Original mozzarella named with "Mozzarella di bufala campana" D.O.P which become a trade mark in 1996 and must come from 100 percent raw buffalo milk. The milk should come from the area in which buffaloes have been held for centuries and the same local producing technique should be used. This organization approved by Italian government and EU. 24 percent of buffalo milk could produce mozzarella compared to bovine milk, where only 13 percent of milk is suitable for mozzarella production. Furthermore, buffalo mozzarella has higher fat and protein content compared with mozzarella produced from cattle milk. This certification also labeled on the mozzarella wrapping which help the costumer to buy the original tasty mozzarella. Cheese production in India is quite young either production or consumption. Many different soft cheeses for examples Indian soft cheese and Paneer which produce from direct acid coagulation without salt are common in India (Borghese 2005).

2.3.1 Fat-rich milk products

Churning product of the buffalo milk is faster and it gives higher overrun than the cow cream. Buffalo milk has bigger globules size and higher proportion of solid fat. This specificity makes the cream easier

and faster to separate from the residue. Buffalo butter has higher stability compare to cow, because it has higher fat and slower rate of fat hydrolysis (Patil and Nayak, 2003).

3 Literature review

3.1 Milk composition

3.1.1 Milk

Fat, solids-not-fat (SNF) and a fluid carrier are three main components of the raw cow milk. Milk could be a food by its own because of high nutrient contents (Walstra et al., 1999; Patton, 2004). 12 percent of milk's weight comes from a natural oil-in-water which is contains lipids, proteins, lactose and mineral. Water can hold these entire components in soluble form. Milk whey protein contains fat globules and casein micelles when the light scattered by them it cause white color in milk (Ribadeau-Dumas and Grappin, 1989).

Importance of milk protein in human nutrition influences the properties of dairy products (Huth et al., 2006). Milk protein consists mainly of whey, casein and milk fat globule membrane proteins (MFGM) which is 1 percent of total milk protein. Casein is approximately 80 percent of milk protein which is includes α-s1 and α-s2 caseins, β-casein, and κ-casein (Walstra et al., 2006). Whey protein which also called the serum or globular proteins is includes β-lactoglobulin, α-lactalbumin, bovine serum albumin (BSA), and immunoglobulins (Ig). MFGM has the lowest percentage in milk protein, but its plays important role in milk secretion mechanism, enzymatic activity of the lactating cell membranes and interaction with serum proteins in whole milk (Cavaletto et al., 2004).

Milk lipids are in the form of fat globules. Natural creaming of the milk is the process when the fat in the milk because of less density than the milk plasma travels to the top of the milk (Walstra et al., 1999).

Table 8. Milk composition in different mammals (Patton, 2004)

Nutrient	Cow	Goat	Sheep	Buffalo	Human
Water (%)	87.99	88.9	83	81.1	87.5
Fat (g)	3.34	3.5	6	8	4.38
Protein (g)	3.29	3.1	5.4	4.5	1.03
Carbohydrate (g)	4.66	4.4	5.1	4.9	6.89
Energy (kcal)	61	60	95	110	70
Lactose (g)	4.8	4.4	5.1	4.9	7
Cholesterol (g)	14	10	11	8	14
Calcium (mg)	119	100	170	195	32

3.1.1.1 The globule membrane in milk

3.1.1.1.1 Origin

Reticulum of the mammary epithelial cell are synthesizing milk lipid which is mainly triglycerides. First lipids are making very small microlipid droplets (between 0.5 to 4 µm diameter sizes). The monolayer membrane consisting of polar phospholipids and proteins are surrounding the microlipid droplets (Keenan, 2001). Microlipid droplets combine together to form cytoplasmic lipid droplets (CLD). The CLD reaches the apical cell surface and slowly travel to the membrane and is forced out from the mammary gland secreting cell. These final fat droplets (0.2 – 15 µm diameter) are surrounded by a membrane called milk fat globule membrane (MFGM) (Bylund, 1995).

3.1.1.1.2 Composition and structure

The MFGM has three layers which are distributed asymmetrically (Singh, 2006). The first layer is a monolayer consisting of phospholipids and proteins, which are conjugated from the endoplasmic reticulum. The second layer consists of proteins, phospholipids, enzymes, cholesterol and other minor components, which is conjugated from mammary gland secretory cells. Between these two layers is an electron-dense coat that is rich in protein (Evers, 2004; Heid and Keenan, 2005).

3.1.1.1.3 Functional and biological properties

MFGM has an amphiphilic nature which functions as an emulsifier in milk. MFGM can increase the absorption of fat-soluble nutrients like vitamin D_3 and vitamin A. Some studies found the antimicrobial proteins and antiviral properties (Spitsberg, 2005; Lopez et al., 2008).

3.1.1.1.4 Diacylglycerol acyltransferase

Total milk lipid's content varies among species. Over 99 percent of bovine milk lipid is consisting of droplets and triacylglycerol (Neville and Picciano, 1997). There is a variation within bovine milk because of physiological factors like stage of lactation, diet and body lipid content (Neville and Picciano, 1997). Carbohydrate stores in the form of triacylglycerols in the body which make them the most important source of energy for eukaryotic cells (Cases et al., 1998). Mammary secretory epithelium secretes the lipid from the blood to the milk (Neville and Picciano, 1997) (see chapter 4.2).

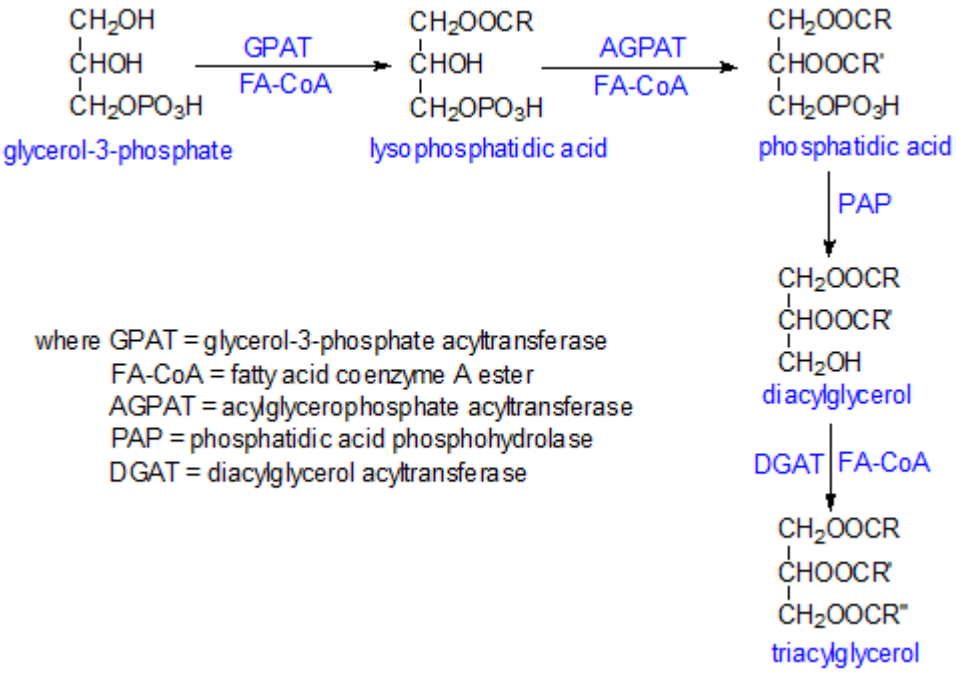
Endoplasmic reticulum is where the biosynthesis of triglyceride occurs. In mammals, this *de novo* pathway is through the glycerol-3-phosphate pathway. By way of taking glycerol-3-phosphate from either glucose or glycerol in the plasma this pathway start. Glycerol-3-phosphate pathway and other pathway (for instance monoacylglycerol pathway) are synthesizing diacylglycerol. Diacylglycerol is converting to triacylglycerol by diacylglycerol acyltransferase (*DGAT*) enzyme (Coleman et al., 2000) (Figure 4). The rate-limiting effect of *DGAT1* on triglyderide secretion is unknown, but it could be studied on knockout mice (Chi-Liang et al., 2008)

50 percent of total milk lipid comes from the fatty acid products of the plasma (Eckel, 1989). Ruminant's milk triacylglycerol contains the short-chain acids butyric and hexanoic acid. These short-chain fatty acids are not found in the non-ruminant's milk (Marshall and Knudsen, 1977). In the final step of the triglyceride synthesis, *DGAT* controls the rate of the triglyceride production (Chen el al., 2002a).

