Pakistan J. Zool., vol. 53(2), pp 705-712, 2021 DOI: https://dx.doi.org/10.17582/journal.pjz/20190929020957

Molecular Cloning and Characterization of E2f3b in Pig

Wentao Wang[1,2], Xu Lin[3], Jianshu Zhuo[3], Dongjie Zhang[2], Xiuqin Yang[3]* and Di Liu[1, 2]*

[1]*College of Wildlife and Protected Area, Northeast Forestry University, Harbin, 150040, China*
[2]*Heilongjiang Academy of Agricultural Sciences, Harbin, 150086, China*
[3]*College of Animal Science and Technology, Northeast Agricultural University, Harbin, 150030, China*

Article Information
Received 29 September 2019
Revised 11 January 2020
Accepted 20 February 2020
Available online 26 February 2021

Authors' Contribution
WWT, LX and ZJS conducted the experiments. WWT drafted the manuscript. ZDJ and YXQ revised the manuscript. LD and YXQ conceived and designed the research.

Key words
Pig, E2F3b, Alternative splicing, Expression, Subcellular localization

ABSTRACT

E2F3 is an important member of E2F transcription factors and has been involved in carcinogenesis. So far, little is known on porcine E2F3. Here, we cloned the complete coding sequence (CDS) of E2F3b and identified seven alternatively spliced transcripts in pigs using molecular biology technique for the first time. The CDS of the canonical transcripts (named V1) of porcine E2F3b is 1023 bp in length, and showed 93.35% and 90.62% identities with the homologues from human and mouse, respectively. The splicing variants were produced by exon skipping, alternative 5' and 3' splice sites alone or in combination. Minigene analysis showed that the splicing of porcine E2F3b is complicated. E2F3b isoforms are expressed in all tissues studied with high level in spleen and muscle. Both of isoforms V1 and 2, containing functional domains of E2Fs, were localized throughout cells. No functional nuclear localization sequence and export signal were characterized through site-directed mutagenesis analysis, although their existence was predicted by bioinformatic methods. The results increase our knowledge of E2F3b mRNA diversity and provide basis for in-depth functional research.

INTRODUCTIONS

The E2F transcription factors (E2Fs), discovered almost 30 years ago, play critical roles in cell differentiation, proliferation, and apoptosis (DeGregori and Johnson, 2006). So far, eight family member genes were identified and named E2F1−8 in the order of discovery (Liu *et al.* 2018). According to their functions in transcriptional regulation, these family members are subdivided into three groups, activator, repressor, and inhibitor E2Fs. E2F1, E2F2, and E2F3a are transcriptional activators, and E2F3b, E2F4, and E2F5 are transcriptional repressors. The remaining family members, including E2F6, E2F7a, E2F7b and E2F8, are classified as transcriptional inhibitor because they compete with activator and repressor E2Fs to bind to target sequences (Ertosun *et al.*, 2016).

E2F3, the major member of the E2F family, is 91.5 kb long and is located on chromosome 6p22 in human. The high expression of E2F3 has been involved in development of melanoma by interfering with the cell cycle, and thus a good target for treatment (Feng *et al.*, 2018). Aberrant E2F3 expression was also related to bladder cancer (Feber *et al.*, 2004), prostate cancer (Bilke *et al.*, 2013), and laryngeal squamous cell carcinoma (Libertini *et al.*, 2011). E2F3 therefores can be used as a factor to predict clinical stage and pathological grade of bladder cancer.

There are several isoforms of E2F3b in humans and mice deposited in GenBank, while no studies on E2F3b cDNA was reported in pig, none the less its alternatively spliced variants. In the present study, we cloned the complete coding sequence (CDS) of porcine E2F3b gene and identified seven transcript variants. At the same time, the subcellular localization and tissue expression profile was analyzed. The results increase our knowledge on E2F3b mRNA diversity and provide basis for in-depth functional research. Additionally, pig is an important animal model for human disease because of its similarity in physiology and disease development, studies on porcine E2F3 could provide a hint for further revealing its role in human.

* Corresponding author: xiuqinyang@neau.edu.cn; liudi1963@163.com
0030-9923/2021/0002-0705 $ 9.00/0

Abbreviations
AS, alternative splicing; CDS, coding sequence; E2Fs, E2F transcription factors; EEJ, exon-exon junction; GFP, green fluorescent protein; NES, nuclear export signal; NLS, nuclear localization sequence; NMD, nonsense-mediated mRNA decay; NVs, novel variants; PTC, premature translation-termination codon; SS, splice site; UTR, untranslated region.

MATERIALS AND METHODS

Samples

Two-month-old Min pigs, obtained from the Institute of Animal Husbandry, Heilongjiang Academy of Agricultural Sciences, Harbin, China, were used. Tissues including lung, liver, heart, spleen, stomach, kidney, bladder, pancreas, intestine, and muscle were collected immediately after the pigs were slaughtered, and snap frozen in liquid nitrogen, then stored at −70℃. All animal experimental treatment was conducted based on the guidelines of the Ministry of Science and Technology of China (2006).

