

Transcriptome Sequencing of Guilan Native Cow in Comparison with *bosTau4* Reference Genome

Research Article

M. Moridi¹, S.H. Hosseini Moghaddam^{1*} and S.Z. Mirhoseini¹¹ Department of Animal Science, Faculty of Agricultural Science, University of Guilan, Rasht, Iran

Received on: 3 Sep 2016

Revised on: 4 Nov 2016

Accepted on: 15 Nov 2016

Online Published on: Jun 2017

*Correspondence E-mail: hosseini@guilan.ac.ir

© 2010 Copyright by Islamic Azad University, Rasht Branch, Rasht, Iran

Online version is available on: www.ijas.ir

ABSTRACT

RNA-sequencing is a new method of transcriptome characterization of organisms. Based on identity and relatedness, there are large genetic variations among different cattle breeds. The goal of the current study was to sequence the transcriptome of Guilan native cow and compare with available reference genome using RNA-sequencing method. Blood samples were collected from 14 Guilan native cows and then were pooled with same ratios of 3 micrograms per sample. Sequencing of the pooled sample was carried out using Illumina Hiseq 2000 from both end and 100 base pair of reading length. Tophat2 software was used to align the reads with reference genome and identify splice junctions and insertions and deletions. Cufflinks software was used to assemble transcripts and calculate their abundances. Total numbers of sequenced RNA fragments were 28434708 and the overall reading map was 87.4 percent. Total numbers of expressed genes were 24616 genes, which 19994 genes from these were protein coding genes and 3825 genes were non-coding. Adenosine triphosphate synthase 6 (*ATP6*) and ribosomal protein, large, P1 (*RPLP1*) genes showed the highest abundances of all expressed genes. The majority of highly expressed genes were involved in ribosomal structures and translational activities; moreover, they were belonging to housekeeping genes. The current study is a report of leukocytes transcriptome sequencing of Guilan native cow which have been not reported, so far. As Guilan native cow has the biggest population among all native populations in Iran, such studies could help to evaluate the genetic potential of this high precise genetic resource in Western Asia.

KEY WORDS cDNA library, Guilan native cattle, Illumina Hiseq 2000, RNA-sequencing.

INTRODUCTION

Cows could be classified in two major groups including taurine breeds (*Bos taurus*) and indicine breeds (*Bos indicus*) (Zhang *et al.* 2015). The indicine breeds belong to tropical regions where natural selection and adaptation to harsh environmental conditions caused them superior in this environment due to disease resistance and food shortage (Wilson, 2009). Iranian native cows are classified in the indicine breeds and could be grouped in the six groups including Sarabi, Golpaigani, Sistani, Dashtyari, Najdi and

Guilan native cows or Taleshi (Iran's country report on farm animal, 2004). Based on the report of Jihad-Keshavarzi organization of Guilan province in 2015 there are 312107 Guilan native cows in Guilan province that includes 70 percent of all cows in this province. According to the reports of Iranian national animal breeding center and promotion of animal products (INABC-PAP), average of test-day milk yield in Iranian Holstein, crossbred, and native cattle populations are 28.1, 10.3 and 4.68 kg/day, respectively. Because of low milk and meat production in Iranian native cattle populations, the crossbreeding pro-

grams have been started around 60 years ago. The basal structure for genetic improvement of Iran cattle populations, such as pedigree registration, recording the traits and artificial insemination has been organized since 50 years ago.

During the last 15 years, while the number of crossbred cattle has increased rapidly (e.g. from 2425000 in 2002 to 4373000 in 2009), the number of native cattle decreased (e.g. from 4337000 in 2002 to 2915000 in 2009). The number of exotic purebred cattle were grown slowly (e.g. from 683000 in 2002 to 961000 in 2009) and keeping exotic purebred cattle had less propensity specially for local breeders, due to management, environmental variation, and feed resource factors (Kamalzadeh *et al.* 2008).