Weiss et al. 1960 identified the activity of DGAT and two genes encoding genes which are *DGAT1* and *DGAT2*. Knockout mice for *DGAT1* show that *DGAT1* is necessary for lipid synthesis within the mammary gland, but also it is important for milk production. These *DGAT1* knockout mice were unable to

produce milk. They conclude that *DGAT1* is a strong candidate gene for milk fat percentage. This study shows the importance of *DGAT1* in milk production and milk fat synthesis (Winter et al., 2002).

Figure 4. Biosynthesis of triglycerides in mammals through glycerol-3-phosphate pathway.



Adapted from http://lipidlibrary.aocs.org

3.2 Genetic analyses

3.2.1 Breeding and selection in dairy industry

In dairy industry, farmers' selection of animals is based on the phenotypic observation for higher milk production in the herd. Artificial insemination (AI) is a helpful solution for reaching higher genetic improvement since 1950's for traits with high heritability. Progeny testing is used for measuring breeding value of top performance sires. Average phenotypic performance from daughters of high performance parents give an estimation of progeny testing in dairy industry (Georges et al., 1995).

Progeny testing has some limitation due to reproduction age and favorable traits are only expressed in one gender. This causes expense and time, because it takes three and a half year and 25,000 dollars per bull which only one out of nine bulls successfully passes the test (Tassell, 2003). The use of molecular genetics instead of phenotypic evaluation could be a good alternative due to the genetic evaluation of each individual without concerning about ages and gender. The combination of genetics with correlation to phenotype is necessary. In this method when favorable genotypes have been define not all individual need to be phenotyped.

Polygenic traits such as milk production and composition are influenced by genetic, epigenetic and environmental factors. Good nutrition could improve milk production and less care about management system could cause a decrease in milk production even in the top performance breeds (Smaragdov, 2006).

Multiple genes and environmental factors are affecting milk fat percentage which is a quantitative trait. Milk fat percentage has heritability (genetic contribution to the variation) between 0.45 and 0.5 (Winter et al., 2002). Heritability is useful for breeding programs to help the breeder predict how offspring genetically perform compared to his or her parent. New selection methods like genotyping will help us to have a better approach in this regard, because milk fat is presenting in one gender and reproductive age limits is another obstacle.

3.2.2 Quantitative trait analysis

A large number of genes with small effect (polygenes) are affecting economically important traits such as milk fat content and composition. Molecular genetics methods are used for identifying the underlying genetic background of these quantitative traits. Marker assisted selection (MAS) is a method for selection based on the presence or absence of the genetic marker that are linked to the specific trait (Smaragdov, 2006). MAS test the animal based on their genetic potential in the early age and independently (Winter et al., 2002).

One of the most important genetic mapping research methods is the use of quantitative trait loci (QTL) mapping. Researchers can track the chromosomal position that correlate with quantifiable complex traits and identify polygenes that are associated with the phenotype by markers (Geldermann, 1975; Smaragdov, 2006). Based on the phenotype trait and genotype correlations positional candidate genes can be identified that are then further studied for functional confirmation. The main purpose is to identify association between positional candidate genes and phenotype trait in one or more populations (Shorten et al., 2004). SNPs (Single Nucleotide Polymorphisms) are efficient genetic markers because of their high frequency and distribution in the genome and the availability of dense SNP arrays that can be used for genome-wide analyses. They are considered the most important genetic markers for genome-wide genetic mapping projects and also for MAS.

3.2.3 DGAT1

A DGAT enzyme uses diacylglycerol and fatty acid CoA as substrates. Triacylglycerol metabolism of DGAT involves this enzyme in intestinal fat absorption, adipose tissue formation, lactation and lipoprotein assembly (Cases et al., 1998). In excess energy situations, glucose and insulin-mediated signalling result in increasing the transcription of genes, because the excess energy will activate fatty acid synthesis and triglyceride formation to store the energy as an adipose tissue (Meegalla et al., 2002).

Glucose increases the *DGAT1* mRNA expression and insulin increase *DGAT2* mRNA expression and consequently increases DGAT activity. In the fasting/re-feeding situation DGAT1 was more active in the process of fat absorption since DGAT2 was shown to primarily participate in VLDL particle assembly in the liver (Meegalla et al., 2002).

Grisart et al., 2004 mapped *DGAT1* on the centromeric end of bovine chromosome 14, which is the location for a QTL related to milk percentage and performance. *DGAT1* is thus a prime positional candidate gene for milk fat percentage based on both functional and positional data. Allele (aa) of *DGAT1* gene has higher milk fat percentage. Two nucleotide polymorphisms (aa to gc) in exon VIII has been reported to be associated with milk production and composition. Exon VIII is highly conserved within and between species except in position 232 which substitute lysine to alanine. Bovine who carries lysine232 polymorphism has higher milk fat (Grisart et al., 2004).

4 Materials and Methods

4.1 *DGAT1* Genotyping

4.1.1 Collection of hair samples and handling

prepared DNA sample from Indonesian animals were used for this experiment. The hair samples from imported Italian River buffalo were collected from Ängsholmens Gårdsmejeri (Floberga 915, 740 47 Harbo, Sweden). 17 River buffalo were available and used for hair collection.

4.1.2 Preparation of genomic DNA from hair

4-6 hairs placed in an eppendorf tube in which roots were down at the bottom of the tube. 100 µl lysis buffer (QIAGEN®, GmbH, Germany) and 2.25 µl proteinase K (8 mg/ml) were added to the hair tubes. Reactions were incubated for 3 hours at 55°C. Proteinase K was deactivated by increase the temperature to 95°C and incubated for 10 min. Tubes were centrifuged for 5 seconds and then the clear supernatants were transferred to new tubes and stored in the freezer at -20°C.

4.1.3 Primer design

River buffalo *DGAT1* gene sequences around 10 kb were obtained from a search in NCBI (http://www.ncbi.nlm.nih.gov/). Designing the primers was done with primer3 online software (http://bioinfo.ut.ee/primer3-0.4.0/) to get the fragments between 300-500 base pairs (bp). The primers were analyzed by BLAT on the UCSC cattle genome (http://genome.ucsc.edu/cgi-bin/hgBlat) to confirm the specificity and position of the primers in the genome. Candidate primers were chosen based on the melting temperature around 59°C to 61°C with 1°C differences between forward and reverse primer. The optimal primer size was between 18-22 bases.

4.1.4 Polymerase Chain Reaction (PCR)

Diluted genomic DNA was used for performing PCR (1ng/µl) with primer pairs specific for buffalo *DGAT1* based on the following circumstances: 25 µL total volume containing 2 µl DNA (10 ng/µl), 03 µl AmpliTaq® Gold DNA polymerase (Applied Biosystems, Brachburg, NJ), 0.2 mM dNTPs (Invitrogen, Carlsbad, CA), 2.5 µl GeneAmp® 10x PCR Buffer (Applied Biosystems), 2.5 µl $MgCl_2$ (Applied Biosystems) and 16.9 µl PCR-grade water. Thermocycling was performed on a Applied Biosystems 2720 Thermal cycler under the following PCR protocol: initial denaturing at 94°C for 5 min; followed by 10 cycles at 94°C for 30 seconds, 65°C for 30 seconds, 72°C for 30 seconds then 30 cycles at 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 30 seconds and final extension at 72°C for 7 min and cooled down and stayed at 4°C.

To determine successful PCR amplification 4 µl of PCR products and 2 µl SYBR® safe (Invitrogen©) was loaded onto the 1% agarose gel (0.5 X TBE) (Invitrogen©, UltraPure™ Agarose) containing ethidium bromide and electrophoresed at 90V for 30-45 min (Based on the length of the fragment). Molecular size ladder (100bp) was loaded at the same time with the samples for proving the correct length of fragment (Figure 5 and 6). For visualizing the gel we used UV transilluminator (BIO RAD laboratory, Inc, universal Hood II,). The remaining PCR products were stored in the cold room at 4°C for subsequent sequencing.

4.1.5 Sequencing

PCR products were cleaned from unwanted amplification with EXOSAP product for high quality sequencing. Two hydrolytic enzymes utilizes by ExoSAP, Exonuclease I and Shrimp Alkaline Phosphatase (SAP). In a specially formulated buffer they remove unwanted dNTPs and primers from PCR

products. Remaining dNTPs removes by SAP. Residual single-stranded primers and extraneous single-stranded DNA product removes by Exonuclease I. ExoSAP reaction mix was containing: 0.08 µl Alkaline Phosphatase, EXO1 0.1 µl, EXO1 buffer 10 µl which make the total volume became 2.18 µl. 2 µl ExoSAP reaction mix added to the PCR well. Thermocycling was performed at 37°C for 60 min, 85°C for 15 min and hold at 4°C in the Thermocycle machine. 3 µl of ExoSAP product, 3 µl primers (only forward or reversed) and 12 µl PCR-grade water added to the new well and sent for Sanger sequencing at Rudbeck Laboratories (Uppsala Genome Center, Sweden).