Reverse transcription-PCR

Total RNA was isolated from tissues using Trizol reagent (Invitrogen, Carlsbad, CA, USA), and reverse-transcribed into cDNA using PrimeScript RT kit (Perfect Real Time) (Takara, Dalian, China) according to the manufacturer's instructions. Specific primers (Forward: 5'- atgcccttacagcagcag-3'; Reverse: 5'-tcaactacacataaagtcttcc-3') were designed for cloning porcine E2F3 cDNA based on electronically predicted sequences, EST in pigs and sequence similarity among species including human, mouse. PCR was performed in a final volume of 25 µl containing 1 U *Ex* Taq polymerase (Takara), 1×*Ex* Taq Buffer (Mg^{2+} Plus), 2 µl dNTP Mixture (2.5 mM each), 1 µl of each primer (10 µM), and 1 µl cDNA. PCR conditions were as follows: 95℃ for 5 min, followed by 30 cycles of 95℃ for 30 s, 60℃ for 30 s, 72℃ for 1 min, and 72℃ for 7 min.

Sequence analysis

The PCR products were subcloned into pMD18-T vector (Takara). Recombinant plasmids were first identified by colony PCR and then sequenced by Beijing Genomic Institute (BGI, Beijing, China). DNAMAN package (Version 5.2.2; Lynnon Biosoft, Quebec, Canada) and the BLASTN program at the National Center for Biotechnology Information (NCBI) website were used to analyze the sequence similarities among species. Chromosome localization, genomic structure and splice site (SS) analysis were performed using Blat program in UCSC database (http://genome.ucsc.edu/). The ExPASy database (http://www.expasy.org) was used to predict functional domain and sequence characteristics at the aa level. The nuclear localization sequence (NLS) and nuclear export sequence (NES) were predicted with the online softwares (http://nls-mapper.iab.keio.ac.jp/, http://www.cbs.dtu.dk/services/NetNES-1.1/).

Plasmids

Standard molecular cloning techniques, described elsewhere (Sambrook and Russell, 2001), were used to construct recombinant plasmids for analyzing the subcellular localization of porcine E2F3b. Complete coding sequence (CDS) of isoforms V1 and V2 were amplified by PCR using their respective cloning vector templates constructed above, *Ex* Taq polymerase (Takara), and specific primer pairs. The resulting products were subcloned into pMD18-T vector (Takara), and then transferred into pEGFP-N1 vector (Clontech, Moutain View, CA, USA) by using HindIII and KpnI enzymes (Takara). The recombinant plasmids, named pEGFP-V1 and pEGFP-V2, respectively, were verified by double-digestion and sequenced by BGI (Beijing) for confirmation.

To verify the bona fide existence of one NES and two NLSs predicted by bioinformatic methods, overlapping extension PCR was performed to create mutations in isoform V1 as described previously (Li *et al*., 2011). A total of two mutants in which NES and both NLSs were mutated, respectively, were obtained. For NES mutagenesis, two overlapping fragments, each containing mutant sequences at one end, were first amplified using pEGFP-V1 template, primer pairs with mismatched base pairs and high fidelity pfu DNA polymerase (Transgene, Beijing, China), and then spliced into a big one by PCR which brought the mutation into the interior. For double mutation of NLSs, NLS1, located in the 5' end of isoform V1, was first mutated with one-step PCR using forward primer containing mismatched bases and normal reverse primer. And then, overlapping extension PCR was performed to mutate NLS2 with NLS1 mutant template as described above.

To further identify alternative splicing (AS) variants of porcine E2F3b, a minigene containing partial sequences of genomic DNA was constructed. Two fragments spanning from exon 4 through the 5' end of intron 4 and from the 3' end of intron 4 through exon 6 were PCR amplified, respectively, and then ligated with restriction endonuclease EcoRI introduced into by the 5' end of primer. The resultant products were subcloned into pMD18-T vector (Takara), and then transferred into pcDNA3.1+ (Invitrogen) by using HindIII and XhoI enzymes. The primers used for plasmid construction are listed in Table I and the primer pairing scheme is given in Table II.

Phylogenetic tree construction

The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The optimal tree with the sum of branch length = 1.13863267 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) are shown next to the branches (Felsenstein, 1985). The tree is drawn to scale, with branch lengths in

the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of the number of amino acid substitutions per site. The analysis involved 19 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 177 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar *et al*., 2016).

Table I. Primers used for plasmid construction.

Name	Sequences (5'-3')	Purpose
A	F: aagcttatgcccttacagcagcagg R: ggtacccgactacacataaagtcttcc	pEGFP-V1 construction
B	F: the same as AF R: ggtaccctatgggtccttgggtactt	pEGFP-V2 construction
E	F: ccactactactactacgttaaccgaggatt R: ttaacgtagtagtagtagtggtgcagcttt	NES mutation
C	F: aagcttatgcccttacagcagcagg-cgaattgctggctg	NLS1 mutation
D	F: caaaacgaggagatttacgacatca R: aaatctcctcgttttgcaccttca	NLS2 mutation
M1	F: aacgttgggctgcagtctgtctgaggac R: gaattccatcagtttttggtgttgcga	Minigene construction First fragment
M2	F: gaattcttactctctttaatagccatggcc R: ctcgagctttggaagtgggtttagggata	Minigene construction Second fragment

The enzyme sites were underlined; the mutated sequences were boxed.