The indigenous cattle breeds play an important economic role in the rural areas by providing milk and meat (Nimbkar *et al.* 2008). However, the production of their milk is too low compared to *Bos taurus* breeds. Due to limitations of forage crop production in Guilan province and genetic resistance of native cattle, still there is an increasing interest in keeping of native cattle and also crossbred of native cattle with taurine (*Bos taurus*) including Holstein and Simmental purebreds. Furthermore, entrance of purebred exotic breed did not succeed in this region due to feed shortages. Therefore, native breeds as an important genetic resource in livestock should be conserved without crossbreeding. A large pool of genetic resource of indigenous breeds justifies the importance of their conservation at the global level (Rewe *et al.* 2015). Approximately, 8 percent of reported livestock breeds have become extinct and an additional 21 percent are considered to be at risk of extinction. Moreover, the situation is presently unknown for 35 percent of breeds, most of which are reared in developing countries (FAO, 2011).

RNA-sequencing (RNA-seq) is a new method for transcriptome characterization of all organisms. This method changed the view of eukaryotic transcriptome complexity and components. RNA-seq provides more accurate and expanded measurement of transcripts and different isoforms of them compared to other methods (i.e. real-time PCR, microarray).

The transcriptome is the complete set of transcripts in a cell, and their quantity, for a specific developmental stage or physiological condition. Understanding the transcriptome is essential for interpreting the functional elements of the genome and revealing the molecular constituents of cells and tissues, and also for understanding development and disease.

RNA-seq is the most powerful method to evaluate genes expression and determine new transcripts, splice junctions and nucleotide variations in RNA sequences (Wang *et al.* 2009). Generally, in RNA-seq method a population of RNA

(total RNA or a portion of total RNA like RNA's with poly-A tail) is transformed into cDNA library with single or both end attached adaptors. Using high-throughput manner each cDNA molecule is sequenced to obtain short sequences. Sequenced reads have length between 30 to 400 base pair. After sequencing, the read results were aligned. The reads were put together to draw genome map includes transcriptome structure and translation levels of genes (Wang *et al.* 2009).

Base on the origin (*Bos taurus* or *Bos indicus*) of different cow breeds there are large genetic variation among different cow breeds (The Bovine HapMap Consortium, 2009). Regardless of large studies on the field of cow transcriptome, the information about different genes expression in various populations and breeds are minor (Huang *et al.* 2012).

Because of the importance of gene expression in shaping the phenotype, understanding of transcriptomic differences of various breeds are important. Although many molecular studies have been performed on Iranian Holstein and native cattle breeds (Ebrahimi *et al.* 2015a; Ebrahimi *et al.* 2015b; Kharrati Koopaei *et al.* 2012; Mohammadabadi *et al.* 2010; Pasandideh *et al.* 2015), but the transcriptome of Guilan native cow has not studied, hence the goal of the current study was sequencing the transcriptome of Guilan native cow and its comparison with available cow reference genome using blood total RNA-seq method.

MATERIALS AND METHODS

The blood samples were obtained from the tail vein of 14 Guilan native cows from station located in Hossein-Kooh of Foman city (Guilan province, Iran). Samples were collected in heparinized venoject tubes and were transferred on ice (4 °C) to the laboratory immediately for RNA extraction. The time interval between collecting the samples and RNA extraction was less than one hour. At least, 4 mL of blood was used for RNA extraction from each sample. First, buffy coat of blood samples was isolated after centrifugation at 2000 × g for 10 min at 4 °C. White blood cells (WBC) were washed with red blood cell (RBC) lysis buffer (1X) twice and centrifuged at 2000 × g for 10 min at 4 °C to obtain white pellet of WBC. Total RNA was extracted using Trizol (Invitrogen) method from WBC (Jiang *et al.* 2013).

The quality and quantity of extracted RNA samples was determined using NanoDrop-2000 (Thermo Scientific) spectrophotometer based on 260 to 280 spectrophotometer ratio and electrophoresis on agarose gel. The same concentration of each sample (3 micrograms/sample) were pooled together to obtain biological average of Guilan native population.

