EXCELLENT PUBLISHERS

Original Research Article

Identification of SNP in Lipoprotein Lipase Gene and its Validation and Association with Milk Production in Goats

K. Mathivathani*, G. Radhika and T. V. Aravindakshan

Department of Animal Genetics and Breeding, College of Veterinary and Animal Sciences, Mannuthy, Thrissur – 680 651, India

*Corresponding author

A B S T R A C T

In Kerala, goat farming forms an important source of income and nutritional security to the rural poor and unemployed sections of society. Milk production, is one of the important economic trait in goats. In Kerala, two native goat breeds, Malabari and Attappady Black vary significantly in milk production. Double digest Restriction Associated DNA Sequencing (ddRADseq) is a modern approach for identification of Single Nucleotide Polymorphisms (SNP) across populations. This ddRADseq technique was performed in three pooled samples (ten each) of high and low milk producing Malabari and Attappady Black goats. The SNPs discovered were analysed using bioinformatics tools. A total of 42 candidate genes affecting milk quantity and quality were selected from literature and the entire data was screened for variants in these candidate genes. Out of these candidate genes, LPL gene (*LPL*)was selected for validation in the present study as researches showed that *LPL* polymorphism was found to be significantly associated with milk yield (kg) in goats. In the current study, Intronic variant c.243 + 1082C>T *LPL,* was validated using Polymerase Chain Reaction-Single Strand Conformation Polymorphism (PCR-SSCP) and further sequencing. It had significant (p≤0.05) association with daily average milk yield and CT (494.72 ± 11.60 mL) genotype showed higher milk yield. Current study emphasized the relevance of ddRADseq in identifying novel SNP in candidate genes influencing economic traits in goats.

Introduction

India has a goat population of 148.88 million, which is about 20 per cent of the total global goat population. There are 34 well recognized breeds of goats in India. Kerala, the southernmost state of India, possess two indigenous goat breeds namely, Malabari and Attappady Black. In India, annual milk production from goats was reported to be 15.2 million metric tonnes, and they provide income throughout the year through milk, meat, skin, manure etc., since ancient times, goat milk has gained importance in traditional medical practices and is believed to play a notable role in enhancing human health. Goat milk is usually recommended for aged persons, infants and convalescent people due to its better digestibility, which can be attributed to smaller diameter of fat globules

found in goat milk when compared to cow milk. Recent advances in molecular techniques provide opportunities for determining the order of nucleotides in DNA. It is considered that sequencing methods were important for SNP discovery, subsequent genotyping and for further association studies with economic traits. Modern advancements in next generation sequencing (NGS) methods provide cost effective sequencing of whole or partial genomes.

In reduced representation sequencing, representative portion of genome is sequenced. Modified reduced representation sequencing method, called double digest Restriction Associated DNA sequencing (ddRADseq) uses two restriction enzymes for skipping the random shearing step of RAD Sequencing. Gradually ddRADseq acquired important role in high density SNP discovery and genotyping, in model and non-model organisms with high efficiency and low cost.

Materials and Methods

Selection of animals for DDRADSEQ

Blood samples were collected from 10 each of high milk producing Malabari, low milk producing Malabari and Attappady Black goats in University Goat and Sheep Farm, Mannuthy, Thrissur. Genomic DNA was extracted from whole blood using standard phenol chloroform method (Sambrook and Russell, 2001). The concentration, purity and quality of DNA were checked by Nanodrop TM 1000 spectrophotometer (Thermo Scientific, USA). The purity of DNA was verified by measuring absorbance at 260 nm and 280 nm. The DNA samples with value between 1.7-1.8 were used for further downstream processing. Library preparation was done with *SphI* and *MIucI* restriction enzymes and by using suitable bioinformatics software, SNPs were identified. Based on

ddRADseq results, a total of 4141 variants were identified in Malabari and 2192 variants in Attappady Black Goats. From the SNPs, 37 variants were identified in 42 candidate genes affecting milk production, and *LPL* gene was selected for further wet lab analysis. In current study, Intronic variant c.243 + 1082C> T of *LPL* identified by ddRADseq, was validated and association studies carried out by using Polymerase Chain Reaction - Single Strand Conformation Polymorphism (PCR-SSCP).

