Journal of Reproduction and Development, Vol. 64, No 1, 2018

—Original Article—

Identification and expression analysis of cDNA encoding insulin-like growth factor 2 in horses

Kohta KIKUCHI[1)*], Keisuke SASAKI[1)*#], Hiroki AKIZAWA[1)], Hayato TSUKAHARA[1)], Hanako BAI[1)], Masashi TAKAHASHI[1)], Yasuo NAMBO[2)##], Hiroshi HATA[3)] and Manabu KAWAHARA[1)]

[1)]*Laboratory of Animal Genetics and Reproduction, Research Faculty of Agriculture, Hokkaido University, Hokkaido 060-8589, Japan*

[2)]*Equine Science Division, Hidaka Training and Research Center, Japan Racing Association, Hokkaido 057-0171, Japan*

[3)]*Field Science Center for Northern Biosphere, Hokkaido University, Hokkaido 060-0811, Japan*

[#] *Present: Department of Bioscience, Tokyo University of Agriculture, Tokyo 156-8502, Japan*

[##]*Present: Department of Clinical Veterinary Sciences, Obihiro University of Agriculture and Veterinary Medicine, Hokkaido 080-8555, Japan*

Abstract. Insulin-like growth factor 2 (IGF2) is responsible for a broad range of physiological processes during fetal development and adulthood, but genomic analyses of *IGF2* containing the 5′- and 3′-untranslated regions (UTRs) in equines have been limited. In this study, we characterized the *IGF2* mRNA containing the UTRs, and determined its expression pattern in the fetal tissues of horses. The complete equine *IGF2* mRNA sequence harboring another exon approximately 2.8 kb upstream from the canonical transcription start site was identified as a new transcript variant. As this upstream exon did not contain the start codon, the amino acid sequence was identical to the canonical variant. Analysis of the deduced amino acid sequence revealed that the protein possessed two major domains, IlGF and IGF2_C, and analysis of *IGF2* sequence polymorphism in fetal tissues of Hokkaido native horse and Thoroughbreds revealed a single nucleotide polymorphism (T to C transition) at position 398 in Thoroughbreds, which caused an amino acid substitution at position 133 in the IGF2 sequence. Furthermore, the expression pattern of the *IGF2* mRNA in the fetal tissues of horses was determined for the first time, and was found to be consistent with those of other species. Taken together, these results suggested that the transcriptional and translational products of the *IGF2* gene have conserved functions in the fetal development of mammals, including horses.
Key words: Amino acid sequence, Horse, Insulin-like growth factor 2, Untranslated region

(**J. Reprod. Dev. 64:** 57–64, 2018)

Insulin like growth factor 2 (IGF2) regulates a wide range of important processes, such as cell growth and proliferation, and metabolic activities at the cellular and physiological levels. Moreover, it is also associated with a variety of productive traits, including milk and meat production, and progeny weight [1–6]. Hence, the gene encoding IGF2 is of great interest to animal breeders because of the critical roles performed by this protein [7–9].

The expression pattern of *IGF2* is unique in mammals. Although the progeny inherits two sets of chromosomes from its parents through fertilization, *IGF2* expresses primarily from the paternal allele, whereas the maternal allele is silenced during fetal development [6, 10]. Furthermore, the expression pattern of *IGF2* varies with tissue type. That *IGF2* expresses mainly from the paternal allele appears to be universal among species, however [11–13], and therefore, characterization of *IGF2* across species may provide further insight into the significance of genomic imprinting in mammals.

Previous studies have isolated and sequenced horse *IGF2* cDNA [14, 15], but genetic analyses of the equine *IGF2* gene, including examination of the 5′- and 3′-untranslated regions (UTRs) of *IGF2* and its protein domains, are scarce. The UTRs, which contain important regulatory elements [16], play important roles in both transcriptional and post-transcriptional regulation of gene expression [17, 18]. Recently, it has been demonstrated that the different 5′-UTRs of mouse *Igf2* variants mediated translational control in embryonic stem and neural precursor cells [19]. Moreover, microRNAs that bind to the 3′-UTR of mouse *Igf2* have been identified and shown to play an important role in the regulation of *Igf2* during placentation [20]. These and other studies highlight the need for further analyses of UTR sequences in order to gain a better understanding of the regulation of *IGF2* expression.