4.2 Data preparation

4.2.1 Sequence analysis

Sequence analysis was done by CodonCode Aligner 4.2.2 (CodonCode Corporation) for possible Single Nucleotide Polymorphism (SNP) between and within two subtypes (River and Swamp). The genomic sequence of SNP positions was compared against human genome in UCSC genome browser for possible conserved region between 29 mammals data base available on USCS genome browser. The reason we used human genomes is variety of options and most complete data available on the UCSC genome browser.

4.2.2 Format conversion

Data generated by UCSC genome browser was in Multi Alignment Format (MAF) which converted to FAST Alignment (FASTA) by convertor available in Galaxy online package (https://main.g2.bx.psu.edu/). The converted file read by Jalview software (http://www.jalview.org/) for percentage of identity between sequences.

4.2.3 Promoter analysis

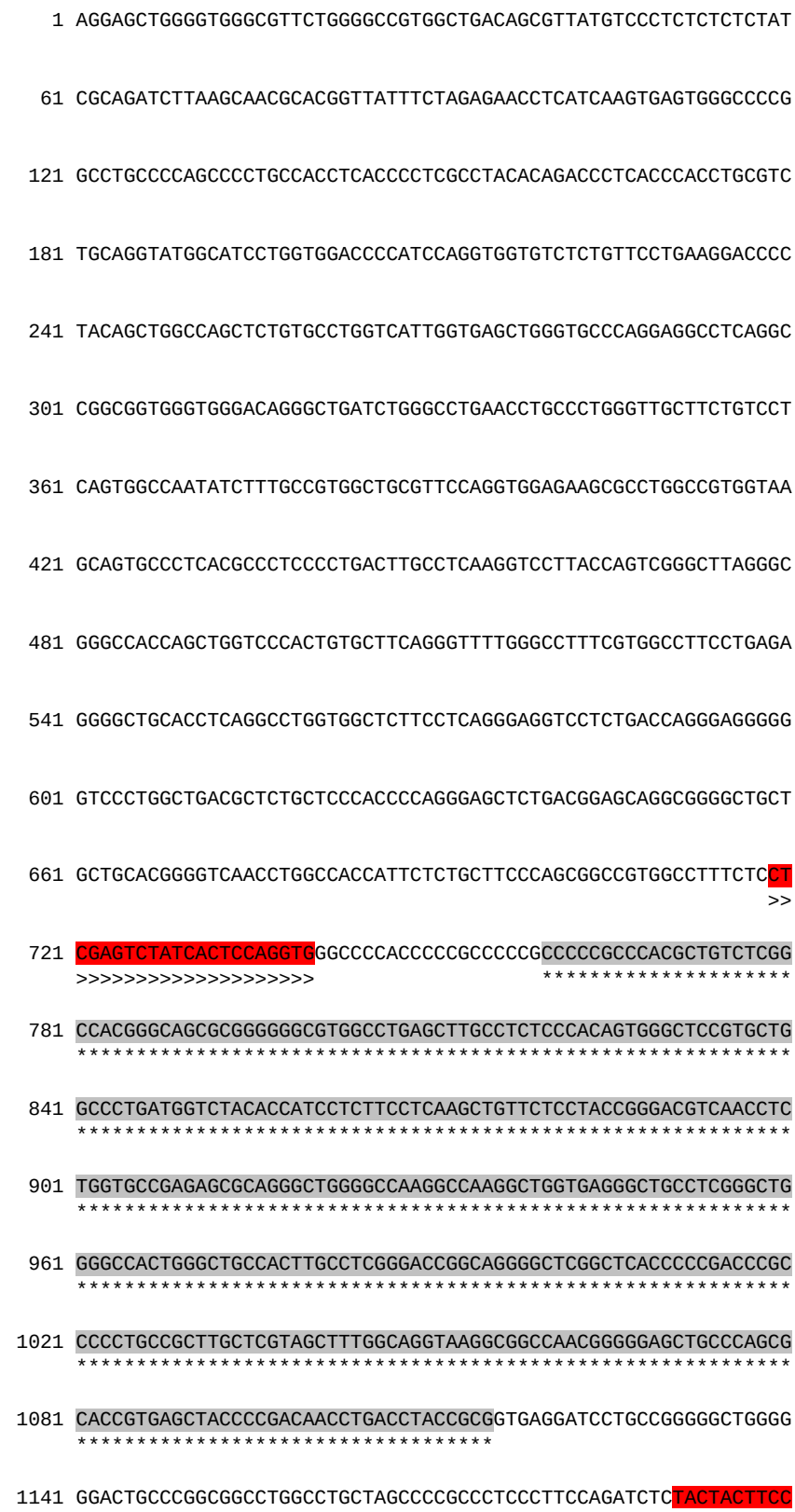
1.5 Kilobase (KB) region was hypothesized as a potential promoter upstream of the *DGAT1* translation start site (Methionine) then the genomic region was downloaded from UCSC genome browser for all available mammals on the data base (29 mammals). Consensus motifs analysis in the SNP position was done by MEME software (http://meme.nbcr.net/meme/) then the results were analyzed by TOMTOM motifs comparison tools (http://meme.nbcr.net/meme/cgi-bin/tomtom.cgi) for identifying possible transcription factor binding site.

5 Results and discussion

5.1 *DGAT1* genotyping

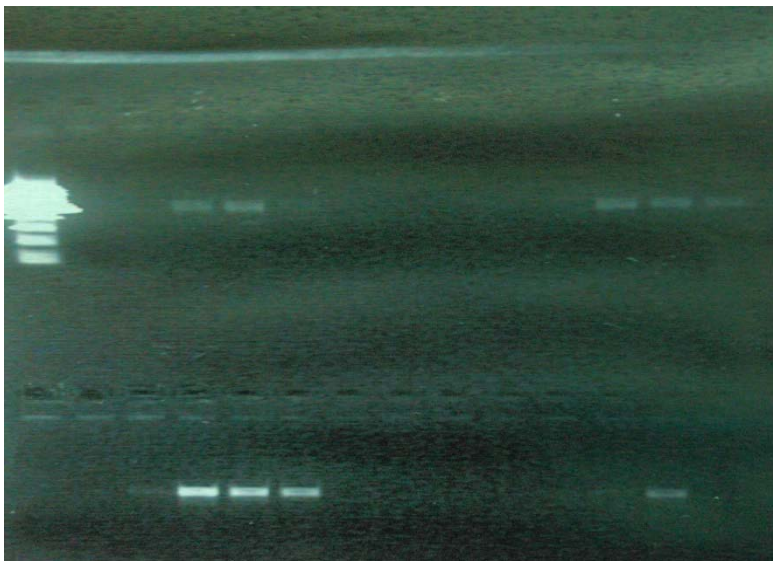
The gel photos show the quality of the PCR fragment. Designing several primers and changing the protocol and concentration of the master mix couldn't improve the quality of the bands for this specific region. That might be because of unavailable buffalo genome sequence. We tested the region of primer based on the UCSC genome browser for cattle genome, which may be slightly different from buffalo genome. The other obstacle could be quality of the DNA from hair samples.

Figure 5: partial *DGAT1* gene specifying exon 8-9 (gray) where K232A located , primer sites (red) from NCBI on PubMed.



```
                                                          <<<<<<<<<<

1201 TCTTCGCCCCCACCCTGTGCTACGAGCTCAACTTCCCCCGCTCCCCCCGCATCCGAAAGC
     <<<<<<<<<<

1261 GCTTCCTGCTGCGGCGACTCCTGGAGATGGTGAGGCGGGGCCTCGCGGGCCAGGGTGGGC

1321 GGGCCTGCCGGCACCCGGCACCGGGGCTCAGCTCACTGTCCGCTTGCTTCCTTCCCCAGC

1381 TGTTCCTCACCCAGCTCCAGGTGGGGCTGATCCAGCAGGTACGTGCCCGGGGGGGGGGGG

1441 GGGGGGGGGGGGGGGGACGGGGACTCTGGGGCCGTTGGGGAGCTGACTCTGCGCTTTTTG

1501 CAGTGGATGGTCCCGGCCATCCAGAACTCCATGAAGCCCTTCAAGGTGAGCAGGCAGGCC

1561 TGGCAGGGTGGGTTCCGGGGTCAGGGCTGAGG
```

Figure 6: A 1% agarose gel photo, result from PCR to amplify *DGAT1* exon 8-9. Lane 1: 100 BioRad Ladder, Lane 4-6 from second row shows the PCR fragment at 492 bp.