Table II. Primer pairing scheme for overlapping extension PCR.

Sites	First round PCR		Second round PCR
	Reaction 1	Reaction 2	
NES	AF/ER	EF/AR	AF/AR
NLS1	CF/AR		
NLS1+2	AF/DR	DF/AR	AF/AR

Cell culture, transfection, and microscopy

PK-15 cells were cultured in Dulbecco's modified Eagle medium (Solarbio, Shanghai, China) supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin (Invitrogen) as described previously (Li *et al*., 2011). Transient transfection was performed with Lipofectamine 2000 reagent (Invitrogen). At 24 h after transfection, the cells were collected directly for minigene analysis or treated as described previously (Li *et al*., 2016), that is, stained with DAPI (Beyotime, Shanghai, China), and viewed with a Zeiss Axiovert 200 M fluorescence microscope (Carl Zeiss AG, Jena, Germany). The images were obtained with an AxioCam MRc5 color camera and collected using AxioVision Rel software.

Real-time quantitative and nested PCR

Real-time quantitative PCR (qPCR) was performed to characterize the tissue expression profile of porcine E2F3b gene as described previously (Li *et al*., 2016). Briefly, the reaction was performed in a final volume of 10 μL with SYBR Green detection (Takara) in accordance with the manufacturer's instruction. The relative mRNA levels were measured with the $2^{\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001), and β-actin was used as a reference. Each sample was run in triplicate to rule out between-run variations. Primer pairs used for qPCR are listed in Table III.

Table III. Primers used for real-time quantitative PCR.

Name	Sequences (5'-3')	Purpose
R1	F: ccagtgtcaaggcctgtcaaaa R: ttctggaggggctttcacaact	Exon 5-containing transcripts amplification
R2	F: tatcctttgtgccacctgttcat R: attcccattgtggtcttggttgt	V2 amplification
R3	F: catcaccatcggcaacga R: gcgtagaggtccttcctgatgt	β-actin amplification

Nested PCR was first performed with pcDNA3.1+ specific primers, T7/SP6, and cDNA from cell samples. Then the resultant products were used as template to perform second round PCR in which the primers were E2F3b specific, M1F/2R. The final products were sub cloned into pMD18-T vector (Takara). Positive clones were first identified with colony PCR together with 3% agarose gel electrophoresis. Those clones carrying fragments with different length were sequenced by BGI (Beijing).

RESULTS

RT-PCR amplification

The RT-PCR products, amplified with specific primers and equally mixed cDNAs obtained from various tissues including lung, heart, spleen, and muscle, were a mixture of fragments with different length as revealed by 1.5% agarose gel electrophoresis (Fig. 1). Considering the potential for AS of porcine E2F3b gene, these bands were purified as a whole and subcloned into pMD18-T vector. After screened by colony PCR, eight fragments (named V1−V8) with different length were obtained and sequence analysis showed that all of them were transcript variants of porcine E2F3b gene. These sequence data have been submitted to the GenBank database under accession numbers MH143790−91 for V1 and V2, MH086520 for V3, and MH117934−38 for V4−8.

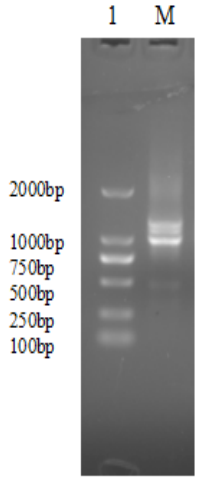

Fig. 1. Electrophoresis pattern of RT-PCR products. Marker was DL 2000 composed of 2000-, 1000-, 750-, 500-, 250-, and 100-bp DNA fragments.