Isolation of dna

A total of 200 goats with milk production records were selected for validation of selected SNP in *LPL* gene discovered through ddRADseq. Blood Samples and milk production data were collected from 100 Attappady Black goats and 100 Malabari goats maintained at University Goat and Sheep Farm, Mannuthy, Thrissur. Isolation of DNA and quality check were done. Fragment of *LPL* (147bp) enclosing the selected variants were amplified using specific set of primers, forward sequence AACCACACTCTTCT GCAAGG and reverse sequence CAGCAGA CCGCATCTCACTA which were designed from template sequences available at GenBank, using online tool Primer3. PCR amplification and agarose gel electrophoresis was done.

The concentration, purity and quality of DNA were checked by NanodropTM 1000 spectrophotometer (Thermo Scientific, USA). The purity of DNA was verified by measuring absorbance at 260 nm and 280 nm. The DNA samples with value between 1.7-1.8 were used for further downstream processing. Single Strand Conformation Polymorphism (SSCP) and Polyacrylamide Gel Electrophoresis (PAGE) of (147 bp) LPL Fragment were done. Fragment of *LPL* (147bp) showed three different banding

patterns. For confirmation, PCR products from different banding patterns were selected and outsourced for sequencing. Sequencing was performed by automated sequencer using Sanger's dideoxychain termination method at AgriGenom Labs Pvt. Ltd.

Statistical analysis

Popgene 32 Version 1.32 was used to estimate the genotype and allele frequencies. Milk quantity and association of genotypes with milk production and the observed genotypes were analysed using fixed General Linear Model (GLM) of SPSS V.24. Posthoc

tests (Least Significant Difference (LSD) and Duncan's Multiple Range Tests (DMRT)) were used to identify homogeneous subsets.

Results and Discussion

Validation studies for selected SNPS

Figure 1indicates the concentration, purity and quality of DNA, which were checked by NanodropTM 1000 spectrophotometer. The PCR-SSCP was conducted to find out SNP in intronic variant of caprine *LPL* gene, obtained from ddRADseq. The147bp fragment of amplified PCR products are given in Plate 1.

Table.1 Allelic and genotypic frequencies of SNPs in *LPL*

Parameter	Malabari (N=100)	Attappady Black(N=70)	Total (N=170)
Genotype frequency			
CC	0.35 (35)	0.61 (43)	0.46 (78)
CT	0.49 (49)	0.21 (15)	0.38 (64)
TT	0.16 (16)	0.18 (12)	0.16 (28)
Allele frequency			
C	0.59	0.72	0.65
T	0.41	0.28	0.35

Table.2 Least Square means of daily milk yield in three genotypes of *LPL*

Trait	Daily Average milk yield (mL)			
	Genotype	Genotype	Genotype	p-value
Milk yield (Total)	$437.12^{a} \pm 12.85$ (78)	$494.72^{c} \pm 11.60$ (64)	$467.63^{b} \pm 19.86$ (28)	0.03
Milk yield (Malabari)	$554.63^{b} \pm 7.43$ (35)	$542.64^{a} \pm 4.42$ (49)	$557.57^{b} \pm 12.22$ (16)	0.01
Milk yield (Attapady Black)	341.46 ± 5.25 (43)	338.20 ± 9.68 (15)	345.91 ± 6.81 (12)	0.8

Figures in parenthesis shows number of observations
Means with different superscripts in same rows differ significantly (*p-value≤0.05)