Here, we determined the 5′- and 3′-UTRs of equine *IGF2* via rapid amplification of cDNA ends (RACE) using Hokkaido native horse fetus, and identified the coding sequence (CDS) of this gene. We then also compared the nucleic acid and predicted amino acid sequences of equine *IGF2* with those of other mammals, and examined *IGF2*

Received: September 12, 2017
Accepted: October 30, 2017
Published online in J-STAGE: November 17, 2017

Correspondence: M Kawahara (e-mail: k-hara@anim.agr.hokudai.ac.jp)
* K Kikuchi and K Sasaki contributed equally to this work.

polymorphism in the fetal tissues of Hokkaido native horse and Thoroughbreds. Finally, we performed a quantitative polymerase chain reaction (qPCR) analysis of *IGF2* using primary tissues derived from an equine fetus retrieved from an early pregnant mare. Horses are monotocous, in contrast to polytocous species such as mice and pigs; moreover, gestation periods of large livestock, including that of equines, is generally far longer than many other experimental animals, making collection of fetal samples extremely difficult, and as the demand for horsemeat is considerably lower than that of beef, opportunities for obtaining samples from slaughterhouse facilities is limited. In addition, because horse breeding is seasonal, collecting samples of fetuses of synchronized embryonic days is highly labor and cost intensive. Because of these challenges, data on equine genetics in official databases are usually insufficient for determining genetic differences between breeds. In cattle, a single fetal sample has been analyzed for global gene expression using microarrays [21]; thus, although the sample size in the present study is comparatively small, our discovery of a novel splicing variant of equine *IGF2* led us to perform comparative analyses between different breeds in an effort to better understand the characteristics of equine *IGF2*, not only from the perspective of animal breeding but also with respect to the role of this gene in developmental and evolutionary biology.

Materials and Methods

Rapid amplification of cDNA ends (RACE) for determining the full-length IGF2 cDNA

A conceptus (embryonic day: E56) was recovered from a Hokkaido native horse mare that was given a lethal injection following an accidental leg fracture (Fig. 1A). Because the imprinted gene transcripts are enriched in extraembryonic tissues, the yolk sac of the conceptus was homogenized using a BioMasher® (Nippi, Tokyo, Japan), and total RNA was extracted with a NucleoSpin® RNA II kit (MACHEREY-NAGEL, Düren, Germany) following the manufacturer's instructions; RNA was quantitated using a spectrophotometer (NanoDrop ND-2000, Thermo Scientific, Wilmington, DE, USA) and stored at –80°C until use. To determine the complete equine *IGF2* mRNA sequences, the 5′- and 3′-UTRs were amplified and cloned using a GeneRacer Kit (Invitrogen, Carlsbad, CA, USA), in accordance with the manufacturer's instructions. The primer sets used for this purpose are shown in Table 1. Based on the known sequence of the annotated equine *IGF2* cDNA (GenBank ID: NM_001114539.1), the eqIGF2-114F and GeneRacer 3′ primers were designed for 3′- RACE PCR. For 5′- RACE, the eqIGF2-216R and GeneRacer 5′ primers were used. To further screen more specific amplicons for 5′- and 3′- UTRs, nested PCRs were performed using GeneRacer 5′-nested and GeneRacer 3′-nested primers. The *IGF2* CDS was also amplified by GoTaq Polymerase (Promega) using eq*IGF2* CDS F and P primers designed from within the UTRs, with PCR conditions consisting of 40 cycles of 94°C for 30 sec, 65°C for 1 min, and 72°C for 3 min. The PCR amplicon was ligated into pGEM-T Easy Vector (Promega, Madison, WI, USA). Plasmid DNAs were sequenced with an Applied Biosystems 3130 Genetic Analyzer (Applied Biosystems, FosterCity, CA, USA) and a BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystems). The *IGF2* mRNA sequence was then compared to available high-throughput RNA sequencing (RNA-seq) datasets from a fetus at E34 (SRX273058, SRX277447), spermatozoa (SRX154655, SRX154656), testes (SRX277445, SRX273056), and placental villi at 0 day post birth (SRX273055, SRX277444) in the Sequence Read Archives (SRA; https://www.ncbi.nlm.nih.gov/sra) using the Nucleotide BLAST tool.