For covering the entire *DGAT1* gene and promoter region we initially designed fifteen primer pairs of which 5 of them for promoter region gave a clear amplification product, but other primer pairs were not successful. We focused on the high quality PCR fragments for finding SNPs between two populations by the use of Codon code software. Primer pair designed for covering the region between exon/intron 11,12,13 which includes intron and exon 12 in *DGAT1* gave a high quality fragment which after the sequencing were shown to be identical between populations (Figures 7 and 8).

Figure 7: Agarose gel photo results from primer covering exons/intron 11, 12,13 for *DGAT1*. First and last lanes are ladders. The product size is 400bp.

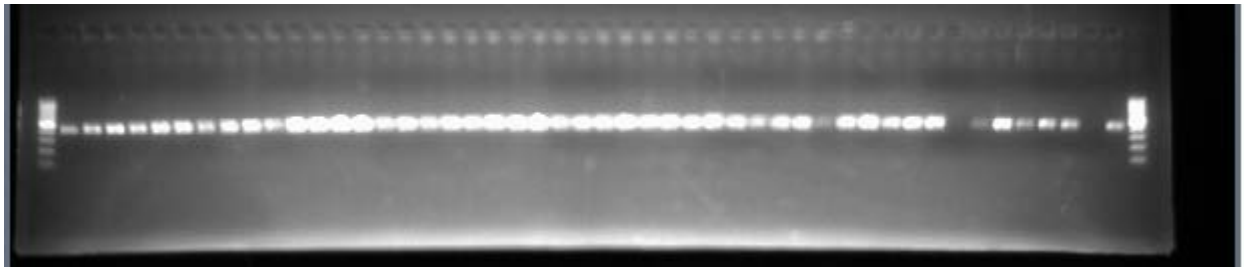


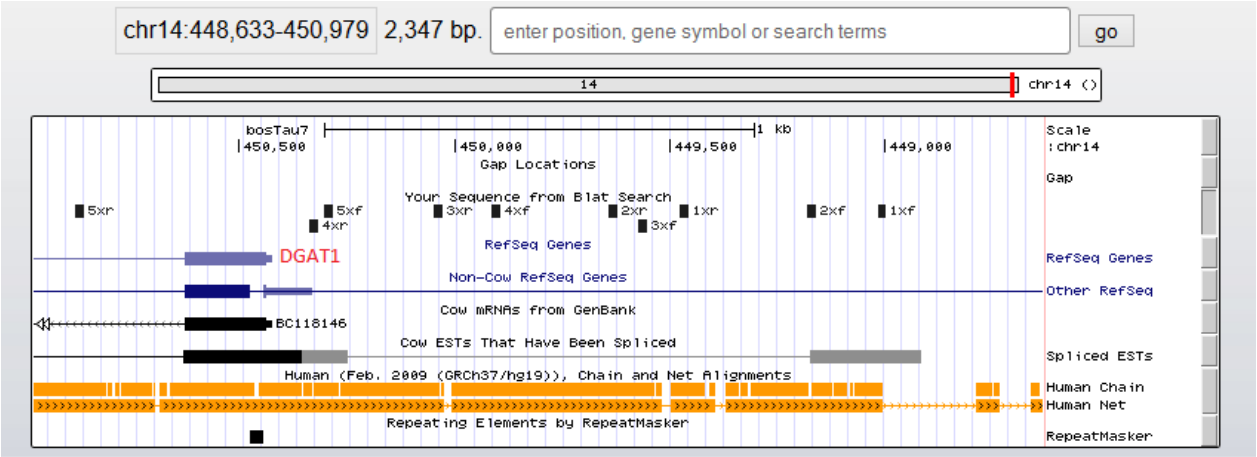
Figure 8: partial *DGAT1* gene specifying exon/intron 11, 12,13 (gray), primer sites (red) from NCBI on PubMed

```
1 GTTGTCGCTGTGGCCAGGGGCGCTGACCCGCTTGGTGGGACGGACGGCCGCTGGGCAGCA

   61 GGTTTCTTCTGCCACGGTGGCACAGGCACCTGGGGTTGTGGTTGGCTCCAGGCGGGCGGG

  121 GGCTGCGTGCCCCTGCGCAGGCACATAGGCCGTGGGTGGGGAGTCTCAGAGCTTGGCGTG

  181 AGGTCCCACAGGGCTGGGCCTGCAGGATGGAGGCCACTGTCCTGAGCTGCAGGTGCTGGC

  241 AGGAGCTGGGGTGGGCGTTCTGGGGCCGTGGCTGACAGCGTTATGTCCCTCTCTCTCTAT
                                          >>>>>>>>>>>>>>>>>>>>

  301 CGCAGATCTTAAGCAACGCACGGTTATTTCTAGAGAACCTCATCAAGTGAGTGGGCCCCG
        ************************************************************

  361 GCCTGCCCCAGCCCCTGCCACCTCACCCCTCGCCTACACAGACCCTCACCCACCTGCGTC
      ************************************************************

  421 TGCAGGTATGGCATCCTGGTGGACCCCATCCAGGTGGTGTCTCTGTTCCTGAAGGACCCC
      ************************************************************

  481 TACAGCTGGCCAGCTCTGTGCCTGGTCATTGGTGAGCTGGGTGCCCAGGAGGCCTCAGGC
      ************************************************************

  541 CGGCGGTGGGTGGGACAGGGCTGATCTGGGCCTGAACCTGCCCTGGGTTGCTTCTGTCCT
      ************************************************************

  601 CAGTGGCCAATATCTTTGCCGTGGCTGCGTTCCAGGTGGAGAAGCGCCTGGCCGTGGTAA
      ************************************************************<<<<

  661 GCAGTGCCCTCACGCCCTCCCCTGACTTGCCTCAAGGTCCTTACCAGTCGGGCTTAGGGC
      <<<<<<<<<<<<<<<

  721 GGGCCACCAGCTGGTCCCACTGTGCTTCAGGGTTTTGGGCCTTTCGTGGCCTTCCTGAGA

  781 GGGGCTGCACCTCAGGCCTGGTGGCTCTTCCTCAGGGAGGTCCTCTGACCAGGGAGGGGG

  841 GTCCCTGGCTGACGCTCTGCTCCCACCCCAGGGAGCTCTGACGGAGCAGGCGGGGCTGCT

  901 GCTGCACGGGGTCAACCTGGCCACCATTCTCTGCTTCCCAGCGGCCGTGGCCTTTCTCCT

  961 CGAGTCTATCACTCCAGGTGGGCCCCACCCCCGCCCCCGCCCCCGCCCACGCTGTCTCGG
```

5.2 Identification and genotyping of the *DGAT1* promoter

The result from UCSC genome browser for five different primers shows that primers cover all the promoter area of the DGAT1 gene (Figure 9).

Figure 9: UCSC result for primers covering promoter area (1.5 kb) of DGAT1.



The PCR products from 3 individuals as a trial had high quality which clearly shows the bands, so these products were sequenced as described above (Figure 10). As it is clear, primer 5x did not show any amplification product.

Figure 10: PCR result of three random individuals (2 Swamp and 1 River). First band from left is ladder and then from left to right are fragments (477-493) from primer 1x-5x respectively.

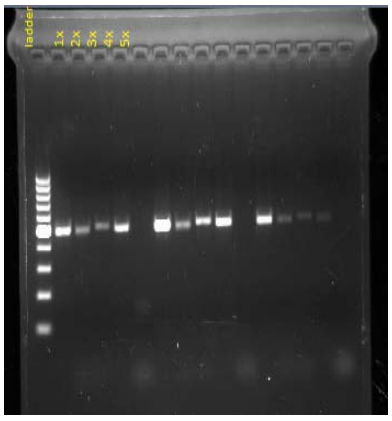


Table 9: list of primers covering 1.5 kb upstream (promoter) of DGAT1

	Length	TM	GC %	Product Size	Seq
1XF	20	59	50	481	gcccaactctgcctatgaat
1XR	20	59	50	481	gcaaagagcatggagtgtga
2XF	20	61	55	480	agcctgccaggtaaggatgt
2XR	20	60	50	480	ttaggatcctggggtgttga
3XF	20	59	45	493	ttgcttcattcaacgactgg
3XR	21	60	42	493	ggatttcgtcatttctcagca
4XF	21	59	47	477	cggtggtaattagtgggttca
4XR	20	61	60	477	accgcctctactacgccact
5XF	18	60	50	600	gatttgcacgcgctgatt
5XR	20	58	55	600	agggcgtctagaagttgagc

Sequencing results were investigated by Codon Code software for detection of possible SNPs. In the area of the genome that covered by primer (3x) we found one SNP, which was different between the two populations (Swamp vs. River). SNP positioned at 857 bp 5' of the translation initiation codon (ATG).