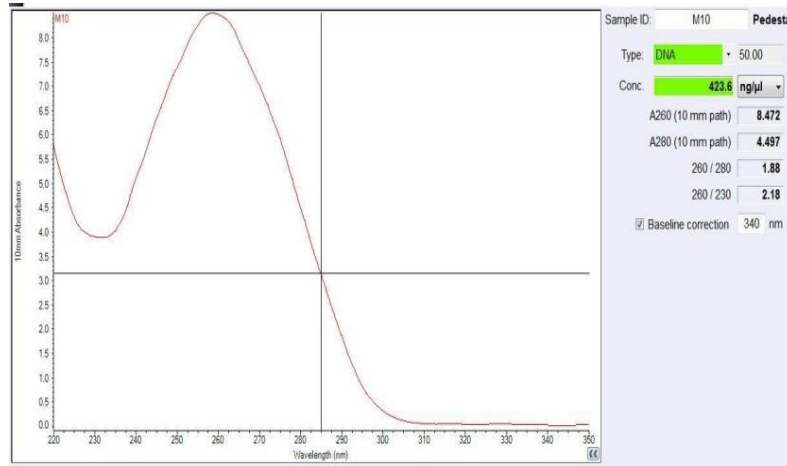**Fig.1** Quality and quantity of DNA measured in Nano drop
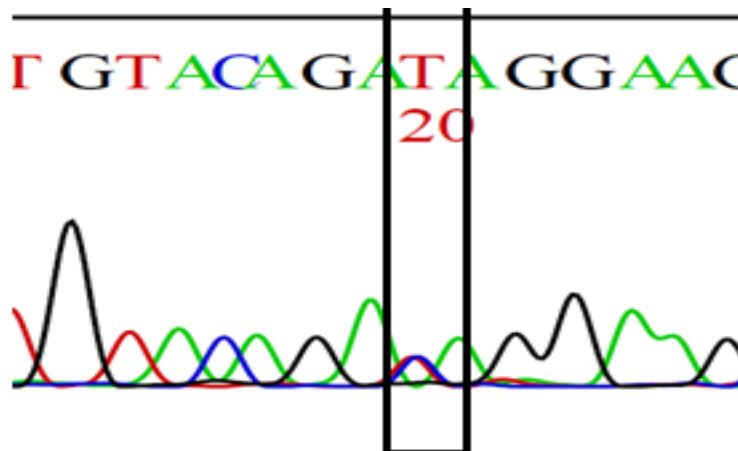

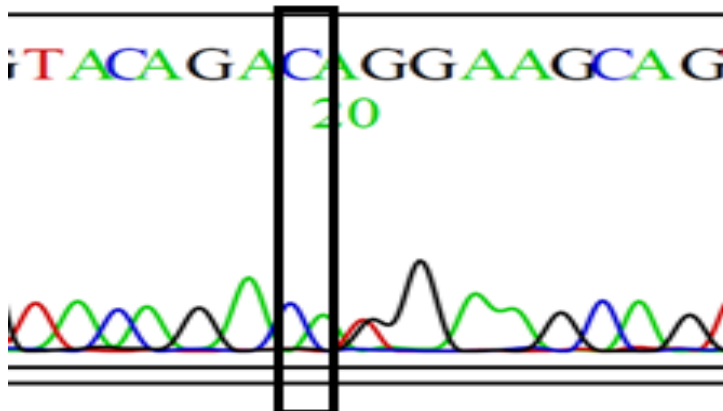**Fig.2** CT genotype (LPL)

Fig.3 CC genotype (LPL)
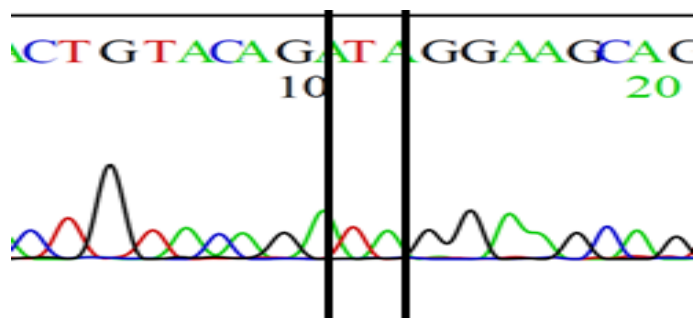
Fig.3 TT genotype (LPL)

