Gene 551 (2014) 183–188

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Identification of a new non-coding exon and haplotype variability in the cattle *DEFB103* gene

Ali Mirabzadeh-Ardakani [a,*], Philip Griebel [b], Sheila M. Schmutz [a]

[a] Department of Animal and Poultry Science, University of Saskatchewan, Saskatoon, Canada
[b] Vaccine and Infectious Disease Organization, University of Saskatchewan, Saskatoon, Canada

ARTICLE INFO

Article history:
Received 30 June 2014
Received in revised form 15 August 2014
Accepted 26 August 2014
Available online 27 August 2014

Keywords:
Copy number variation
Defensin
Haplotype
Splicing

ABSTRACT

The *DEFB103* gene is a member of the β-defensin gene family. In this study, we applied multiple sets of primers to characterize the *DEFB103* transcript. RT-PCR was used to determine the cDNA boundaries and it indicated that the cDNA start point is at least 514 bp before the start codon and not further than 678 bp. In addition, the length of the 3′UTR was determined to be at least 53 bp after the stop codon. Seven SNPs were located in the 5′UTR, and comprised 4 different haplotypes in genomic DNA. Using these haplotype data, it could be proven that at least two complete copies of *DEFB103* with an ATG start codon are present in cDNA in most cattle. Additionally haplotype data indicated that there are also multiple incomplete copies in most cattle. A non-coding exon 1a, and a 261 bp intron 1a were identified in cattle, and subsequently predicted in sheep and goats. *DEFB103* sequence assemblies and partial cloning sequences revealed two types of deletion (4 bp and 8 bp) in the 5′UTR. These observations could prove that these copies are not assembly artifact.

1. Introduction

Defensins are members of a cationic protein family (3–4 kDa) with six conserved cysteine residues that create three cysteine–disulfide bonds (Ganz and Lehrer, 1994). These conserved cysteines provide three different patterns for the disulfide bonds, that divide the defensin family into three subfamilies called α, β (Lehrer and Ganz, 1999), and θ-defensins (Tang et al., 1999). β-Defensins are in a single gene cluster in cattle and they were mapped to chromosome 27 (Gallagher et al., 1995). β-Defensin 103 (*DEFB103*) is a multi-functional gene influencing innate immunity (Bevins, 2006), coat color (Candille et al., 2007), and feeding behavior (Candille et al., 2007). All members of the β-defensin gene family in humans have two exons and one intron, with the exception of *DEFB105* (GenBank NM_152250.2) which has three coding exons and two introns (Pazgier et al., 2006).

Data from mRNA-seq showed that alternative splicing occurs in more than 90% of human genes (Pan et al., 2008; Wang et al., 2008). Genes with alternative promoter regions tend to generate alternative splicing (Xin et al., 2008). *Microphthalmia-associated transcription factor* (*MITF*) with nine exons (Tassabehji et al., 1994) and *insulin-like growth factor 2* (*IGF2*) with 10 exons (Goodall and Schmutz, 2007) are such genes. In most genes with alternative splicing, some of the exons (Goodall and Schmutz, 2007) or part of the exons (Udono et al., 2000) are non-coding.

β-Defensins are variable in their copy number (Hollox et al., 2003). *DEFB103* copy number variation has been reported in humans (2–12 copies per diploid genome) (Hollox et al., 2003) and dogs (2–4 copies per diploid genome) (Leonard et al., 2012). Cattle *DEFB103* occurs in multiple copies in at least some individuals, based on sequence results from a cloning experiment (GenBank KM347983, KM347984, KM347985, and KM347986). Five SNPs were reported in the 5′UTR of the *DEFB103* by Dreger and Schmutz (2009) (c.−319A>G, c.−264C>T, c.−69A>G, c.−42A>G, and c.−34G>A), and two additional SNPs (c.−383A>G, c.−241G>A) were identified in the 5′UTR of the *DEFB103* (Mirabzadeh-Ardakani et al., in press). The current research was conducted to investigate gene structure and haplotype variation of *DEFB103* in different tissues collected from dairy and beef cattle.