5.3 Bioinformatics analysis of promoter and SNP position

Observation by Galview software for comparative genomics of the 2 kb upstream of the translation start site (hypothetic promoter) showed that the area near translation start site is highly conserved between species available as expected (Figure 11). This block of highly conserved sequence between species is a strong indication of its functional importance. The length differences of promoters which is the thinner blue area in the region upstream of the genes (figure 12) between human and cattle might be because of limited functional studies performed on the cattle *DGAT1* gene for clearly identify the promoter of *DGAT1* gene. Based on the blat result of the 1.5 kb region upstream of the translation start site of buffalo *DGAT1* they have 97.2 % similarity (Figure 13). This difference between *DGAT1* in these two species might be important for different control of milk fat production in these two species. Future study on the nature of the buffalo DGAT1 gene will be necessary to further define the mechanisms underlying such potential differences. Clear sequence of the promoter of *DGAT1* and recording data for both River and Swamp subtype can explain the big differences between cattle and buffalo milk fat percentage (3.5 % vs. 6-8%).

Result from MEME analysis and then by using the TOMTOM database for identification whether the SNP influenced consensus binding site for transcription factors did not reveal a particular consensus motif. The result from the area around the SNP position for potential binding site didn't show conclusive results as well. The only difference is that our River population had different SNP on that position T instead of G on the online database (Figure14 and 15). That might be important, but we had only access to a small sized population of river buffalo that were imported from Italy. The other obstacle was that we did not have milk recording of the population and access to the progeny of our population. The result however, reveal low sequence similarity between buffalo and cattle compared to the input file which was included around 10 slightly related animal based on the phylogenetic tree.

Result from transfac professional (TRANSFAC®) on both promoter region of cattle and buffalo showed some unrelated motifs to *DGATI* gene which it could not be useful for this project. Future high quality sequencing and functional study need to find the specific binding region for *DGAT1*.

Figure 11: Galview photo from 50 bp upstream of the translation start site. Darker blue shows percentage of identity.

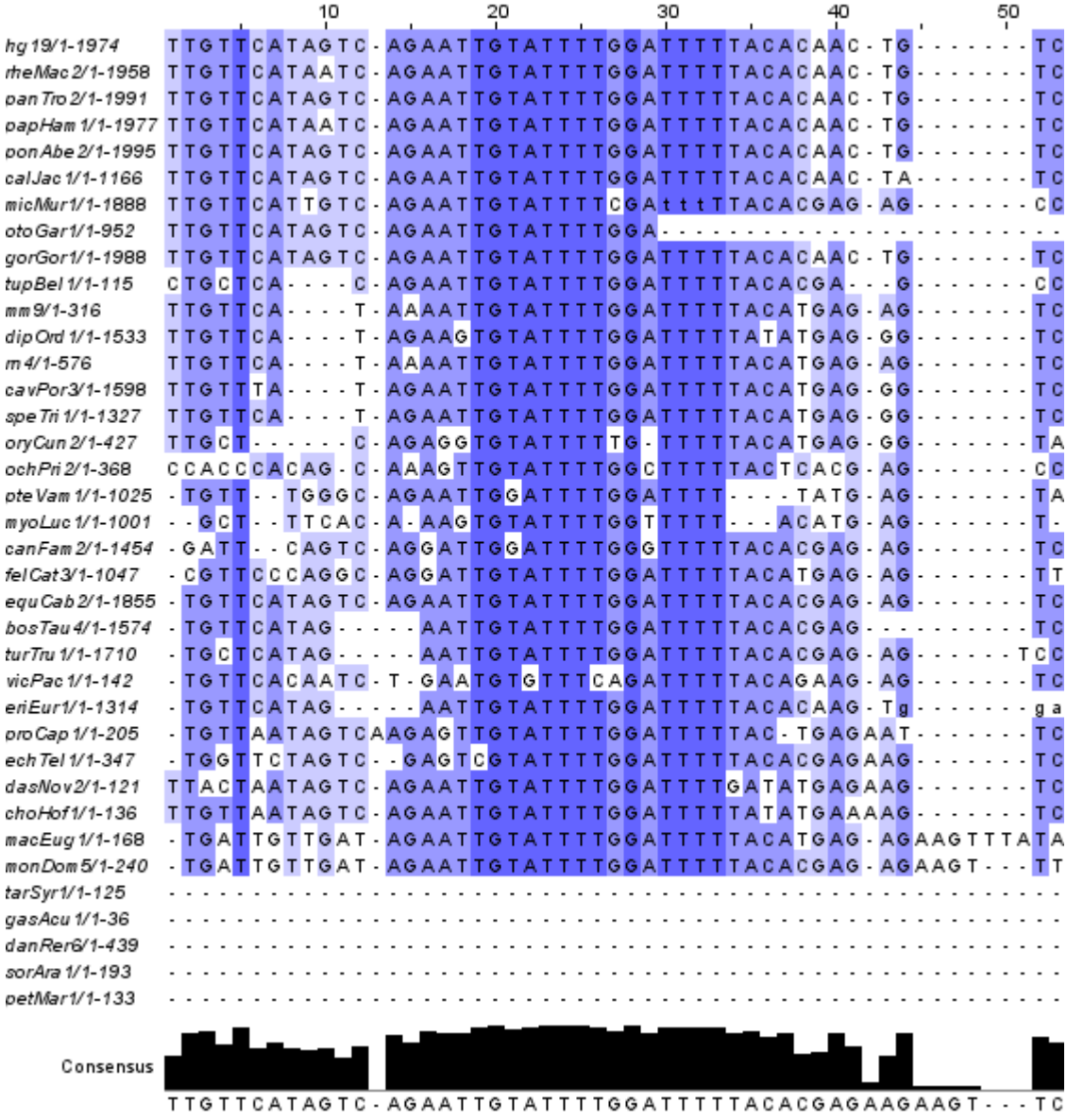


Figure 12: promoter length comparision between human and cow genome. The tinner line shows the promoter region compare to the exon region which is a thicher blue line. UCSC genome browser.

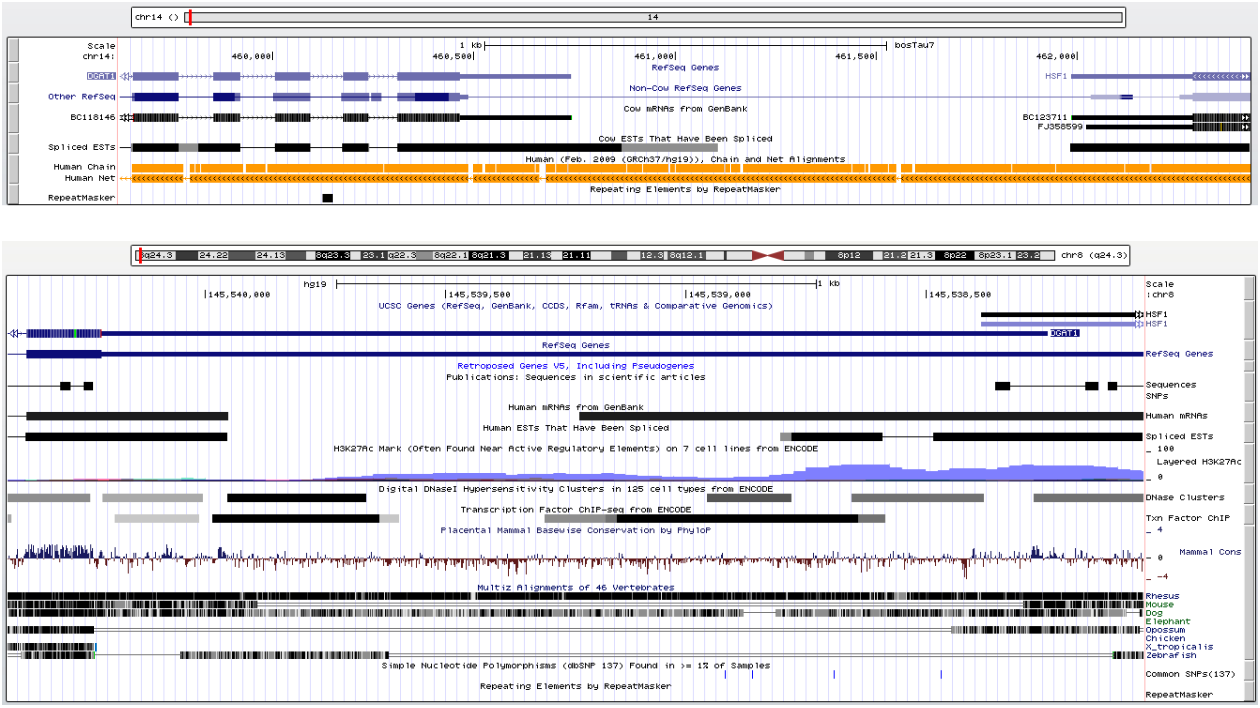


Figure 13: 97.2 %similarity between cattle and buffalo sequence (River) in 1.5 kb upstream of the DGAT1 gene. UCSC genome browser.