Sequence of canonical transcript

Transcript variant V1, 1023 bp long, contains a complete CDS encoding a polypeptide of 340 aa (Fig. 2). The sequence identities of the CDS with those of E2F3b in humans (NM_001243076.2) and mice (GenBank No. NM_001289920.1) are 93.35% and 90.62%, respectively. At the aa level, the identities are 95.88% and 95.59%, respectively. The identities between V1 obtained and human E2F3a (NP_001940.1) is 70.32%.

```
1    ATGCCCTTACAGCAGCAGGCGAAGCGCAGGCTGGAGCTAGGAGAGAGCGGTCAGCAGTAC
1      M  P  L   Q  Q  Q  A  K  R  R  L  E  L  G  E  S  G  Q  Q  Y
61   CTCGCAGATGGTTTAAAAACCCCCAAGGGCAAAGGGAGAGCTGCGCTGCGAAGTCCAGAT
21   L  A  D  G  L  K  T  P  K  G  K  G  R  A  A  L  R  S  P  D
121  AGCCCCAAAACTCCAAAATCTCCCTCAGAAAAAACACGGTATGATACATCCCTTGGTCTG
41   S  P  K  T  P  K  S  P  S  E  K  T  R  Y  D  T  S  L  G  L
181  CTCACCAAGAAGTTCATTCAACTACTGAGCCAATCACCTGATGGGGTCTTGGATTTGAAC
61     L  T  K  K  F  I  Q  L  L  S  Q  S  P  D  G  V  L  D  L  N
241  AAAGCAGCAGAGGTGCTGAAGGTGCAAAAGAGAAGGATTTACGACATCACCAATGTACTG
81     K  A  A  E  V  L  K  V  Q  K  R  R  I  Y  D  I  T  N  V  L
301  GAAGGCATCCACCTCATTAAGAAGAAGTCTAAAAACAACGTGCAGTGGATGGGCTGCAGT
101    E  G  I  H  L  I  K  K  K  S  K  N  N  V  Q  W  M  G  C  S
361  CTGTCTGAGGACGGGGGCACGCTGGCCCAGTGTCAAGGCCTGTCAAAAGAAGTGACCGAG
121    L  S  E  D  G  G  T  L  A  Q  C  Q  G  L  S  K  E  V  T  E
421  CTCAGTCAGGAAGAGAAGAAATTAGATGAACTGATCCAAAGCTGCACCCTGGACCTCAAA
141    L  S  Q  E  E  K  K  L  D  E  L  I  Q  S  C  T  L  D  L  K
481  CTGTTAACCGAGGATTCAGAGAATCAAAGGTTAGCTTATGTTACATATCAAGATATTCGA
161    L  L  T  E  D  S  E  N  Q  R  L  A  Y  V  T  Y  Q  D  I  R
541  AAAATTAGTGGCCTTAAAGACCAAACTGTTATAGTTGTGAAAGCCCCTCCAGAAACAAGA
181    K  I  S  G  L  K  D  Q  T  V  I  V  V  K  A  P  P  E  T  R
601  CTTGAAGTGCCTGACCCAATAGAGAGCCTACAAATACATTTGGCAAGTACCCAAGGACCC
201    L  E  V  P  D  P  I  E  S  L  Q  I  H  L  A  S  T  Q  G  P
661  ATTGAGGTTTATTTGTGTCCAGAAGAGACTGAAACACACAGTCCAATGAAAACAAACAAC
221    I  E  V  Y  L  C  P  E  E  T  E  T  H  S  P  M  K  T  N  N
721  CAAGACCACAATGGGAATATCCCTAAACCCACTTCCAAAGACTTGGCTTCAACCAACTCA
241    Q  D  H  N  G  N  I  P  K  P  T  S  K  D  L  A  S  T  N  S
781  GGACATAGTGATTGCTCGATTTCTATGGCAAACCTTTCTCCTTTGGCCTCCCCAGCCAAC
261    G  H  S  D  C  S  I  S  M  A  N  L  S  P  L  A  S  P  A  N
841  CTTTTACAGCAGACTGAGGACCAAATTCCTTCCAACCTAGAAGGACCATTTGTGAACTTA
281    L  L  Q  Q  T  E  D  Q  I  P  S  N  L  E  G  P  F  V  N  L
901  CTGCCTCCCCTGCTCCAAGAAGACTATCTCCTAAGCCTTGGGGAGGAAGAAGGCATCAGC
301    L  P  P  L  L  Q  E  D  Y  L  L  S  L  G  E  E  E  G  I  S
961  GATCTCTTTGATGCTTATGATTTGGAAAAGCTCCCACTGGTGGAAGACTTTATGTGTAGT
321    D  L  F  D  A  Y  D  L  E  K  L  P  L  V  E  D  F  M  C  S
1021TGA
341  *
```

Fig. 2. cDNA sequences of transcript variant 1 of E2F3b in pig. Putative nuclear localization sequences were indicated in bold, and nuclear export signal was boxed. E2F_TDP domain was underlined. E2F_DD domain was shadowed.

The predicted polypetide has a theoretical molecular mass of 37.62 kDa and an isoelectric point of 5.00. It also has characteristic structure of E2F transcription factors. There are an DNA-binding domain, E2F_TDP, at aa position 54−118, and a dimerization domain, E2F_DD, at position 129−229. Additionally, two NLSs located at aa position 6−34 and 87−96 (named NLS1 and NLS2), respectively, and one NES at position 157−161 were predicted by bioinformatic methods (Fig. 3).

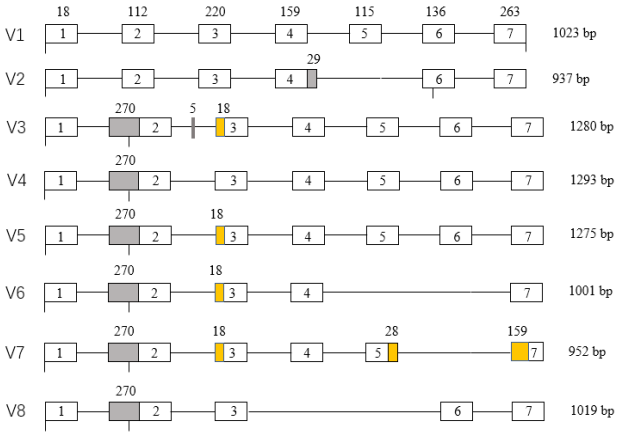

Fig. 3. Structure of eight transcript variants of porcine E2F3b gene. Boxes indicate exons with closed gray and yellow boxes indicating the sequence added or deleted, respectively. Lines indicate introns. Short vertical lines under boxes indicate the position of start and stop codon sequentially.