2. Materials and methods

2.1. Sample collection

Animals used for this study were humanely killed according to Canadian Council of Animal Care guidelines and protocols approved by the University of Saskatchewan Animal Care Committee. Tissue samples were collected from seven neonatal Holstein calves and five 6–8 month old Angus-cross beef calves. The tongue, buccal epithelium, conjunctiva, rumen, skin, trachea, and liver from neonatal calves and

Abbreviations: DEFB103, β-defensin 103; mRNA-seq, messenger ribonucleic acid sequencing; MITF, microphthalmia-associated transcription factor; IGF2, insulin-like growth factor 2; cDNA, complementary deoxyribonucleic acid; PCR, polymerase chain reaction; RT-PCR, reverse transcription polymerase chain reaction; UTR, untranslated region; SNP, single nucleotide polymorphism; CNV, copy number variation.

* Corresponding author at: Department of Animal and Poultry Science, University of Saskatchewan, 51 Campus Dr., Saskatoon, SK S7N 5A8, Canada.
E-mail address: a.mirabzadeh@usask.ca (A. Mirabzadeh-Ardakani).

http://dx.doi.org/10.1016/j.gene.2014.08.055
0378-1119/

tongue, buccal epithelium, trachea, and hard-palate from Angus-cross calves were collected immediately after euthanasia and stored in RNA*later*® (Applied Biosystem-Ambion) in −80 °C. Whole blood samples (5 ml) were collected from the jugular vein for each animal in a BD Vacutainers® with EDTA as an anticoagulant.

2.2. DNA and total RNA extraction

DNA was extracted from the buffy coat using a SQ blood DNA kit (OMEGA Bio-Tek). Total RNA was extracted from collected tissues stabilized by RNA*later*® using the TRIzol method (Invitrogen). After the first step of extraction, the extracted RNA was treated with DNase-I (RNase free) (Invitrogen) to exclude any DNA contamination. The RNA quantity and quality were determined using a NanoDrop 2000 (NanoDrop, Technologies, DE, USA) spectrophotometer after the first extraction, and using an Agilent 2100 Bioanalyzer (G2938B, Agilent Technologies, Mississauga, Ontario) after the second RNA extraction. RNA with an integrity number higher than 7 was considered acceptable as intact RNA for cDNA preparation.

2.3. Reverse transcription and cDNA synthesis

Total RNA (250 ng) was reverse-transcribed using qScript™ cDNA SuperMix (Quanta Biosciences™), following the manufacturer's protocol. The cDNA was diluted to 1:10 as a template for the reverse transcription-polymerase chain reaction (RT-PCR).

2.4. Semi nested-PCR and RT-PCR amplification and genotyping

A step wise walking procedure was used to design a 5′ forward primer approximately every 100 bp. The furthest one that could amplify a fragment was selected for the experiment. Primers (Table 1) were designed using Primer3web (http://bioinfo.ut.ee/primer3/) with the exception of KNEST1-F and KNEST1-R primers (Dreger and Schmutz, 2009). Since beta-defensin genes have multiple copies in cattle (Elsik et al., 2009), a semi nested-PCR was used for genomic DNA (gDNA) in an attempt to exclude multiple products or possible pseudogenes. In the primary PCR reaction, KNEST1-F forward and KNEST1-R reverse primers (Table 1) were used to amplify a 1925 bp fragment of genomic *DEFB103* that includes the 5′UTR, both previously reported exons, the intron, and the 3′UTR. The primary PCR product was used as a template for the semi-nested PCR. In the semi-nested amplification reaction, the same forward (KNEST1-F) and the KDEFEX1-R reverse primers (Table 1) were used to amplify the 5′UTR, and exon 1 (786 bp).

In the RT-PCR, three primer pairs (Table 1) were used to characterize the cDNA of the *DEFB103* gene. All the PCR products were separated using 1% agarose gels, and then amplified products were extracted using a Gel-Extraction kit (OMEGA Bio-Tek). The extracted purified PCR products were sequenced (NRC-PBI, Saskatoon), and then analyzed using Sequencher software (version 4.8) to determine genotype and then haplotype for the seven SNPs in the 5′UTR of the *DEFB103* gene.