The inherent biological variance between different samples is an important factor to make sound conclusions. Sample pooling will allow us to average out this variance. The quality of pooled sample was assessed using Agilent 2100 Bioanalyzer device and Agilent RNA 6000 Nano kit. Total RNA was digested by DNaseI (NEB) and purified by oligo-dT beads (Dyna beads mRNA purification kit, Invitrogen, USA), then poly (A)-containing mRNA was fragmented into 130 base pair with first strand buffer. First-strand cDNA was synthesized using N6 primer, first strand master mix, and super script II reverse transcription (Invitrogen, USA). Reaction conditions were 25 °C for 10 min, 42 °C for 40 min, and 70 °C for 15 min. The second strand master mix was added to synthesize the second-strand cDNA (16 °C for 1 h). The cDNA was purified using Agencourt® AMPure® XP PCR purification kit (Agencourt, USA). The final library was quantitated in two ways for validation. First, the average of molecule length was determined using the Agilent 2100 Bioanalyzer instrument (Agilent DNA 1000 Reagents), and then the quantification of the library was performed by real-time quantitative PCR (RT-qPCR) using TaqMan Probes. The qualified libraries were amplified on cBot to generate the cluster on the flow cell (TruSeq PE cluster kit V3-cBot-HS, Illumina). The amplified flow cell was sequenced, pair ended, on the HiSeq 2000 device (TruSeq SBS KIT-HS V3, Illumina) with read length of 100 base pair (BGI-Shenzhen, China). The whole dataset is available in NCBI with GSM2101067 GEO accession number.

To assess the quality of sequenced sample we used FastQC software before any analysis. Tophat2 with Bowtie2 (version 2.0.4) was used to align mRNA sequence reads to the reference genome (bosTau4), segment mapping algorithm to discovering splice junctions, and indel search to discover insertions and deletions (Kim *et al.* 2013). The accepted hits outputs of Tophat2 were applied to Cufflinks for assemble transcripts and estimation of their abundances (Trapnell *et al.* 2010). Cufflinks estimates gene expression (fragments per kilo base exon per million mapped fragments, or FPKM). Reference Ensembl *Bos taurus* annotation file was used as guide for assembling transcripts, bias correction using reference genome, and multi-read correction. The minimum for cutoff of abundance was defined 0.1.

RESULTS AND DISCUSSION

In this research, all samples of each population have been pooled to optimize biological differences of samples. Decreased biological variation and increased detection power of differentially expressed genes are the results of samples pooling.

The pooling bias can occur via difference between the calculated value in pooled sample and the average value in each single sample (Rajkumar *et al.* 2015).

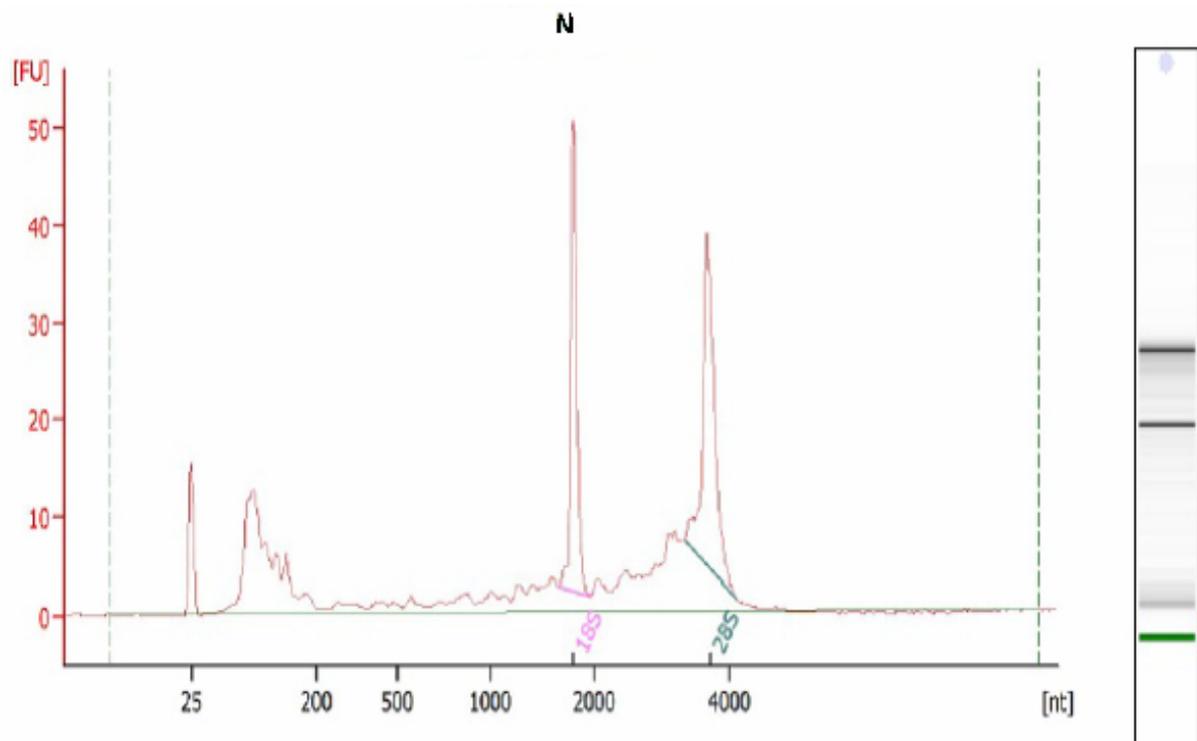
Extracted RNA's were shown the average of 350 nanograms per microliter concentration and average of 1.9 for 260/280 ratio. Using Bioanalyzer device, the RIN (RNA Integrity Number) for pooled sample was 8 which showed us a good quality of sample for RNA-seq (Figure 1).