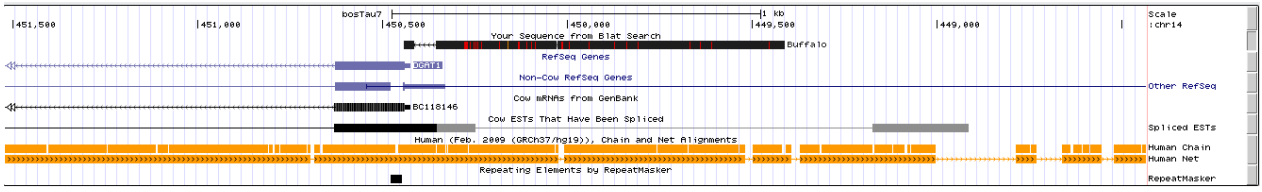


Figure 14: schematic view of the SNP position on MEME result.

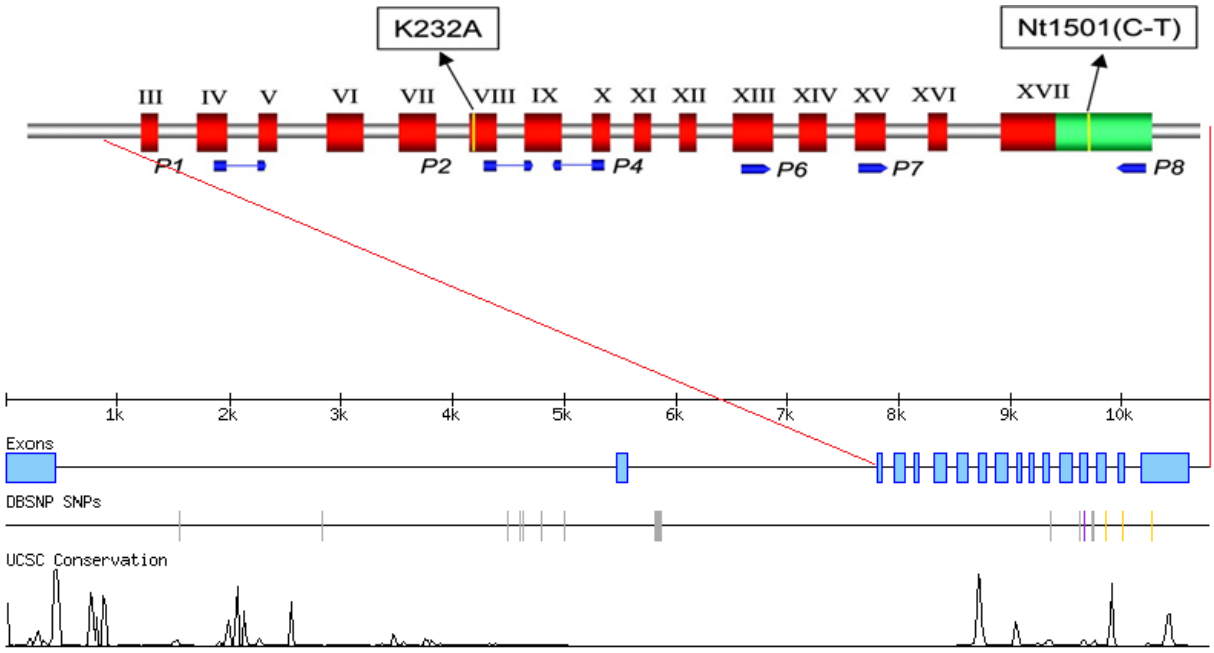
Figure 15: Alignment of the SNP position between different species from MEME database.

Name	Strand	Start	*p*-value	Sites
Swamp	+	2	4.79e-32	T TTGCTTCATTCAACGACTGGATGACTGCCGACTGTGCAATGAAACAGAAA
Buffalo	+	136	4.79e-32	GGTGTATTGG TTGCTTCATTCAACGACTGGATGACTGCCGACTGTGCAATGAAACAGAAA
cow	+	203	4.79e-32	GGTGTATTGG TTGCTTCATTCAACGACTGGATGACTGCCGACTGTGCAATGAAACAGAAA
River	+	2	7.78e-32	T TTGCTTCATTCAACGACTGGATGACTGCCTACTGTGCAATGAAACAGAAA
SNP	+	2	7.78e-32	T TTGCTTCATTCAACGACTGGATGACTGCCTACTGTGCAATGAAACAGAAA
Dolphin	+	84	1.36e-30	GGTGTATTAG TTGCTTCATTTAACGACTGGATGACTGCAGACTGTGCAATGACACAGAAG

SNP

River: Swedish River , Buffalo: online sequence,

Figure 16: K232A located in exon VIII of DGAT1 gene (mutdb.org)



6 Conclusion and future research

The *DGAT1* gene was mapped to the centromeric region of cattle chromosome 14 (Grisart et al., 2004). The importance of this gene in influencing milk fat percentage has been established in several studies. For review see Winter et al., 2002. The high potential of buffalo milk production is becoming increasingly important because of disease resistance and fat rich milk. Most researchers focus on the River buffalo because of the bigger global population compared to the exotic Swamp subtype which has a limited geographical distribution. Swamp buffalo has even higher milk fat percentage, but because of less concern they have low milk production, which traditionally have made them less interesting for use in industrial purposes. The first step which is already taken in this process is whole genome sequencing of River buffalo. If the Buffalo genome sequence be available it will open a new window to identify genetic factors with potential of this species as a production animal. Concerning the *DGAT1* gene we need high quality reference sequences of both gene body and promoter region from Swamp as well as River buffalo. Sequencing of *DGAT1* cDNA obtained from mRNA prepared from mammary gland and liver tissue of both subtypes will be important to identify and define the transcription start site and 5′ UTR region of the buffalo *DGAT1* mRNA. The *DGAT1* gene is located in a gene rich area in the genome. The nature and function of some of these genes have been studied but some genes still remain to be characterized. It is possible that the promoter of *DGAT1* is bi-directional and can co-regulate the neighboring gene. This can cause the milk fat percentage changes in different situation like heat shock or different environment. Electrophoretic mobility shift assay (EMSA) will be the next approach for finding the protein-DNA interaction in the SNP position that was identified in this study.

Proper recording system for Swamp buffalo is crucial in this process. Without recording of milk production and milk content it not possible to find the differences between/within population. Breeding system for Swamp buffalo is also important for success in this process. Improving of traits without strong breeding system is impossible. The identification of the SNP in the buffalo *DGAT1* promoter allows future studies to define the functional significance of this genetic variation.

7 Acknowledgement

I would like to thank my supervisor Göran Andersson for all his help during my master degree. He led me to pursue my dream from the beginning of my master degree and this project as well. I express my sincere gratitude to my other supervisor Erik Bongcam-Rudloff. I would say thanks to the following: people that helped me to do my project in IMBIM/HGEN lab at BMC, Ängsholmens Gårdsmejeri for cooperating in collecting Swedish buffalo samples, Awaisa Ghazal for helping me in bioinformatics analysis. I express my gratitude to my parents for all their support during my whole life, especially during my stay in Sweden.