Comparison of equine IGF2 nucleotide and predicted amino acid sequences between breeds and species

Equine *IGF2* CDS was compared between Hokkaido native horse and Thoroughbreds. The cDNAs from the livers of two Thoroughbred fetuses (E79 and E88) were prepared in the same manner as that described above for the Hokkaido native horse fetus. We also compared the predicted amino acid sequences of the equine IGF2 between the two breeds. Protein domain analysis was performed based on two public collections of conserved domain models, namely Pfam (http://pfam.xfam.org/) [22] and SMART (http://smart.embl-heidelberg.de/) [23]. To investigate similarities in the *IGF2* CDSs and in the predicted amino acid sequences among different species, the determined equine *IGF2* sequence was compared to that of humans (GenBank ID: NM_001127598.2), mice (GenBank ID: NM_010514.3), rats (GenBank ID: NM_031511.2), cattle (GenBank ID: NM_174087.3), and pigs (GenBank ID: NM_213883.2). The IGF2 amino acid sequences were also compared among the species using ClustalW algorithm (http://www.genome.jp/tools-bin/clustalw) [24].

qPCR analysis in horse fetal tissues

Total RNA was extracted from different tissues using an RNeasy Mini Kit (QIAGEN, Tokyo, Japan), including the neck, yolk sac, kidney, intestine, heart, leg, lung, umbilical cord, liver, and brain of the Hokkaido native horse fetus, following the procedures described above. Each cDNA was synthesized using ReverTra Ace qPCR RT Master Mix (TOYOBO, Osaka, Japan), after which qRT-PCR was performed using a LightCycler480 (Roche Applied Science, Penzberg, Germany). The reaction mixtures were prepared by adding THUNDERBIRD SYBR qPCR Mix (TOYOBO) at a final concentration of 0.5 μM for each primer. Thermal cycling conditions consisted of one cycle at 95°C for 30 sec (denaturation), followed by 45 cycles at 95°C for 10 sec (denaturation), 61°C for 15 sec (primer annealing), and 72°C for 30 sec (extension). Relative mRNA abundance was calculated by the ΔΔCt method, with *GAPDH* as the reference gene. To improve reproducibility of the results, the examined sample from the same cDNA source was analyzed in triplicate for each PCR.

Statistical analysis

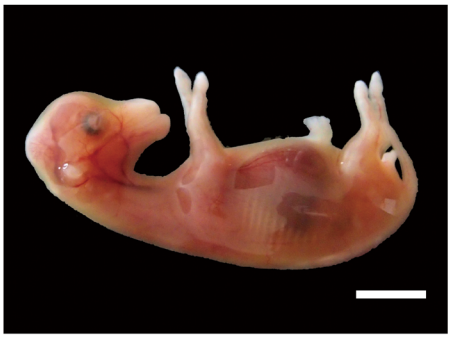
The expression levels of *IGF2* in the equine fetal tissues were compared to the means of the *IGF2* expression level in the yolk sac using Student's *t*-test. All analyses were performed with StatView software (Abacus Concepts, Berkeley, CA, USA), and a value of $P < 0.01$ was considered significant.