The quality assessment of sequenced sample using FastQC software showed the high sequencing quality in both forward and reverse directions. Summary of mapping reads using reference genome are available in Table 1. Total sequenced fragments and overall read mapping rate were 28434708 and 87.8%, respectively. Numbers of insertions, deletions, splice junctions and aligned pairs with concordant alignments were 144258, 140262, 192100 and 18234548, respectively. Huang *et al.* (2012) reported 21078477, 21358931 and 20940063 total sequenced fragments and 64.4, 67.3 and 78.3 overall read mapping rate for Holstein, Jersey, and Cholistani breeds, respectively. The Holstein and Jersey are *Bos taurus* origin, while Cholistani same as Guilan native cow is classified as *Bos indicus* breeds.

Generally, we observed 24616 expressed genes, from which 19994 genes were protein coding genes and 3825 genes were classified in non coding genes. These non coding genes produce non translated transcripts and have regulatory function in genes expression (Table 2). The results showed us 26740 gene transcripts for Guilan native cow population. The total number of 13409, 13787 and 13666 expressed genes were reported in blood leukocytes transcriptome for Holstein, Jersey, and Cholistani populations, respectively (Huang *et al.* 2012).

The majority of highly expressed genes were housekeeping genes. Housekeeping genes are involved in basic cell maintenance and are expected to maintain constant expression levels in all cells and conditions (Eisenberg and Levinson, 2013). The five more expressed annotated transcripts are represented in Table 3. The majority of these genes categorized in housekeeping genes. The adenosine triphosphate synthase 6 (*ATP6*) and ribosomal protein, large, P1 (*RPLP1*) transcripts were showed highest expression among all the other transcripts. The ATP6 is one subunit of ATP synthase in mitochondria that is responsible for final step of oxidative phosphorylation (Habersetzer *et al.* 2013).

The ATP6 gene transcript showed higher differences with the other genes transcripts. *RPLP1* gene encodes a ribosomal phosphoprotein that is a component of 60S subunit of ribosomes (Martinez-Azorin *et al.* 2008). B2M gene encodes beta-2-microglobuline protein, a component of MHC class I molecules (Hall *et al.* 2015).