Figure 1. Bovine subfamily (Marnoch Yindee, 2010)

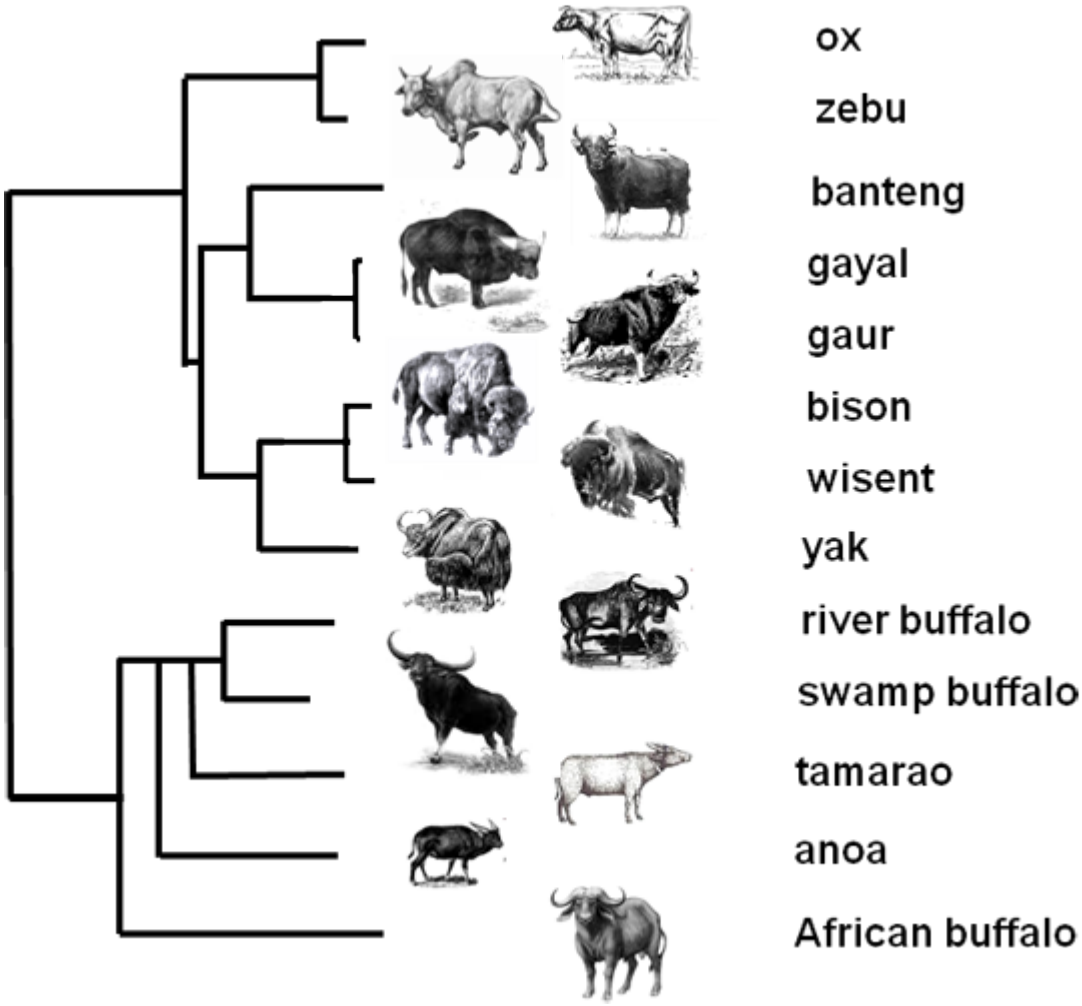


Figure 2. Murrah breed

provided by Dr. K.L. Dahiya

Figure 3. Chinese Buffalo breeds

Binhu buffalo

Enshi buffalo

Shanghai buffalo

Fuling buffalo

Fuzhong buffalo

Dongliu buffalo

(Borghese & Mazzi. 2005)

8 References:

Alexiev, A. 1998. The Water buffalo. St. Kliment Ohridski University Press.

Borghese, A., Di Giacomo, A., & Mormile, M. (1996). Meat eating quality in bovine and buffalo young bulls of one year age. Interna- tional symposium on buffalo products, 1±4 December, Paestum EAAP, publication no. 82.

Borghese, A and marco mazzi. 2005. Buffalo production and research. Fao regional office for europe inter-regional cooperative research network on buffalo (escorena). Reu Technical Series 67.

Borghese, A. 2010. Development and perspective of Buffalo and Buffalo market in Europe and Near East.

Borghese, A. and Moioli, B. 2000. Feasibility of introducing buffaloes into some African countries. Third All African Conf. Anim. Agric., Alexandria, Egypt, 6 to 9 Nov.

Borghese, A., Moioli, B. and Tripaldi, C. 2000. Buffalo milk: processing and product development in Mediterranean countries. Proc. Third Asian Buffalo Congress, Kandy, Sri Lanka, March 27-31: 37-46.

Bylund, G. (1995). Dairy Processing Handbook. Lund, Sweden, Tetra Pak Processing Systems AB.

Cases, S., Y. Zheng, H. Myers, S. Lear, E. Sande, S. Novak, C. Collins, C. Welch, A. Lusis, S. Erickson and R. Farese (1998). "Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis." Proceedings of the National Academy of Sciences 95: 13018 - 13023.

Chen, H., Z. Ladha, D. Jensen, L. Ferreira, L. Pulawa, J. McGuire, R. Pitas, R. Eckel and R. Farese (2002a). "Increase insulin and leptin sensitivity in mice lacking acyl CoA:diacylglycerol acyltransferase 1." Journal of Clinical Investigation 109(8): 1049 - 1055.

Cavaletto, M., M. Giuffrida and A. Conti (2004). "The proteomic approach to analysis of human milk fat globule membrane." Clinica Chimica Acta 347: 41 - 48.

Cases, S., Y. Zheng, H. Myers, S. Lear, E. Sande, S. Novak, C. Collins, C. Welch, A. Lusis, S. Erickson and R. Farese (1998). "Identification of a gene encoding an acyl CoA:diacylglycerol acyltransferase, a key enzyme in triacylglycerol synthesis." Proceedings of the National Academy of Sciences 95: 13018 - 13023.

Chi-Liang Eric Yen, Scot J. Stone, Suneil Koliwad,§, Charles Harris, and Robert V. Farese, Jr. DGAT enzymes and triacylglycerol biosynthesis. Journal of Lipid research. JLR Papers in Press, August 29, 2008.DOI 10.1194/jlr.R800018-JLR200

Dhanda, O.P. 2004. Developments in water buffalo in Asia and Oceania. Proc. of the Seventh World Buffalo Congress, Manila, Philippines, 20 to 23 Oct.: 17-28.

Di Berardino, D. and Iannuzzi, L. 1981. Chromosome banding homologies in Swamp and Murrah buffalo. J. Hered., 72: 183-188.

Eckel, R. (1989). "Lipoprotein lipase. A multifunctional enzyme relevant to common metabolic diseases." New England Journal of Medicine 320(16): 1060-1068.

Evers, J.M. (2004). "The milkfat globule membrane - compositional and structural changes post secretion by the mammary secretory cell." International Dairy Journal 14(8): 661-674.

Fikri, El-Kirabi, 1995. Buffalo population and production in Egypt. FAO/RNE, Buffalo Newsletter, 3: 8.

Grisart, B., F. Farnir, L. Karim, N. Cambisano, J. Kim, A. Kvasz, M. Mni, P. Simon, J. Frere, W. Coppieters and M. Georges (2004). "Genetic and functional confirmation of the causality of the DGAT1 K232A quantitative trait nucleotide in affecting milk yield and composition." Proceedings of the National Academy of Sciences 101: 2398 - 2403.

Georges, M., D. Nielsen, M. Mackinnon, A. Mishra, R. Okimoto, A. Pasquino, L. Sargeant, A. Sorensen, M. Steele and X. Zhao (1995). "Mapping quantitative trait loci controlling milk production in dairy cattle and exploiting progeny testing." Genetics 139: 907 - 920.

Geldermann, H. (1975). "Investigations on inheritance of quantitative characters in animals by gene markers." Theoretical and Applied Genetics 46: 319-330.

Huth, P., D. DiRienzo and G. Miller (2006). "Major Scientific Advances with Dairy Foods in Nutrition and Health." Journal of Dairy Science 89: 1207-1221.

Hedges, S. 1996. Bubalus bubalis. In: IUCN 2007. IUCN Red List of Threatened Species. Gland, Switzerland

Heid, H. and T.W. Keenan (2005). "Intracellular origin and secretion of milk fat globules." European Journal of Cell Biology 84: 245-258.

Keenan, T.W. (2001). "Milk lipid globules and their surrounding membrane: A brief history and perspecitives for future research." Journal of Mammary Gland Biology and Neoplasia **6**(3): 365-371.

Lopez, C., V. Briard-Bion, O. Menard, F. Rousseau, P. Pradel and J.M. Besle (2008). "Phospholipid, Sphingolipid, and Fatty Acid Compositions of the Milk Fat Globule Membrane are Modified by Diet." Journal of Agricultural and Food Chemistry 56(13): 5226-5236.