Results

Identification of a new transcript variant of the equine IGF2 mRNA

The *IGF2* mRNA sequence containing the UTRs was determined using RACE-PCR, and was deposited in the DNA Data Bank of Japan

A



B

cgctgttcggtttgcgatacgcagcagggaggtgcgcggcagccgcgccggcttccag
ataccaATGGGGATCCCGGTGGGGAAGTCGCTGCTGATGCTGTTCACTTTCTTGGCTT
TTGCCTCGTGCTGCATTGCTGCTTACCGCCCCAGTGAGACTCTGTGTGGCGGGGAGCT
GGTGGATACCCTCCAGTTTGTCTGTGGGGACCGCGGCTTTTACTTCAGCAGGCCCGCA
AGCCGCATCAACCGCCGCAGCCGTGGCATCGTGGAAGAGTGCTGTTTCCGCAGCTGCG
ACCTGGCCCTCCTGGAGACCTACTGTGCCACCCCCGCCAAGTCCGAGAGGGACGTGTC
GACCCCTCCGACCGTGCTTCCGGACGACTCCCCCAGATACCCCGTGGTCAAGCTCTTC
CAGTACAACGCCTGGAAGCAATCCACCCAGCGCCTGCGCAGGGGCCTGCCTGCCCTCC
TGCGTACCCGCCGGGGTCGCATGCTCGTCAAGGAGCTCGAAGCGTTCAGGGAGGCCCA
ACGTCACCGTCCCCTGATCGCCCTGCCCACCGAGGACCCCACTCCCCACGGGGCCGCC
TTTGTCGAGGTGTCCAGCGATCTGCAGTGAgccaaattgtcgtatttttgcaaaatgg
cactttccaccttgtgccctcttcttgaccagggccccttccaccgggtccacctctg
aaatctctctgtaccgtcccccagggcgggcctccccccccagcccccttgccccaac
ctcctcatgtcaggcacttcggccctgtccatcgggctgaaggaatcatagtaacatc
atc

C

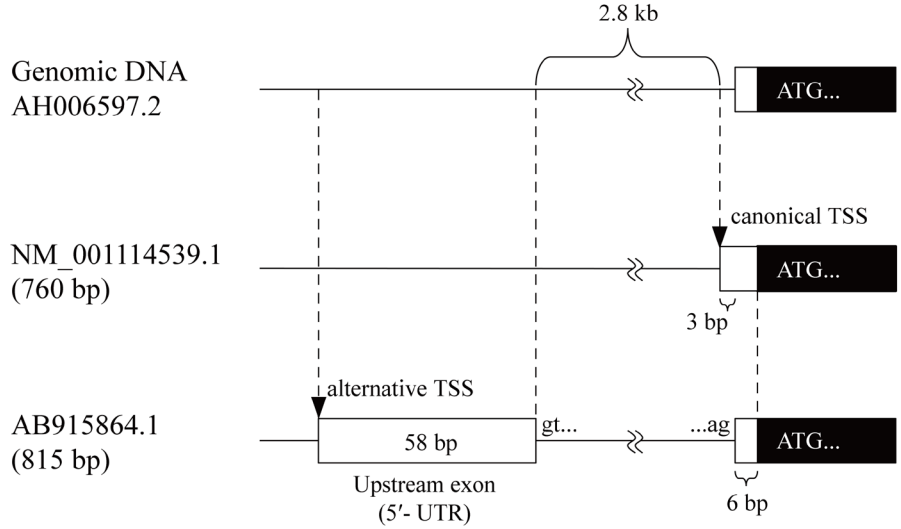


Fig. 1. Determination of the full-length equine *IGF2* mRNA sequence. A: the Hokkaido native horse fetus used in the analyses. The possible embryonic day of conceptus was evaluated as day 56, based on the fetal crown-rump-length. Bar = 10 mm. B: The full-length *IGF2* mRNA was 815 bp, and contained a 5′-untranslated region (UTR) of 64 bp, a coding sequence (CDS) of 546 bp, and a 3′-UTR of 205 bp. The UTR sequences are shown in lower-case letters and red font; uppercase letters indicate the CDS. C: Structural comparison of the identified mRNA with sequences available in public database repositories. The identified complete *IGF2* mRNA sequence had an additional exon located ~2.8 kb upstream from the canonical transcription start site (TSS). An alternative TSS determined by 5′ RACE was also located in the upstream UTR. The upstream 5′-UTR consisted of 58 bp, and did not contain the translation start site (ATG). White and black boxes indicate intron and exon region of *IGF2*, respectively.