Overall Result for sample : **N**

RNA Integrity Number (RIN): 8

Fragment table for sample : **N**

Name	Start Size [nt]	End Size [nt]
18S	1,648	1,958
28S	3,346	4,128

Figure 1 Electropherogram of pooled sample from Bioanalyzer device

All 28S, 18S, and 5S bands absorptions and lengths and their gel qualities are obvious

The vertical axe represents the light absorption of nucleotides and the horizontal axe shows the nucleotides length (N= Guilan native cow population)

The right side is the agarose gel electrophoresis of pooled sample

The *B2M* gene was the third highest expressed genes in our case (Table 3). The NADH dehydrogenase 2 (ND2) protein is a subunit of NADH dehydrogenase (ubiquinone), which is located in the mitochondrial inner membrane and is the largest of the five complexes of the electron transport chain (Zickermann *et al.* 2015). MicroRNAs (miRNAs) are small non-coding RNAs of approximately 19 to 25 nucleotides that post-transcriptionally regulate gene expression. MiRNAs bind to the UTR (3'-untranslated region) of their target mRNAs messenger (RNAs) through complementary recognition, which then leads to degradation or repression of protein expression (Yan *et al.* 2014).

Much micro-RNAs have been found to involve in the physiological and pathological processes of angiogenesis (Suarez *et al.* 2008).

Among the top transcripts there was also the micro RNA MiR-126 which is a specific and highly expressed micro-RNA in vessel endothelial cells and has key roles in controlling arteriogenesis (increase in the diameter of existing arterial vessels) and angiogenesis (the physiological process through which new blood vessels form from pre-existing vessels) (Van Solingen *et al.* 2009). In current study MiR-126 was one of the top five expressed transcripts in Guilan native cow population (Table 3).

Table 1 Summary of RNA-seq alignment to the reference genome (bosTau4) for Guilan native cow population

Left reads	Guilan native cattle
Total sequenced Fragments	28434708
Mapped to reference genome	25421000 (89.4% of input)
Sequences with multiple alignments	2170039 (8.5% of input)
Right reads	
Total sequenced Fragments	28434708
Mapped to reference genome	24278459 (85.4% of input)
Sequences with multiple alignments	2064643 (8.5% of input)
Overall read mapping rate	87.40%
Aligned pairs	23140289
Aligned pairs with multiple alignments	1956181 (8.5%)
Aligned pairs with discordant alignments	744972 (3.2%)
Aligned pairs with concordant alignments	18234548 (78.8%)
Number of insertions	140262
Number of deletions	144258
Number of splice junctions	192100

Table 2 Summary of mapped reads of Guilan native cow population

Items	Guilan native cattle
Total number of expressed genes	24616
Number of coding genes	19994
Number of non-coding genes	3825
Number of pseudogenes	797

Table 3 Top five highly expressed transcripts in Guilan native cow population

	Transcript ID	Gene symbol	Value (FPKM*)
1	ENSBTAT00000060539	<i>ATP6</i>	36230.2
2	ENSBTAT00000024376	<i>RPLP1</i>	8343.5
3	ENSBTAT00000016359	<i>B2M</i>	7248.9
4	ENSBTAT00000060548	<i>ND2</i>	6235.6
5	ENSBTAT00000036924	<i>bta-miR-126</i>	4944.1

* FPKM: fragments per kilobase of exon per million fragments mapped.

Out of first five highly expressed genes, two (*ATP6* and *ND2*) genes belong to electron transport chain in mitochondria and could be the indicative of the importance of this cyclic energy creation biological pathway in cows. The most of the highly expressed genes involved in ribosomal structures and translational activities that could be explained by the housekeeping functions of these genes.

RNA-seq has many advantages over traditional cDNA microarray technologies. RNA-seq is free from probe design or bias from hybridization issues and is more sensitive in detecting genes with very low expression and more accurate in detecting expression of extremely abundant genes (Croucher and Thomson, 2010; Marioni *et al.* 2011).

With manufacturers predicting increased read lengths, reduced costs and faster sequencing relative to existing platforms, the future of RNA-Seq technology appears to be promising and routinely affordable. It is expected that once the issues with the widespread use of RNA-Seq are overcome (e.g. higher cost, high data-storage requirements, and the absence of a standard for analysis) this technique will become the predominant tool for transcriptome analysis.

Marioni *et al.* (2011) showed that the correlation between RNA-Seq and qRT-PCR could reach 0.93, offering that the RNA-Seq technique is accurate and reproducible. This technique may also consider for discovering variations reasons in economically important traits and an accurate tool for designing of a cattle breeding program.

