This is an Accepted Manuscript, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this Accepted Manuscript with the edited and formatted Advance Article as soon as it is available.

You can find more information about Accepted Manuscripts in the Information for Authors.

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal’s standard Terms & Conditions and the Ethical guidelines still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this Accepted Manuscript or any consequences arising from the use of any information it contains.
Tetra-primer ARMS–PCR is an efficient SNP genotyping method: with an example from SIRT2

Mingxun Li\textsuperscript{1}, Xiaomei Sun\textsuperscript{1}, Jing Jiang\textsuperscript{1}, Yujia Sun\textsuperscript{1}, Xianyong Lan\textsuperscript{1}, Chuzhao Lei\textsuperscript{1}, Chunlei Zhang\textsuperscript{2}, Hong Chen\textsuperscript{1}\textasternote

\textsuperscript{1}.College of Animal Science and Technology, Northwest A\&F University, Shaanxi Key Laboratory of Molecular Biology for Agriculture, Yangling, Shaanxi 712100, China.

\textsuperscript{2}.Institute of Cellular and Molecular Biology, Jiangsu Normal University, Xuzhou, Jiangsu 221116, China.

*Corresponding author: Hong Chen

Tel: +86-029-8709 2102

Fax: +86-029-8709 2164

E-mail: chenhong1212@263.net

Address: No.22 Xinong Road, College of Animal Science and Technology, Northwest A\&F University, Yangling, Shaanxi 712100, China.
Abstract: Tetra-primer amplification refractory mutation system PCR (T-ARMS-PCR) offers fast detection and extreme simplicity at a negligible cost for SNP genotyping. SIRT2, the family member (sirtuins, SIRT1–7) with the greatest homology to the silent information regulator 2 (Sir2), is the most abundantly expressed sirtuins in adipocytes and has been implicated in promoting fatty acid oxidation (FAO) by deacetylating various substrates. In the current study, we have successfully genotyped a new identified bovine SIRT2 SNP g.4140A>G by T-ARMS-PCR method and validated the accuracy by PCR-RFLP assay using 1255 animals representing the five main Chinese breeds. The concordance between the two different methods was 98.8%. Individuals with discordant genotypes were retyped by direct DNA sequencing. 40% of the discrepancies could be attributed to incomplete digestion in the PCR–RFLP assay. 60% of discordant genotypes were resulted from allele failure in the T-ARMS-PCR assay. Chi-square test showed that the frequencies of SNP g.4140A>G are in Hardy–Weinberg equilibrium in the all samples (P>0.05), which suggested that the five populations are almost a dynamic equilibrium even in artificial selection. Association analysis showed that the g.4140A>G polymorphism is significantly related to 24-months-old body weight in Nanyang cattle. Our results provide direct evidence that T-ARMS-PCR is a rapid, reliable, and cost-effective method for SNP genotyping and g.4140A>G polymorphism in bovine SIRT2 is associated with growth efficiency traits. These findings may be used for marker-assisted selection and management in feedlot cattle.

Key Words: T-ARMS-PCR, accuracy, SIRT2, SNP, growth traits
1 Introduction

Cattle have been serving as a valuable model for studying mammalian metabolism, reproduction, physiology and comparative genomics, as well as providing critical sources of human dietary protein and economic security. The publication of a complete draft genome sequence, based on the DNA of a partially inbred individual from the Hereford breed and her sire, and follow-on efforts to characterize genetic variation have resulted in the discovery of more than ten millions of single nucleotide polymorphisms (SNPs). Differences in just one of these SNPs may affect the function of a gene and mean the difference between a highly productive and a poorly performing animal. Therefore, efficient and cost-effective SNP genotyping methods will be required for routine cattle selective applications once disease- or growth-predisposing genes have been identified and the allelic variants that predict disease or improve diagnostics have been specified.

To date, there are numerous methods for SNP genotyping. Among them, tetra-primer amplification refractory mutation system PCR (T-ARMS-PCR) is an inexpensive, time-saving genotyping method. The technique employs two primer pairs to amplify two alleles in one PCR reaction. The region flanking the mutation is amplified by two outer primers, producing a non–allele-specific control amplicons. Two allele-specific (inner) primers are designed in opposite orientation and, in combination with the outer primers, can simultaneously amplify both the wild-type and the mutant amplicons. The two allele-specific amplicons have significantly different lengths allowing for easily separated by agarose gel electrophoresis.