2.5. Haplotype determination

The 5 reported SNPs by Dreger and Schmutz (2009) formed two specific haplotypes. Two new SNPs in the 5′UTR increased the number of haplotypes from 2 to 4 (Mirabzadeh-Ardakani et al., in press). About 205 animals were genotyped for the seven SNPs in the 5′UTR of *DEFB103* (Mirabzadeh-Ardakani et al., in press). Haplotypes were initially identified in homozygous cattle. No animal presented a specific genotype that could not be explained by these 4 haplotypes (Supplementary Table 1).

3. Results

3.1. SNP haplotypes in genomic DNA

All cattle in this study were first genotyped for the seven known SNPs in the 5′UTR of *DEFB103*. Four haploid haplotypes were determined for these seven SNPs (c.−383A>G, c.−319A>G, c.−264C>T, c.−241G>A, c.−69A>G, c.−42A>G, and c.−34G>A) in Holstein cattle (Mirabzadeh-Ardakani et al., in press) (Table 2). Among the four haploid haplotypes that were identified, the haploid haplotypes 2, 3, and 4 were present in the Holstein calves, and the haploid haplotypes 2, and 4 were present in beef calves genotyped.

3.2. Characterization of cDNA start and end positions

Two primer pairs (KNEST1-F and DEFB103Ex2-R2; DEFB103-5′UTR-3 and DEFEX1-R) (Table 1, Fig. 1) were used to determine the start (in the 5′UTR) and the end (in the 3′UTR) of the *DEFB103* cDNA transcript in a tongue cDNA sample. A RT-PCR product (511 bp) was obtained using the KNEST1-F and the DEFB103Ex2-R2 primers. However, no fragment was obtained using the DEFB103-5′UTR-3 and the DEFEX1-R primers. The DEFEX1-R reverse primer had been designed in exon 1 and it must have a complementary site in the cDNA transcript. Therefore the DEFB103-5′UTR-3 forward primer in the 5′UTR did not have a complementary site in the cDNA transcript. As a result, the start of the *DEFB103* transcript was determined to be at least 514 bp before the start codon (−514, Fig. 1), and not further than 678 bp. The length of the 3′UTR was determined to be at least 53 bp after the stop codon (*53, Fig. 1).

3.3. SNP haplotypes in cDNA

At the beginning of this study, KNEST1-F and DEFEX1-R primers (Table 1) were used to amplify a 573 bp fragment of *DEFB103* and genotype the seven SNPs in the 5′UTR. However, two PCR products (573 bp and 312 bp) were amplified by using cDNA from the oral cavity tissues (tongue, buccal, hard palate), conjunctiva, rumen, bladder, trachea and liver as a template (Figs. 2a and b), but only one product was amplified (573 bp) from the skin and lung cDNA (Fig. 2a). Both PCR products were sequenced and were aligned to genomic DNA. This allowed us to

Table 1
Primers used in this study.

Primer name	Starting bp[a]	Primer sequence 5′- … -3′	Template and product size (bp)	Temp (°C)
KNEST1-F	−514	GCGTATCCATGCAGGTTCAG	Genomic, 1925	63
KNEST1-R	*545	AGGCCTGAGGTCAGTCAGAACA		
KNEST1-F	−514	GCGTATCCATGCAGGTTCAG	Genomic, 786	60
KDEFEX1-R	58+188	GATTAGTCTCTTAAGAGTGAATCAGAA		
KNEST1-F	−514	GCGTATCCATGCAGGTTCAG	cDNA, 573 or 312	63
DEFEX1-R	38	CCTGGAACAGGCAGCAAGAACA		
DEFB103-5′UTR-3	−687	AGACACCTTCCTGACTCACG	cDNA, 746	60
DEFEX1-R	38	CCTGGAACAGGCAGCAAGAACA		
KNEST1-F	−514	GCGTATCCATGCAGGTTCAG	cDNA, 511	60
DEFB103Ex2-R2	*53	CTGGAGGAGGCATTTTCACG		

[a] Based on coding DNA reference sequence.