Table 1. Primer sets used in this study

Analysis	Name	Primer sequence (5'-3')
RACE	GeneRacer 5′	CGACTGGAGCACGAGGACACTGA
	GeneRacer 3′	GCTGTCAACGATACGCTACGTAACG
	GeneRacer 5′ Nested	GGACACTGACATGGACTGAAGGAGTA
	GeneRacer 3′ Nested	CGCTACGTAACGGCATGACAGTG
CDS sequencing	eq*IGF2* CDS F	TGTTCGGTTTGCGATACGCAG
	eq*IGF2* CDS R	GACATGAGGAGGTTGGGGCAA
qPCR	eq*IGF2*-103bp-F	GGATACCCTCCAGTTTGTCTG
	eq*IGF2*-103bp-R	GAAACAGCACTCTTCCACGAT
	GAPDH 135bp F	CGACCACTTTGTCAAG
	GAPDH 135bp R	TCCTTCTCTTGCTGGGTGAT

A

Nucleotide position	398
(Amino acid position)	(133)
Database	T
(NM_001114539.1)	(Leu)
Hokkaido native horse examined in this study	T
	(Leu)
Thoroughbred_1	C
	(Pro)
Thoroughbred_2	T
	(Leu)

B

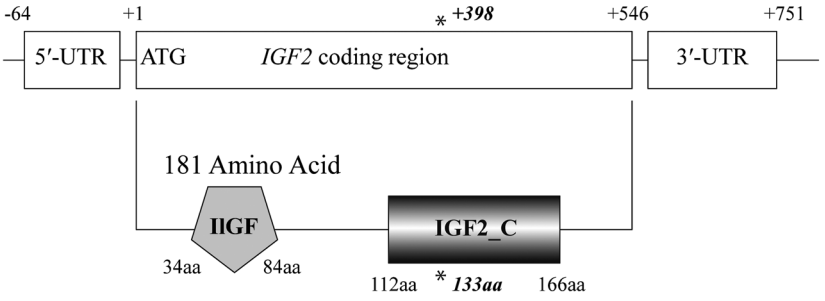


Fig. 2. Polymorphism analysis using *IGF2* sequences from three equine fetuses and the protein structure of equine IGF2. A: Nucleotide and amino acid substitution of equine *IGF2* determined in this study. Nucleotide position indicates the number from the transcription start site. The sequences corresponding to the protein domains are aligned under the horse sequence. B: A schematic showing the nucleotide sequence (upper) and protein domains (lower), IlGF and IGF2_C, predicted from the Pfam and SMART databases. Asterisks represent the nucleotide and amino acid substitution site in the equine *IGF2* gene, as shown in A.

(DDBJ) and National Center of Biotechnology Information (NCBI) (GenBank ID: AB915864). The complete *IGF2* mRNA was 815 bp, with a 5′-UTR of 64 bp, a CDS of 546 bp, and a 3′-UTR of 205 bp. The 5′- and 3′-UTR sequences of the *IGF2* mRNA are shown in Fig. 1B. In addition, the CDS of the equine *IGF2* was also identified. The *IGF2* mRNA sequence was 55 bp longer than the 5′-UTR sequence deposited previously (GenBank ID: NM_001114539.1). Compared to the equine genomic sequence (GenBank ID: AH006597), the 58 bp sequence identified in the present study was located approximately 2.8 kbp upstream of the canonical transcriptional start site (TSS) (Fig. 1C). Thus, as the nucleotide sequence beginning with "GT" and ending in "AG" is generally regarded as an intron, the 58 bp sequence identified here was derived from the exon coding *IGF2* mRNA differently than in the sequence deposited previously. Furthermore, the determined complete equine *IGF2* mRNA sequence was compared to the RNA-seq datasets obtained from Thoroughbred fetal tissue, spermatozoa, testes, and placental villi. Consequently, almost every short read was unmatched in the upstream exon, suggesting that the equine *IGF2* mRNA sequence identified in the present study represents a novel alternative TSS in horses.