Silent information regulator 2 (SIR2) has been implicated in the regulation of lifespan under calorie restriction in lower organisms, including yeast, *Caenorhabditis elegans*, and *Drosophila melanogaster*. In mammals, the homologs of SIR2 have been named sirtuins (SIRT), with seven members in a family termed SIRT1 through SIRT7. They share a conserved central deacetylase domain but have different N- and C termini and display distinct subcellular localization, suggesting different biological functions. Sirtuin family of proteins possess NAD⁺-dependent deacetylase and ADP ribosyltransferase activities. They play an important role in aging, inflammation, and metabolism regulation and have been thought to “sense” the beneficial effects of caloric restriction on physiology.

SIRT2 proteins, a most abundant sirtuin in adipocytes, are distributed throughout the cytoplasm, mainly colocalizing with microtubules and functioning as an α-tubulin deacetylase. SIRT2 expression level is elevated in the white adipose tissue and kidney of caloric restricted mice and downregulated during preadipocyte differentiation in 3T3-L1 cells. Overexpression of SIRT2 inhibits differentiation and promotes lipolysis in mature adipocytes, whereas reducing SIRT2 expression promotes adipogenesis. Mechanistically, SIRT2 suppresses adipogenesis by deacetylating FOXO1 to promote FOXO1’s binding to PPARγ and subsequent repression on PPARγ transcriptional activity.

Because of the important role of SIRT2 in metabolism and studies in humans have concluded that polymorphisms in SIRT2 affect height among elderly Japanese
subjects, here, we developed a sensitive, quick and low-cost T-ARMS-PCR assay to reliably detect bovine SIRT2 g.4140A>G polymorphism, evaluated the accuracy of this method, and further investigated the genetic effects of this polymorphism on economic growth traits of Nanyang cattle.

2 Material and methods

All animal experiments were performed in compliance with the relevant laws and institutional guidelines and were approved by Northwest A&F University Institutional Animal Care and Use Committee.

2.1 Animals and genomic DNA isolation

Blood samples were collected from 1255 Chinese cattle representing five breeds: Nanyang (NY, n=210), Qinchan (QC, n=224), Luxi (LX, n=168), Jiaxian (JX, n=416), and Chinese Red Steppe cattle (CRS, n=237). These five groups represent the main breeds of China and are reared in the provinces of Henan, Shaanxi, Shandong, Henan, and Jilin, respectively. Among them, the NY, JX, QC and LX were used for beef production, while CRS is a dual purpose (beef and dairy) breed. All of the NY cattle in the present study were reared at one farm under the same conditions and weaned at 6 months of age and then fed a concentrate and straw diet ad libitum until 24 months of age. Their growth traits (body weight, body height, body length, chest girth, hucklebone width and average daily gain) were recorded at birth, 6, 12, 18, and 24 months of age.

Genomic DNA was isolated from 2% heparin-treated blood samples as previously described. DNA quantity and purity (A260/A280 ratio) for each sample was assessed using the NanoDrop™ 1000 Spectrometer (Thermo Scientific, Waltam, MA, USA).

2.2 SNP discovery

In an effort to identify SNPs in a cost-effective manner, SNP discovery was implemented by sequencing pooled PCR products, which were amplified from DNA of 50 individuals that were randomly chosen from each cattle breed. Primers used to screen genetic polymorphisms were designed based on the GenBank sequence NC_007316.4. PCR amplicons were sequenced (3730 DNA Analyzer, Applied Biosystem) in forward and reverse directions for SNP discovery. Gene sequences and polymorphisms were assembled and annotated using Vector NTI advance 11.5.1 software (Invitrogen Corporation). The polymorphism identified in our SNP discovery analysis was compared with the bovine NCBI dbSNP database (http://www.ncbi.nlm.nih.gov/SNP/index.html) using BLAST (Basic Local Alignment Search Tool).

2.3 T-ARMS-PCR method

The SIRT2 gene DNA region containing a new identified g.4140A>G polymorphism (NCBI ss#770757955) was used to design two sets of PCR primers (Fig. 1). The primers were designed utilizing the web-based program accessible from...
http://www.cedar.genetics.soton.ac.uk/public_html/primer1.html, following the rules to introduce a second deliberate mismatch at position -2 from the 3' terminus. Details of primer sequences and PCR conditions were given in Table 1. The PCR reactions were performed in a total volume of 10 µl, containing 10 pmol of each of the inner primers, 2 pmol of each of the outer primers, 200 µM of each dNTP, 2 mM of MgCl₂, 1× PCR buffer, 50 ng of DNA and 0.2 U of Taq DNA polymerase (MBI, Fermentas). To increase the specificity of the reaction, a touchdown profile was performed with 94 °C for 4 min; 18 cycles of 30 s at 94 °C, 30 s at 68°C decreasing by 1°C per cycle, and 40 s at 72 °C; 20 cycles of 30 s at 94 °C, 30 s at 50°C, and 40 s at 72 °C; and a final extension of 10 min at 72 °C. The PCR products were electrophoresed on a 3% agarose gel and stained with ethidium bromide (Fig. 2).