Alternative splicing

Compared with V1, transcript V2 results from exon 5 skipping combined with alternative 5' SS of exon 4 leading to 29 bp intron retention. The amplified V2 cDNA is 973 bp long containing a CDS of 579 bp encoding a polypeptide of 192 aa and a 3' untranslated region (UTR) of 394 bp. The first 509 bp of V2 is the same as that of V1 resulting in the first 170 aa of the two polypeptides identical. Thus, isoform V2 also contains the same putative NLSs, NES, and E2F_TDP domain as V1. All the remaining six transcript variants (V3−8) have alternative 3' SS in exon 2 resulting in a 270 bp intron retention, which leading to the formation of premature translation-termination codon (PTC). V3−8 encode the same C-terminal truncated polypeptides of 59 aa having putative NLS1 site. Skipping of exons including 3, 4, 5, 6, or 7 (partial or complete) was also found in the splicing of V3-8. Additionally, a 5 bp sequence from inner of intron 2 was present in V3 (Fig. 3).

Among the seven exons of V1, six (exons 2-7) are involved in alternative splicing. By using Blat program, SSs of each alternative exons were identified (Table IV).

Exons 2, 3, and 7 have two 3' SSs, and exons 4 and 5 have two 5' SSs. While exon 6 have no alternative SS in that it was completely skipped in alternative splicing of porcine E2F3b gene. Besides classical GT-AG sequences in the SSs, there are AT, AA and TT in 5' SSs, and TT, AA and AC in 3' SSs. The presence of GT-AG is dominant.

Table IV. Splice site of alternative exons in porcine E2F3b gene.

Exons	3' splice sites		5' splice sites	
	Number	Sequences	Number	Sequences
2	2	tt/aa	1	gt
3	2	ag/ag	1	gt
4	1	ag	2	at/gt
5	1	ac	2	gt/aa
6	1	ag	1	gt
7	2	ag/ag	1	tt

The first splice site is present in variant V1.

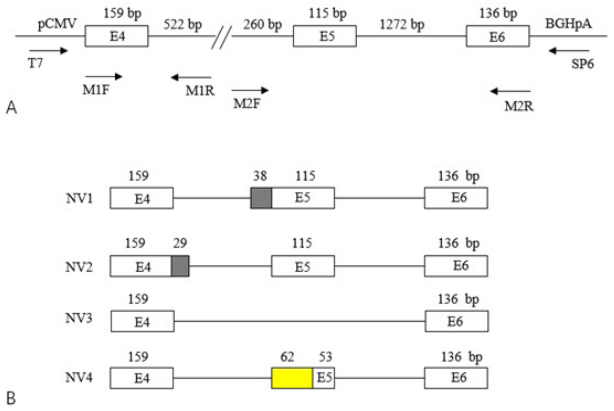

Fig. 4. Identification of alternative splicing of porcine E2F3b gene by minigene construct. A, Schematic of the minigene construct. Boxes indicate exons, lines indicate introns. Arrows indicate the location of primers. B, Sequence composition of novel alternative splicing variants (NV1-4). The black and yellow boxes indicate the sequences are added and deleted, respectively.

Minigene analysis

We used minigene system to further identify splicing patterns of porcine E2F3b gene. The minigene included complete sequences of exons 4−6 and intron 5, and partial sequences of intron 4 (Fig. 4A). By using nested PCR, canonical transcript composed of complete exons 2, 3, and 4, was obtained exclusively from minigene, indicating that the system contains necessary splice signals including SSs and branch points. A total of four novel variants (NVs) were amplified here, and the splicing pattern includes alternative 5' SSs, 3' SSs and exon skipping. Two novel alternative SSs were identified in exon 5 (Fig. 4B).

Evolutionary relationship of E2f3b

To gain insight into the evolutionary relationship of E2F3b among species, a phylogenetic tree was constructed using seven representative species including human, pig, chicken, platypus, African frog, Atlantic salmon, and pike (Fig. 5). We found that the isoforms of E2F3b are not paralogs; they all came from the common ancestor in the species and are highly conservative. Two isoforms, X1 and X2 that were identified by us, exist in pig and are highly homologous to isoform X3 in human and chicken. We therefore concluded that pig E2F3b belonged to isoform X3 of E2F3 family.

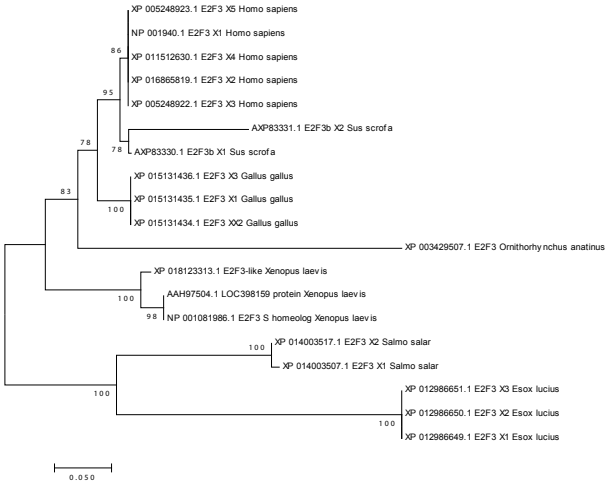

Fig. 5. Evolutionary relationships of E2F3 isforms.