Plate.1 PCR product of LPL fragment resolved in two per cent agarose gel
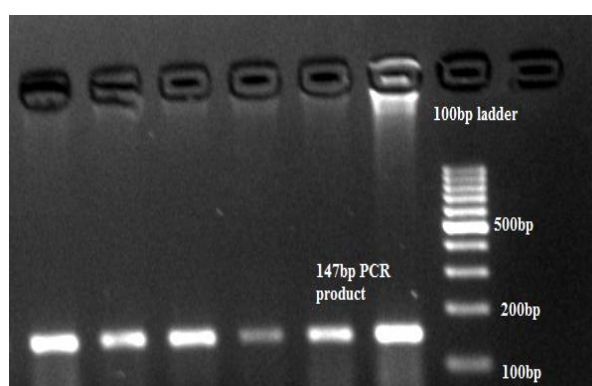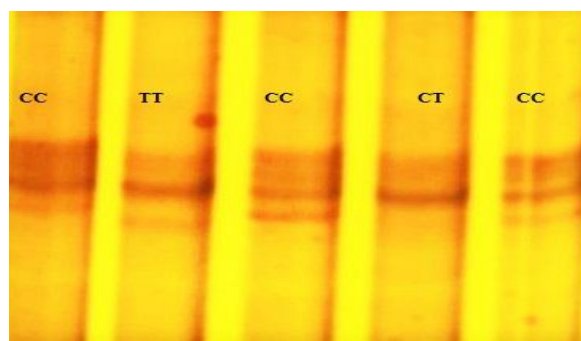

Plate.2 SSCP pattern of LPL fragment

The PCR-SSCP analysis of the 147bp amplified *LPL* product revealed three banding patterns (homozygous CC, homozygous TT and heterozygous CT) given in Plate 2. Amplicons obtained from PCR were sequenced and a snapshot of chromatogram depicting the CT, CC and TT genotypes are given in Fig 2, 3 and 4 respectively. Genotype and allelic frequencies of intronic variant in *LPL* is given in Table 1. Three genotypes CC, CT and TT *s*howed significant association with milk yield thereby indicating the influence of the SNP in milk production. The least square means with standard error for average milk yield for CC, CT and TT genotypes are presented in the Table 2.

ddRADseq

The RADseq has already been employed in

studies of allele frequency estimation (Van Tassell *et al.,* 2008), association studies (Parchman *et al.,* 2012), genetic mapping (Baird *et al.,* 2008; Andolfatto *et al.,* 2011). Peterson *et al.,* (2012) developed a new RAD-Seq platform using double digestion of genomic DNA with two Restriction enzymeswhich is used for SNP discovery and genotyping. Few candidate genes affecting milk production in goats were selected from the available literature. Out of these candidate genes, *LPL* was selected for validation study. Intronic variant c.243+1082C>T *LPL,* was validated using PCR-SSCP.

Association and validation studies

The *LPL* gene intronic variant c.243 + 1082C> T was identified in sixth intron of *LPL,* by ddRADseq. Lipoprotein Lipase is an enzyme in milk which was responsible for hydrolysis of fatty acids from triglycerides and phospholipids in milk. Intronic variant of *LPL* gene was selected from ddRADseq and PCR-SSCP was performed in 100 each of Malabari and Attappady Black goats. The results revealed all the three genotypes CC, CT and TT, thus validating the c.243 + 1082C> T variant in the studied goat population, discovered through ddRADseq. According to Yu *et al.,* (2016) ddRADseq can be successfully utilised for discovery of SNPs, in model and non-model species. In the current study, *LPL* gene polymorphism in noncoding region showed significant association with milk yield. Rychtarova *et al.,* (2017) discussed that *LPL* polymorphism in intronic region, was associated with milk yield, protein content and fat content in Czech white short haired goat breed.In the current scenario, researchers are more concentrating on SNPs in noncoding regions and GWAS revealed more association of SNPs in intergenic regions with biological traits than in coding region (Hrdlickova *et al.,* 2014). Further studies on this novel SNP identified through ddRADseq is essential.