2.4 Validation of genotypes scored by T-ARMS-PCR

To validate the accuracy of genotype scoring by T-ARMS-PCR, conventional PCR was performed in all samples, followed by restriction endonuclease digestion. PCR was carried out under the condition used in our previous study. A 5 µl aliquot of the PCR product was mixed with 1× buffer and 3 units of MspI (TaKaRa Biotech Co, Dalian, China) at 37°C for 12 h and then separated on 3% agarose gels (Fig. 2).

2.5 Statistical Analyses

Specific genotype frequencies, allele frequencies and goodness of fit tests for Hardy–Weinberg equilibrium were calculated using the HelixTree software version 6.3.1(Golden Helix Inc). Gene heterozygosity (He), as a measure of the degree of genetic variability, was determined according to Nei's methods. The effective number of alleles per locus (Ne) was calculated according to Morgante’s formula. The polymorphism information content (PIC) value is commonly used in genetics as a measure of polymorphism for a marker locus used in linkage analysis and was calculated according to Botstein's methods. The formulas were as follows:

\[ H_e = 1 - \sum_{i=1}^{n} P_i^2 \]
\[ Ne = \left( \frac{\sum_{i=1}^{n} P_i^2}{n} \right) \]
\[ PIC = 1 - \sum_{i=1}^{m} P_i^2 - \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} 2P_i P_j \]

where \( P_i \) is the frequency of the \( i^{th} \) allele and \( n \) is the number of alleles.

The effects of genotype on the growth traits of NY cattle were analyzed by the least-square method as applied in the general linear model (GLM) procedure of SPSS 21 (IBM, Armonk, NY, USA) according to our previously reported statistical model:

\[ Y_{ij} = \mu + A_i + G_j + e_{ij} \]

where \( Y_{ij} \) was the trait measured on each of the \( ij^{th} \) animal, \( \mu \) was the overall mean for each trait, \( A_i \) was fixed effect due to the \( i^{th} \) age, \( G_j \) was the fixed effect of \( j^{th} \) single SNP marker genotype, and \( e_{ij} \) was the random error. This linear model did not include the season of birth effect since preliminary analysis indicated that it did not have a significant effect on the measured traits.

3 Results
3.1 Results of T-ARMS-PCR

T-ARMS-PCR method was successfully applied to genotype SIRT2 g.4140A>G polymorphism. The amount of inner to outer primers at a ratio of 5:1 was the most favorable to ensure amplification efficiency and allele specificity in our analysis. In addition, the use of touchdown PCR also permitted enhancement of the amplification of the allele-specific fragments and reduced artificial products.

To validate the accuracy of genotype scoring by T-ARMS-PCR, conventional genotyping assay (PCR-RFLP) was performed in all samples, examples of which were shown in Fig. 2. The concordance between the two different methods was 98.8%, based on a total of 1,255 genotypes. Individuals with discordant genotypes were retyped by direct DNA sequencing. 40% of the discrepancies could be attributed to incomplete digestion in the PCR–RFLP assay (Table 1). 60% of discordant genotypes were resulted from allele failure in the T-ARMS-PCR assay (Table 1).

3.2 Genetic diversity analyses

Genotype and allele frequencies for SNP g.4140A>G were presented in Table 2. Frequencies for alleles were not balanced, the A allele being predominant over the G allele in all tested breeds, especially in CRS population. The minor allele frequencies were 0.381, 0.353, 0.416, 0.457 and 0.017 for NY, QC, JX, LX, and CRS, respectively. Chi-square test showed that the frequencies of SNP g.4140A>G are in Hardy–Weinberg equilibrium in the all samples ($P>0.05$), which suggested that the five populations are almost a dynamic equilibrium even in artificial selection.

In order to measure the informativeness of the identified SNP, the polymorphic information content (PIC value) was calculated. PIC values ranged from 0.032 to 0.373. The LX population revealed the highest PIC value (0.373), which corresponds to the highest He (0.496), while the CRS had a lowest genetic diversity (PIC=0.032). The NY, QC, JX and JX are Chinese local breeds with a long history while the CRS is a crossbred breed developed in China since the 1920's. Not surprisingly, the former had richer genetic polymorphisms compared to the latter.