Subcellular localization

To gain insights into the subcellular localization of V1 and V2, two isoforms of porcine E2F3b encoding polypetides with functional domain and putative NLSs and NES, the cDNAs of ORFs were fused to the N-terminus of the green fluorescent protein (GFP). In transiently transfected PK-15 cells, fluorescence produced by each of the two fusion proteins was distributed throughout the cells, similar to that produced by empty vector, pEGFP-N1, used as a control (Fig. 6A), suggesting that the two isoforms do not have functional NLSs.

To further characterize the role of putative NLSs, the putative NES in isoform V1 was mutated. When the fusion protein carrying mutant type V1 was transfected into PK-15 cells, the fluorescence was also distributed throughout the cells. At the same time, we mutated the both putative NLSs simultaneously in isoform V1 to analyze the role of

NES sequence. Similarly, the fluorescence produced by the mutant fusion protein distributed throughout the cells (Fig. 6B). These results indicate that the putative NLSs or NES are not bona fide.

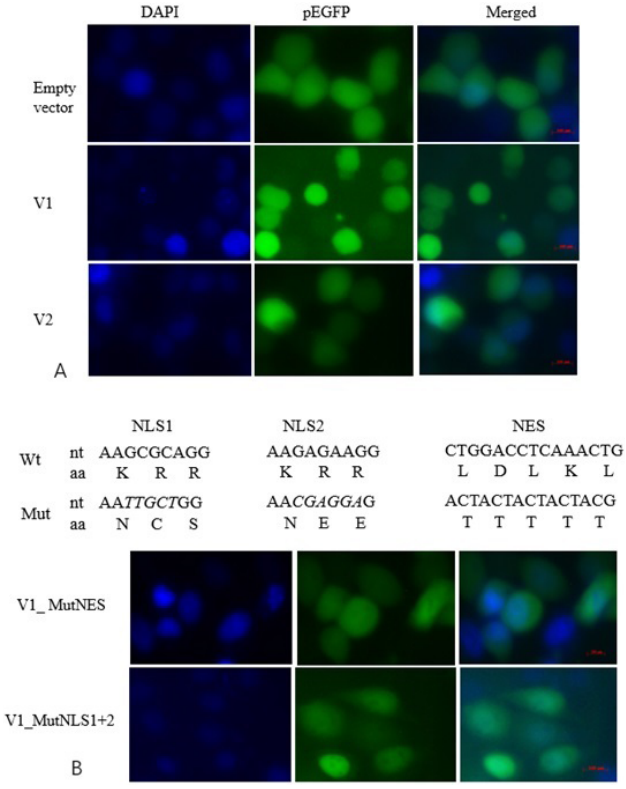

Fig. 6. Subcellular localization and putative NLSs/NES identification. A, Subcellular localization of two isoforms of porcine E2F3b. B. Effect of putative NLSs/NES on E2F3b subcellular localization. Wt, wild type; Mut, mutant; nt, nucleotide; aa, amino acids. The bar is 100 μm.

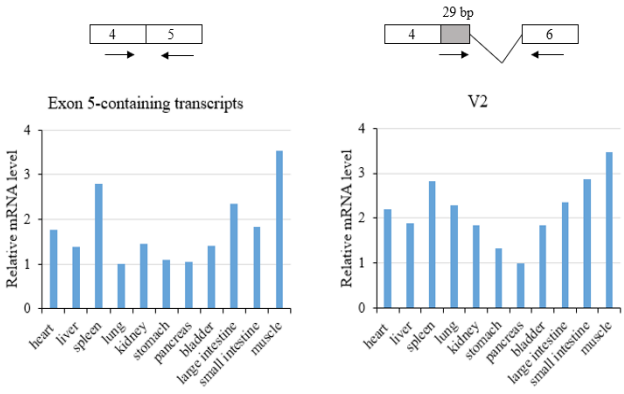

Fig. 7. Expression profiles of porcine E2F3b. In the upper panel, boxes indicate exons, while closed box indicates intron retained. The slashes indicate skipping of exon 5.

Tissue expression profile

To gain insights into the tissue expression profile, qPCR was used to measure relative mRNA level of porcine E2F3b gene. As there are no sequences specific to V1, we can not find a specific primer. Therefore, the expression of exon 5-containing transcripts, including V1, 3-5, and 7 was analyzed as a whole. At the same time, the 29 bp-intron retention make it possible to design specific primers for V2. The results showed that they had similar expression profiles in tissues. Both of exon 5-containing transcripts and V2 are expressed in all tissues studied, with high level in spleen and muscle (Fig. 7).

DISCUSSION

As a transcription factor, E2F3 plays an important role in cell cycle and proliferation, which is closely related to tumor progression (Hoeijmakers, 2000). Here, we first cloned the complete CDS of E2F3b in pig and obtained several AS variants using molecular biology methods. Additionally, the subcellular localization and expression profiles were characterized. The results will contribute to further revealing the role of E2F3 which is important in cancer research and therapy.