Murti, T.W. dan G. Ciptadi. 1986. Kerbau Perah dan Kerbau Kerja: Tatalaksana dan Pengetahuan Dasar Pasca Panen. Ed.1, Cetakan 1. Mediyatama Sarana Perkasa. Yogyakarta.

Marshall, M.O. and J. Knudsen (1977). "Biosynthesis of triacylglycerols containing short-chain fatty acids in lactating cow mammary gland. Activity of diacylglycerol acyltransferase towards short-chain acyl-CoA esters." European Journal of Biochemistry 81(2): 259-266.

Meegalla, R.L., J.T.B. and D. Cheng (2002). "Concerted elevation of acyl-coenzyme A:diacylglycerol acyltransferase (DGAT) activity through independent stimulation of mRNA expression of DGAT1 and DGAT2 by carbohydrate and insulin." Biochemical and Biophysical Research Communications 298: 317-323.

Neville, M. and M. Picciano (1997). "Regulation of Milk Lipid Secretion and Composition." Annual Review of Nutrition 17: 159 - 184.

Pasha. T. N, Hayat. Z, 2009. present situation and future perspective of buffalo production in Asia. The journal of animal and plant sciences, 22(3 suppl.): 2012, page: 250-256. issn: 1018-7081

Patton, S. (2004). Milk: Its Remarkable Contribution to Human Health and Well-Being. New Brunswick, New Jersey, Transaction Publishers.

Patil, G.R. and Nayak, S.K. 2003. Competitive advantages of using buffalo product manufacture. Proc. Fourth Asian Buffalo Congress, New Delhi, India, Feb. 25-28: 183-190.

Ribadeau-Dumas, B. and R. Grappin (1989). "Milk protein analysis." Lait 69: 357.

Soliman I., 2009. Personal Communication by Antonio Borghese. Development and perspective of Buffalo and Buffalo market in Europe and Near East.

Singh, H. (2006). "The milk fat globule membrane - A biophysical system for food applications." Current Opinion in Colloid and Interface Science 11: 154-163.

Sethi, R.K, 2003. Improving riverine and swamp buffaloes through breeding. Proc. of the Fourth Asian Buffalo Congress, New Delhi, India, 25 to 28 Feb.: 51-60.

Spitsberg, V.L. (2005). "Invited Review: Bovine milk fat globule membrane as a potential nutraceutical." Journal of Dairy Science 88(7): 2289-2294.

Siregar, A. R. dan C. Talib. 1991. Menggali Potensi dan Pengembangan Kerbau di Indonesia. Prosiding Seminar Nasional Usaha Peningkatan Produktivitas Peternakan dan Perikanan. Fakultas Peternakan Universitas Diponegoro. Semarang. Hal.77- 83.

Smaragdov, M. (2006). "Genetic Mapping of Loci Responsible for Milk Production Traits in Dairy Cattle." Russian Journal of Genetics 42: 1 - 15.

Shorten, P., T. Pleasants and G. Upreti (2004). "A mathematical model for mammary fatty acid synthesis and triglyceride assembly: the role of stearoyl CoA desaturase (SCD)." Journal of Dairy Research 71: 385 - 397.

Tassell, C. (2003). Genetic Selection Tools in the Genomics Era.

Wiryosuhanto, S. D., Purwandariyanto dan W. Ediyati. 1980. Peternakan Kerbau di Indonesia. Deptan. Jakarta.

Walstra, P., T.J. Geurts, A. Noormen, A. Jellema and M.A.J.S. v. Boekel (1999). Dairy Technology. Principles of milk properties and processes. New York, NY, Marcel Dekker AG.

Weiss, S., E. Kennedy and J. Kiyasu (1960). "The enzymatic synthesis of triglycerides." Journal of Biological Chemistry 235: 40 - 44.

Walstra, P., T.J. Geurts, A. Noormen, A. Jellema and M.A.J.S. v. Boekel (1999). Dairy Technology. Principles of milk properties and processes. New York, NY, Marcel Dekker AG.

Winter, A., W. Kramer, F. Werner, S. Kollers, S. Kata, G. Durstewitz, J. Buitkamp, J. Womack, G. Thaller and R. Fries (2002). "Association of a lysine-232/alanin polymorphism in a bovine gene encoding acyl-CoA:diacylglycerol acyltransferase (DGAT1) with variation at a quantitative trait locus for milk fat conent." Proceedings of the National Academy of Sciences 99: 9300 - 9305.

Yang Bingzhuang, Li Zhongquan, Liang Xianwei and Zou Caixia, 2003. The advance of genetic improvement and the development of the dairy industry in the Chinese Water buffalo. Proc. Of the Fourth Asian Buffalo Congress, New Delhi, India, 25 to 28 Feb.: 27-30.

Yindee Marnoch., Reproduction and genetic diversity of the swamp buffalo. Veterinary Sciences Tomorrow - 10 January 2011

http://faostat.fao.org/site/569/default.aspx#ancor

http://chartsbin.com/view/1482

www.save-foundation.net

http://lipidlibrary.aocs.org/lipids/tag2/index.htm

Table7. River buffalo breeds and performance traits (Borghese, 2010, Sethi, 2003, FAO 2003)

breed	country	Lactation duration (days)	Milk yield(kg)	Milk fat (%)	Milk protein (%)	Average body weight female (kg)	Age at first calving (months)	First lactation 305 days or less yield (kg)	First lactation total yield (kg)	All lactation 305 days or less yield (kg)	All lactation total yield	All lactation length (days)	Average fat (percent)	Average dry period (days)	Service period (days)	Calving interval (days)
Anatolian	Turkey	220-270	700-1 000	6.6-8.1	4.2-4.6	*	*	*	*	*	*	*	*	*	*	*
Azeri or Caucasian	Indo valley (Indian buffalo)	200-220	1 200-1 300	6.6	*	*	*	*	*	*	*	*	*	*	*	*
Bangladeshi	Bangladesh	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Bhadawari						385.5	48.6±0.58	711±25	781±29	812±23	272±4			297±24	179±10	478±11
Bulgarian Murrah	Bulgaria	270-305	1 800	7.04	*	*	*	*	*	*	*	*	*	*	*	*
Egyptian	Egypt	210-280	1 200-2 100	6.5-7.0	*	*	*	*	*	*	*	*	*	*	*	*
Jafarabadi	Gujarat (India)	350	1 800-2 700	8.5		529±13	1 925±196	1 642±283	1 642±283	1 950±79	2 097±110	320.1±11.6	7.7±1.0	159.8±10.9	161.5±14.0	509.8±20.1
Jerangi	border of Orissa with Andhra	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
Kuhzestani or Iraqi	South of Iran and Iraq	200-270	1 300-1 400	6.6	*	*	*	*	*	*	*	*	*	*	*	*
Kundi	Pakistan	320	2 000	7	6	*	*	*	*	*	*	*	*	*	*	*
Lime	Nepal	351	875	7	*	*	*	*	*	*	*	*	*	*	*	*
Manda	Orissa with Andhra Pradesh		4 kg/day	*	*	*	*	*	*	*	*	*	*	*	*	*
Mediterranean	Europe	270	900-4 000	8	4.2-4.6	450-650	*	*	*	*	*	*	*	*	*	*
Meshana	Gujarat, India	305	1 800-2 700	6.6-8.1	4.2-4.6	430	*	*	*	*	*	*	*	*	*	*
Murrah	India	305	1 800	7.2	*	495	50.6±2.0	1 894±44	*	2 183±136	2 226±152	305±16	6.7	144±26	146±27	479±33
Nagpuri	India	243	825	7	*	*	*	*	*	*	*	*	*	*	*	*
Nili-Ravi	Pakistan	305	2 000	6.5		546	39.97	1 565	1 571	1 946	1 969	299	7.1	131	151	443
Parkote	Nepal	351	875	7	*	*	*	*	*	*	*	*	*	*	*	*
Sambalpuri	Bilaspur, India	350	2 400	*	*	*	*	*	*	*	*	*	*	*	*	*
Surti	Gujarat (India)	350	2 090	6.6-8.1	4.2-4.6	462±7.0	53.2±1.7	1 295±57		1 477±42	1 547±50	311±7	8.1	234±21	207±17	510±16
Tarai	India	250	450	6.6-8.1	4.2-4.6	*	*	*	*	501±10.6	*	*	8.22±0.08	*	*	15.74±0.4
Toda	Madras	200	500	*	*	*	*	*	*	*	*	*	*	*	*	*

*: Data not available
Some breeds are really small in population, so those data come from just one experiment.
Breed names are listed alphabetically.