The negative consequences of genetic erosion and inbreeding depression may be manifested by loss of viability, fertility and disease resistance, and the frequent occurrence of recessive genetic diseases (FAO, 2011). Thus, more attention should be paid to animal genetic resources. More accurate genetic information can be obtained to better understanding of existing animal genetic resources by applying for advances in molecular biotechnology (Yang *et al.* 2013).

CONCLUSION

This study investigated the complexity of the blood transcriptome using RNA-seq technique and it was the first report of blood leukocytes transcriptome sequencing of

Guilan native cow population. Guilan native cow has the biggest population among all native cattle populations in Iran and there is a lot of interest to us to explore the genetic potential of this population. Such studies could help to detect and introduce highly expressed genes and unique transcripts in Guilan native cow population and flaunt the genetic potential of this population with other indicine breeds with high resistance to harsh environments, diseases, and feed shortages. Furthermore, the results of this study can be used as a valuable resource for further investigations on cattle aimed at finding genes expression patterns and RNA sequence variations.

ACKNOWLEDGEMENT

This work was supported by Iran national science foundation under grant No. 93012544.

REFERENCES

- Croucher N.J. and Thomson N.R. (2010). Studying bacterial transcriptomes using RNA-seq. *Curr. Opin. Microbiol.* **13**, 619-624.
- Ebrahimi Z., Mohammadabadi M.R., Esmailzadeh A.K., Khezri A. and Najmi Noori A. (2015a). Association of *PIT1* gene with milk fat percentage in Holstein cattle. *Iranian J. Appl. Anim. Sci.* **5**, 575-582.
- Ebrahimi Z., Mohammadabadi M.R., Esmailzadeh A. and Khezri A. (2015b). Association of *PIT1* gene and milk protein percentage in Holstein cattle. *J. Livest. Sci. Technol.* **3**, 41-49.
- Eisenberg E. and Levanon E.Y. (2013). Human housekeeping genes, revisited. *Trans Genet.* **29**, 569-574.
- FAO. (2011). Molecular genetic characterization of animal genetic resources. Food and Agriculture Organization of the United Nations (FAO), Rome, Italy.
- Habersetzer J., Ziani W., Larriou I., Stines-Chaumeil C., Giraud M.F., Brêthes D., Dautant A. and Paumard P. (2013). ATP synthase oligomerization: from the enzyme models to the mitochondrial morphology. *Int. J. Biochem. Cell. Biol.* **45**, 99-105.
- Hall Z., Schmidt C. and Politis A. (2015). Uncovering the early assembly mechanism for amyloidogenic β 2-microglobulin using cross-linking and native mass spectrometry. *J. Biol. Chem.* **291**, 4626-4637.
- Huang W., Nadeem A., Zhang B., Babar M., Soller M. and Khatib H. (2012). Characterization and comparison of the leukocyte transcriptomes of three cattle breeds. *PLoS One.* **7**, 30244-30251.
- Iran's Country Report on Farm Animal Genetic Resources. (2004). Draft Iran's Country Report on Farm Animal Genetic Resources. Animal Science Research Institute of Iran, Tehran, Iran.
- Jiang Z., Uboh C.E., Chen J. and Soma L.R. (2013). Isolation of RNA from equine peripheral blood cells: comparison of methods. *SpringerPlus.* **2**, 1-6.
- Kamalzadeh A., Rajabbaigy M. and Kiasat A. (2008). Livestock production systems and trends in livestock industry in Iran. *J. Agric. Soc. Sci.* **4**, 183-188.
- Kharrati Koopaei H., Mohammad Abadi M.R., Ansari Mahyari S., Tarang A.R., Potki P. and Esmailzadeh A.K. (2012). Effect of *DGAT1* variants on milk composition traits in Iranian Holstein cattle population. *Anim. Sci. Pap. Rep.* **30**, 231-240.
- Kim D., Perteau G., Trapnell C., Harold P., Kelley R. and Salzberg S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome. Biol.* **14**, 36-42.
- Marioni J.C., Mason C.E., Mane S.M., Stephens M. and Gilad Y. (2011). RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome. Res.* **18**, 1509-1517.
- Martinez-Azorin F., Remacha M. and Ballesta J.P. (2008). Functional characterization of ribosomal P1/P2 proteins in human cells. *Biochem. J.* **413**, 527-534.
- Mohammadabadi M.R., Torabi A., Tahmourespour M., Baghizadeh A., Esmailzadeh Koshkoie A. and Mohammadi A. (2010). Analysis of bovine growth hormone gene polymorphism of local and Holstein cattle breeds in Kerman province of Iran using polymerase chain reaction restriction fragment length polymorphism (PCR-RFLP). *African J. Biotechnol.* **9**, 6848-6852.
- Nimbkar C., Gibson J., Okeyo M., Boettcher P. and Soelkner J. (2008). Sustainable use and genetic improvement. *Anim. Genet. Resour.* **42**, 49-70.
- Pasandideh M., Mohammadabadi M.R., Esmailzadeh A.K. and Tarang A. (2015). Association of bovine *PPARGCIA* and *OPN* genes with milk production and composition in Holstein cattle. *Czech J. Anim. Sci.* **60**, 97-104.
- Rajkumar A.P., Qvist P., Lazarus R., Lescai F., Ju J., Nyegaard M., Mors O., Børghlum A.D., Li Q. and Christensen J.H. (2015). Experimental validation of methods for differential gene expression analysis and sample pooling in RNA-seq. *BMC Genom.* **16**, 548-552.
- Rewe T.O., Peixoto M.G.C.D., Cardoso V.L., Vercesi Filho A.E., El Faro L. and Strandberg E. (2015). Gir for the Giriama: the case for Zebu dairying in the tropics-a review. *Livest. Res. Rural Dev.* Available at: <http://www.lrrd.org/lrrd27/8/rewe27150.html>.
- Suarez Y., Fernandez-Hernando C., Yu J., Gerber S.A., Harrison K.D., Pober J.S., Iruela-Arispe M.L., Merkenschlager M. and Sessa W.C. (2008). Dicer-dependent endothelial microRNAs are necessary for postnatal angiogenesis. *Proc. Natl. Acad. Sci. USA.* **105**, 14082-14087.
- The Bovine HapMap Consortium. (2009). Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science.* **324**, 528-532.
- Trapnell C., Williams B.A., Perteau G., Mortazavi A., Kwan G., Van Baren M.J., Salzberg S.L., Wold B.J. and Pachter L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511-515.
- Van Solingen C., Seghers L., Bijkerk R., Duijs J.M., Roeten M.K., van Oeveren-Rietdijk A.M., Baelde H.J., Monge M., Vos J.B.,