AS, generating different mature mRNAs from the same pre-mRNA, is a fundamental mechanism for proteome diversity and complexity in that it yields different protein isoforms with distinct cellular roles (Sanchez *et al*., 2011). It has been shown that approximately 95% of human genes are processed by AS to yield multiple transcript variants (Nilsen and Graveley, 2010). Five basic modes of AS including exon skipping, mutually exclusive exons, alternative 5' SS, alternative 3' SS, and intron retention, are generally recognized. Among which exon skipping is the most common mode in mammals. Additionally, there are two other main modes, multiple promoters and multiple polyadenylation sites, by which transcript variants may be generated (Sammeth *et al*., 2008).

Here, we focused on characterization of AS in CDS, and did not refer to the promoter and polyadenylation sites. Exon skipping, alternative 5' SS, and 3' SS were found in the AS of porcine E2F3b. A total of five distinct alternative SSs including three 5' SSs and two 3'SSs were identified, showing a percentage of alternative 5' SSs > alternative 3'SSs > exon skipping. The result is divergent from previous reports (Bergsma *et al*., 2018). However, we used agarose gel electrophoresis whose sensitivity is low to screen RT-PCR products with different length preliminarily. It can not differentiate fragments with similar length or those with similar length but having different splicing pattern. So, we speculate that there might exist other splicing variants. To verify this, minigene analysis was performed in transiently

transfected PK-15 cells through which several novel splicing pattern of porcine E2F3b were identified.

Studies have shown that minigenes are useful tools to test gene splicing in cells (Li *et al*., 2016), and functionally identify the effect of mutation on splicing (Acedo *et al*., 2015; Fraile-Bethencourt *et al*., 2018). To avoid amplification of endogenous transcripts, nested PCR was performed with outer primers complementary to vector sequence and inner primers specific to E2F3b. Another advantage of nested PCR is to amplify transcripts with very low expression that otherwise can not be amplified by regular PCR. When screened with colony PCR, the canonical transcript, that is, composed of complete exons 4-6, is the most frequently amplified, indicating its abundant expression level, while NVs were only present in very low frequency. This might be the reason that NVs were not identified *in vivo*. Here, we still used agarose gel electrophoresis to screen different splicing patterns, and we did not try our best to identify new variants. Thus, we can conclude that the splicing of porcine E2F3b is complicated and that there should many transcript variants remain to be identified.

The products of AS include not only polypeptides with function but also truncated polypetides less than 100 aa having toxic effects on cells due to the formation of PTC. To avoid the accumulation of truncated polypetides, the cells develop a surveillance mechanism, nonsense-mediated mRNA decay (NMD), to destroy aberrant transcripts with PTC. It was shown that mRNAs with a PTC more than 55 nt upstream of the 3'-most exon-exon junction (EEJ) can be degraded by NMD (McGlincy and Smith, 2008). The products of V3−8 are polypetides of 55 aa and their PTC locates at 220 nt upstream of the 3'-most EEJ, suggesting they are might be degraded by NMD.

Although NLSs and NES were predicted in isoforms V1 and 2 using bioinformatic methods, we could not verify their existence through transient transfection and site-directed mutagenesis. As a transcription factor, it should have two fundamental structures, DNA-binding and dimerization domain. Both of them are present in isoform V1, while isoform V2 only has DNA binding domain, as revealed by sequence analysis. Sequence and structure are the basis for function, we therefore concluded that isoform V2 could not function as a transcription factor. While V2 ubiquitously expressed in all tissues studied, suggesting a role in cell growth and physiology. Further efforts should be made to ascertain the function of V2.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (31741114), Foundation for Improving Innovative Capability of Scientific Institutions, Heilongjiang (YC2016D001), and National Science and Technology Planning Project of "12th Five-Year" in Rural Areas (2015BAD03B02-5).

Statement of conflict of interest

Authors have declared no conflict of interest.

REFERENCES

Acedo, A., Hernández-Moro, C., Curiel-García, Á., Díez-Gómez, B. and Velasco, E.A., 2015. Functional classification of BRCA2 DNA variants by splicing assays in a large minigene with 9 exons. *Hum. Mutat*., **36**: 210-221. https://doi.org/10.1002/humu.22725

Bergsma, A.J., van der Wal, E., Broeders, M., van der Ploeg, A.T. and Pim-Pijnappel, W.W.M., 2018. Alternative splicing in genetic diseases: Improved diagnosis and novel treatment options. *Int. Rev. Cell mol. Biol*., **335**: 85-141. https://doi.org/10.1016/bs.ircmb.2017.07.008

Bilke, S., Schwentner, R., Yang, F., Kauer, M., Jug, G., Walker, R.L., Davis, S., Zhu. Y.J., Pineda, M., Meltzer, P.S. and Kovar, H., 2013. Oncogenic ETS fusions deregulate E2F3 target genes in Ewing sarcoma and prostate cancer. *Genome Res.,* **23**: 1797-1809.