3.3 Associations between SNP g.4140A>G and growth traits of Nanyang cattle

Relationship between SNP g.4140A>G and growth traits (body weight, body height, body length, and chest girth) were analyzed in Nanyang cattle aged 0, 6, 12, 18, and 24 months. As shown in Table 3, a significant association was found between this polymorphism and 24-months-old body weight. Compared with individuals who carried the AA genotype (364.49kg), G carriers (AG, 377.99kg; GG, 385.69kg) had a significantly higher body weight ($P=0.045$, $P=0.008$, respectively). However, there were no significant associations between the polymorphism and other economic growth traits (data not shown).

4 Discussion

The sirtuins represent a complex family of proteins that show homology to the yeast class III NAD-dependent protein/histone deacetylase SIR2. SIRT2 is the most abundantly expressed sirtuins in adipocytes and has been implicated in promoting
fatty acid oxidation (FAO) by deacetylating various substrates.\textsuperscript{11,12} In 3T3-L1 adipocytes, SIRT2 deacetylates FOXO1 and induces FOXO1 binding to the promoter of PPAR\textgamma and subsequently represses PPAR\textgamma activity.\textsuperscript{17} Reduced interaction between SIRT2 and FOXO1 enhances insulin-stimulated phosphorylation of FOXO1 by AKT/PKB, which in turn increases FOXO1 cytosolic localization.\textsuperscript{12} In addition, SIRT2 can deacetylate PGC1-\textalpha in vivo and in vitro, which leads to an increase in fatty acid oxidation and gluconeogenesis and a decrease in glycolysis.\textsuperscript{26} Studies in humans concluded that polymorphisms in SIRT2 affected height among elderly Japanese subjects and were associated with risk of Alzheimer’s disease.\textsuperscript{18,27}

Since its development, T-ARMS–PCR method has become one of the most commonly used methods for SNP genotyping.\textsuperscript{2,3} It is a significant advantage over the commonly used PCR–RFLP method in which partial digests often lead to inaccurate genotype inference. In addition, because T-ARMS–PCR does not need incubation time with a restriction enzyme, it offers fast detection and extreme simplicity at a negligible cost.

In this study, we successfully genotyped SIRT2 g.4140A>G polymorphism by T-ARMS-PCR. The concordance between T-ARMS-PCR method and PCR–RFLP was 98.8%, demonstrated its reliability in SNP genotyping. Allele frequencies, genotype frequencies, and PIC value at g.4140A>G locus were significantly different between CRS and four other beef breeds, implying that this mutation is possibly associated with some quantitative traits. So we next investigated the effect of g.4140A>G on economic growth traits of Nanyang cattle. Our results demonstrated that g.4140A>G mutation was significantly associated with 24-months-old body weight. This SNP located in non-coding regions of bovine SIRT2, which did not result in changes in amino acids. But previously reports still have powerfully proved that intronic polymorphisms could also affect the gene’s function and phenotypic expression. In human, the SNP rs7202116 located in the first intron of the fat mass and obesity associated gene (FTO) gene was significantly associated with phenotypic variability of body mass index.\textsuperscript{28} In pigs, a single nucleotide substitution in intron 3 of insulin-like growth factor 2 (IGF2) abrogates a binding site for a repressor and leads to a 3-fold up-regulation of IGF2 in skeletal muscle. The mutation has major effects on muscle growth, size of the heart, and fat deposition.\textsuperscript{29} Such associations may also be the results of linkage between these SNPs and other genes on the same chromosome that have a significant effect on these production traits.

Our previous study has revealed that polymorphisms in bovine SIRT1 are associated with growth traits in Nanyang cattle.\textsuperscript{25} In the current study, we further demonstrated that genetic variation in bovine SIRT2 is also associated with growth efficiency traits in Nanyang cattle. SIRT1 and SIRT2 share a conserved central deacetylase domain and both of the two proteins inhibit differentiation and promote lipolysis in mature adipocytes. The similar genetic effects can be attributed to their similar function in adipogenesis. These findings implies that the sirtuin family may play an important role in animal growth and development and is useful in marker-assisted selection and management in feedlot cattle.

In conclusion, we provide a rapid, reliable, and cost-effective detection for SIRT2
g.4140A>G mutation without the use of any special equipment, thus improving the accessibility to SNP genotyping for all minimally equipped laboratories. The SNP g.4140A>G was demonstrated to have a significant association with 24-months-old body weight of NY cattle. These findings will benefit for cattle breeding and genetics.

Acknowledgments

This study was supported by the National 863 Program of China (Grant No. 2013AA102505), National Natural Science Foundation of China (Grant No. 31272408), Program of National Beef Cattle and Yak Industrial Technology System (CARS-38), Agricultural Science and Technology Innovation Projects of Shaanxi Province (No.2012NKC01-13).