- de Boer H.C., Quax P.H., Rabelink T.J. and van Zonneveld A.J. (2009). Antagomir-mediated silencing of endothelial cell specific microRNA-126 impairs ischemia-induced angiogenesis. *J. Cell. Mol. Med.* **13**, 1577-1585.
- Wang Z., Gerstein M. and Snyder M. (2009). RNA-seq: a revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57-63.
- Wilson R.T. (2009). Fit for purpose-the right animal in the right place. *Trop. Anim. Health Prod.* **41**, 1081-1090.
- Yan S., Yim L.Y., Lu L., Lau C.S. and Chan V.S.F. (2014). MicroRNA regulation in systemic lupus erythematosus pathogenesis. *Immune. Netw.* **14**, 138-148.
- Yang W., Kang X., Yang O., Lin Y. and Fang M. (2013). Review on the development of genotyping methods for assessing farm animal diversity. *J. Anim. Sci. Biotechnol.* **4**, 2-9.
- Zhang L., Jia S., Plath M., Huang Y., Li C., Lei C., Zhao X. and Chen H. (2015). Impact of parental *Bos taurus* and *Bos indicus* origins on copy number variation in traditional Chinese cattle breeds. *Genome. Biol. Evol.* **7**, 2352-2361.
- Zickermann V., Wirth C., Nasiri H., Siegmund K., Schwalbe H., Hunte C. and Brandt U. (2015). Mechanistic insight from the crystal structure of mitochondrial complex I. *Science.* **347**, 44-49.
-