DeGregori, J. and Johnson, D.G., 2006. Distinct and overlapping roles for E2F family members in transcription, proliferation and apoptosis. *Curr. mol. Med*., **6**: 739-748. https://doi.org/10.2174/156652401060607073 9

Ertosun, M.G., Hapil, F.Z. and Osman Nidai, O., 2016. E2F1 transcription factor and its impact on growth factor and cytokine signaling. *Cytokine Growth Factor Rev.*, **31**: 17-25. https://doi.org/10.1016/j.cytogfr.2016.02.001

Feber, A., Clark, J., Goodwin, G., Dodson, A.R., Smith, P.H., Fletcher, A., Edwards, S., Flohr, P., Falconer, A., Roe, T., Kovacs, G., Dennis, N., Fisher, C., Wooster, R., Huddart, R., Foster, C.S. and Cooper, C.S., 2004. Amplification and overexpression of E2F3 in human bladder cancer. *Oncogene*, **23**: 1627-1630. https://doi.org/10.1038/sj.onc.1207274

Felsenstein, J., 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, **39**: 783-791. https://doi.org/10.1111/j.1558-5646.1985.tb00420.x

Feng, Z., Peng, C., Li, D., Zhang, D., Li, X., Cui, F., Chen, Y. and He, Q., 2018. E2F3 promotes cancer growth and is overexpressed through copy

number variation in human melanoma. *Oncotargets Ther*., **11**: 5303-5313. https://doi.org/10.2147/OTT.S174103

Fraile-Bethencourt, E., Valenzuela-Palomo, A., Díez-Gómez, B., Acedo, A. and Velasco, E.A., 2018. Identification of eight spliceogenic variants in BRCA2 exon 16 by minigene assays. *Front. Genet*., **9**: 188. https://doi.org/10.3389/fgene.2018.00188

Hoeijmakers, J.H., 2001. Genome maintenance mechanisms for preventing cancer. *Nature*, **411**: 366-374. https://doi.org/10.1038/35077232

Kumar, S., Stecher, G. and Tamura, K., 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol*., **33**: 1870-1874. https://doi.org/10.1093/molbev/msw054

Li, H.T., Liu, D. and Yang, X.Q., 2011. Identification and functional analysis of a novel single nucleotide polymorphism (SNP) in the porcine Toll-like receptor (TLR) 5 gene. *Acta Agr. Scan A-AN.,* **61**: 161-167. https://doi.org/10.1080/09064702.2012.656140

Li, L.Z., Wang, Q.S., Han, L.X., Wang, J.K., Shao, S.Y., Wang, L., Liu, D. and Yang, X.Q., 2016. Molecular characterization of Sp110 gene in pigs. *Mol. Genet. Genom*., **291**: 1431-1442. https://doi.org/10.1007/s00438-016-1189-9

Libertini, S.J., Chen, H., al-Bataina, B., Koilvaram, T., George, M., Gao, A.C. and Mudryj, M., 2011. The interleukin 6 receptor is a direct transcriptional target of E2F3 in prostate tumor derived cells. *Prostate*, **72**: 649-660. https://doi.org/10.1002/pros.21468

Liu, Z.L., Bi, X.W., Liu, P.P., Lei, D.X., Wang, Y., Li, Z.M., Jiang, W.Q. and Xia, Y., 2018. Expressions and prognostic values of the E2F transcription factors in human breast carcinoma. *Cancer Manage. Res.,* **10**: 3521-3532. https://doi.org/10.2147/CMAR.S172332

Livak, K.J. and Schmittgen, T.D., 2001. Aanalysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods*, **25**: 402-408. https://doi.org/10.1006/meth.2001.1262

McGlincy, N.J. and Smith, C.W., 2008. Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? *Trends Biochem. Sci*., **33**: 385-393. https://doi.org/10.1016/j.tibs.2008.06.001

Ministry of Science and Technology of China, 2006. *Guidelines on humane treatment of laboratory animals*. Available at http:/ /www.most.gov.cn/fggw/zfwj/ zfwj2006/200609/t20060930_54389.htm. Accessed 30 Sept 2006

Nilsen, T.W. and Graveley, B.R., 2010. Expansion of the eukaryotic proteome by alternative splicing. *Nature*, **463**: 457-463. https://doi.org/10.1038/nature08909

Saitou, N. and Nei, M., 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol*., **4**: 406-425.

Sambrook, J. and Russell, D.W., 2001. *Molecular cloning: A laboratory manual, 3rd edn*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sammeth, M., Foissac, S. and Guigó, R., 2008. A general definition and nomenclature for alternative splicing events. *PLoS Comput. Biol*., **4**: e1000147. https://doi.org/10.1371/journal.pcbi.1000147

Sanchez, S.E., Petrillo, E., Kornblihtt, A.R. and Yanovsky, M.J., 2011. Alternative splicing at the right time. *RNA Biol*., **8**: 954-959. https://doi.org/10.4161/rna.8.6.17336

Zuckerkandl, E. and Pauling, L., 1965. Evolutionary divergence and convergence in proteins. In: *Evolving genes and proteins* (eds. V. Bryson and H.J. Vogel). Academic Press, New York. pp. 97-166. https://doi.org/10.1016/B978-1-4832-2734-4.50017-6