References


Fig. 1. (A) Schematic presentation of the SIRT2 gene (NM_001113531.1). Exons and introns are represented by cylinders and lines, respectively. (B) Tetra primer positions relative to the A>G substitution in intron 3 are shown by arrows. P3 and P4 are outer primers acting as control primer, and P1 and P2 are allele-specific primers. (C) Schematic genotype pattern of T-ARMS-PCR for SNP g.4140A>G.
Fig. 2. (A) SNP g.4140A>G genotyping by T-ARMS-PCR resolved on a 3% agarose gel. 419 bp band represents the common amplicon, whereas the A and G allele-specific bands are represented by the 210 and 262 bp amplicons, respectively. (B) Genotyping pattern for g.4140A>G by PCR–RFLP digesting with MspI. AA, 400 bp; AG, 400+335+65 bp; GG, 335+65 bp. Genomic DNA used in PCR-RFLP method corresponds to the samples used in the T-ARMS-PCR in (A). M represented Marker I.
### Table 1. PCR primers and conditions for identification of SIRT2 g.4140A>G polymorphism.

<table>
<thead>
<tr>
<th>Genotyping methods</th>
<th>Primers (5′–3′)(^a)</th>
<th>Mistyped number</th>
<th>Restriction enzyme</th>
<th>Genotype pattern (bp)</th>
</tr>
</thead>
</table>
| T-ARMS-P CR        | P1: GGAGGCGGGTTAAGGCAGGGGATACG  
P2: TGTCACCCCTGGTGAGGTGAACACT  
P3: GGTGTCATCCCCCTCTCCCCCTCCCTAAC  
P4: CCCAGAGTCTGGAGAAGAAATTTCCGCA | 9                | -                  | 419bp (outer) |
| PCR-RFLP           | F: CGAAGTCACCCCATAGGAGGC  
R: ATCCCGAATGCTGATCTG | 6                | MspI               | 400/335+65\(^b\) |

\(^a\) Specificity is increased by the introduction of a deliberate mismatch at position -2 of the polymorphism site, indicated by underlined letters.

\(^b\) The pattern size was too short to be visible in the gel.
<table>
<thead>
<tr>
<th>Breeds</th>
<th>Genotype frequencies</th>
<th>Allele frequencies</th>
<th>$\chi^2$ (HWE)$^a$</th>
<th>Diversity parameters$^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AA</td>
<td>AG</td>
<td>GG</td>
<td>A</td>
</tr>
<tr>
<td>NY</td>
<td>0.373</td>
<td>0.492</td>
<td>0.135</td>
<td>0.619</td>
</tr>
<tr>
<td>QC</td>
<td>0.44</td>
<td>0.415</td>
<td>0.145</td>
<td>0.647</td>
</tr>
<tr>
<td>JX</td>
<td>0.328</td>
<td>0.511</td>
<td>0.16</td>
<td>0.584</td>
</tr>
<tr>
<td>LX</td>
<td>0.319</td>
<td>0.448</td>
<td>0.233</td>
<td>0.543</td>
</tr>
<tr>
<td>CRS</td>
<td>0.967</td>
<td>0.033</td>
<td>0</td>
<td>0.983</td>
</tr>
</tbody>
</table>

Note: $\chi^2_{0.05}$ (df=2) = 5.99, $\chi^2_{0.05}$ (df=1) = 3.84.

$^a$ $\chi^2$ (HWE), Hardy–Weinberg equilibrium $\chi^2$ value.

$^b$ He, gene heterozygosity; Ne, effective allele numbers; PIC, polymorphism information content.
Table 3. Effects of the SIRT2 g.4140A>G SNP on Nanyang body weight

<table>
<thead>
<tr>
<th>Age (months)</th>
<th>Body weight (Means± SE)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AA</td>
<td>AG</td>
</tr>
<tr>
<td>0</td>
<td>30.21±0.34</td>
<td>29.65±0.37</td>
</tr>
<tr>
<td>6</td>
<td>160.84±2.94</td>
<td>157.89±3.17</td>
</tr>
<tr>
<td>12</td>
<td>208.89±7.48</td>
<td>222.67±3.14</td>
</tr>
<tr>
<td>18</td>
<td>293.08±4.94</td>
<td>301.58±4.58</td>
</tr>
<tr>
<td>24</td>
<td>364.49±2.75</td>
<td>377.99±7.78</td>
</tr>
</tbody>
</table>

Values with different superscript letters within the same row differ significantly at $P < 0.01$ (A, B) and $P < 0.05$ (a, b).