In conclusion the ddRADseq is a robust, efficient and economical reduced representation sequencing technique, which surveys on representative samples from genome for discovery of high confidence SNPs across groups or populations. In the current study, ddRADseq is applied to highly milk producing Malabari and low milk producing Attappady Black goat groups in order to identify genetic variants in candidate genes affecting milk production. The *LPL* intronic variant c.243+1082C>T was selected for validation by PCR-SSCP and further sequencing. It had significant ($p \leq 0.05$) association with mean daily milk yield (mL) of 437.12 ± 12.85, 494.72 ± 11.60 and 467.63 ± 19.86 for CC, CT and TT genotypes respectively. Besides the discovery of large number of high confidence SNPs, present study emphasized the relevance of ddRADseq in identifying novel SNP in candidate genes and also promises the possibility of discovering SNPs in unexplored genes which might affect complex traits.

References

Andolfatto, P., Davison, D., Erezyilmaz, D., Hu, T.T., Mast, J., Sunayama-Morita, T. and Stern, D.L., 2011. Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Res.*, 21:610-617.

Baird, N.A., Etter, P.D., Atwood, T.S., Currey, M.C., Shiver, A.L., Lewis, Z.A., Selker, E.U., Cresko, W.A. and Johnson, E.A., 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS one.*, 3(10):3376- 3382.

Hrdlickova, B., De Almeida, R.C., Borek, Z. and Withoff, S., 2014. Genetic variation in the non-coding genome: Involvement of micro-RNAs and long non-coding RNAs in disease. *Biochimica et Biophysica Acta (BBA)- Molecular Basis of Disease.*, *1842*:1910-1922.

Parchman, T.L., Gompert, Z., Mudge, J., Schilkey, F.D., Benkman, C.W. and Buerkle, C.A. 2012. Genome- wide association genetics of an adaptive trait in lodgepole pine. *Mol. Ecol.*, 21: 2991-3005.

Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S. and Hoekstra, H.E. 2012. Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PloS one*. [Online] 7 (2) Available: https://journals. plos.org/ plosone/article id 10.1371 / journal.pone.0131572 [12 Nov 2017].

Rychtarova, J., Sztankoova, Z. and Svitakova, A. 2017. Association of polymorphism at BTN1A1, SCD and LPL gene on somatic cell count in Czech White shorthaired goat breed. *J. Hyg. Engng. Design*., 21: 64-69.

Sambrook, J. and Russell, D.W. 2001. Molecular cloning: A laboratory manual. (3 rd Ed.). Cold spring Harbor Laboratory Press, New York, 1886p.

Van Tassell, C.P., Smith, T.P., Matukumalli, L.K., Taylor, J.F., Schnabel, R.D., Lawley, C.T., Haudenschild, C.D., Moore, S.S., Warren, W.C. and Sonstegard, T.S. 2008. SNP discovery and allele frequency estimation by deep sequencing of reduced representation libraries. *Nat. Methods*., 5: 247- 254.

Yu, H., You, X., Li, J., Liu, H., Meng, Z., Xiao, L., Zhang, H., Lin, H.R., Zhang, Y. and Shi, Q. 2016. Genome-wide mapping of growth-related quantitative trait loci in orange-spotted grouper (*Epinephelus coioides*) using double digest restriction-site associated DNA sequencing (ddRADseq). *Int. J. Mol. Sci*., 17:501- 509.