Comparison of the IGF2 nucleotide and predicted amino acid sequences

The CDSs of the equine *IGF2* were compared with the reference sequence (Thoroughbred, GenBank ID: NM_001114539.1), as well as that of Hokkaido native horse and two Thoroughbred breeds (Fig. 2A). The CDS from the Hokkaido native horse examined in this study corresponded perfectly with the reference sequence and that of one of the Thoroughbred breeds (Thoroughbred_2), whereas sequencing revealed a mutation (T −> C transition) at position 398 in the other Thoroughbred breed (Thoroughbred_1). This single nucleotide polymorphism (SNP) in the CDS caused

an amino acid substitution (Leu −> Pro) at position 133 (133 aa) of the IGF2 amino acid sequence. Analysis of the IGF2 sequence using two public collections of the conserved domain models Pfam and SMART revealed that the equine IGF2 protein possessed two protein domains, IlGF (CDD ID: cd04368) and IGF2_C (CDD ID: pfam08365), indicating that the detected SNP was present in the IGF2_C domain (Fig. 2B).

Next, the complete equine *IGF2* mRNA sequences (from *Equus caballus*) was compared with those of humans (*Homo sapiens*), mice (*Mus musculus*), rats (*Rattus norvegicus*), cattle (*Bos taurus*), and pigs (*Sus scrofa*). Comparisons of the complete mRNA sequences (Supplementary Table 1: online only) revealed that the equine *IGF2* was most similar to that of *Sus scrofa* (85.5%) and least similar to that of *Bos taurus* (82.0%). However, when only CDS were compared, the equine *IGF2* was determined to be most similar to that of *Homo sapiens* (90.3%) and least similar to that of *Mus musculus* (84.7%). Comparison of the individual domains IlGF and IGF2_C revealed the number of species-specific amino acid variations in the IlGF domain to be markedly smaller than in the IGF2_C domain (Fig. 3A). Furthermore, comparisons of the predicted IGF2 amino acid sequences among the different species showed that the equine IGF2 amino acid sequence (from *Equus caballus*) was more similar to the sequences of the livestock *Sus scrofa* and *Bos taurus*, and *Homo sapiens*, than to those of the rodents *Mus musculus* and *Rattus norvegicus* (Fig. 3B).

Expression dynamics of IGF2 mRNA among the fetal tissues

To explore *IGF2* expression dynamics during fetal development in horses, we conducted quantitative gene expression analyses in 10 different tissues, consisting of the yolk sac, neck, kidney, intestine, heart, leg, lung, umbilical cord, liver, and brain, which were collected from a Hokkaido native horse fetus (Fig. 4). A criterion for *IGF2* expression level was interpreted based on the level of expression in the yolk sac, as most imprinted genes are expressed and play critical roles in extraembryonic tissues [25–27]. Of the 10 tissues that were examined, expression levels of *IGF2* were found to be significantly higher ($P < 0.01$) in the neck, leg, umbilical cord, and liver than in the yolk sac, whereas expression levels in the kidney, intestine, heart, and brain were significantly lower than those in the yolk sac ($P < 0.01$) (Fig. 4).

Discussion

In this study, we first identified the complete equine *IGF2* mRNA sequences, which harbor longer 5′-UTR than did the previously annotated sequence. The newly identified 5′-UTR was predicted to be located in a different exon, composed of 58 nucleotides. It has been reported that equine *IGF2* has at least three tissue-specific promoters driving the different transcripts, depending on the respective TSS [14]. The sequence of the *IGF2* transcript identified here did not correspond with the sequences annotated previously. In humans and mice, a variety of *IGF2* transcripts with differing 5′-UTRs have been identified in a tissue-specific manner [28, 29]. We identified a single variant from the yolk sac, but other variants may yet to be discovered in the embryonic tissues.

Generally, IGF2 possesses two major domains, those of IlGF and IGF2_C. Based on the predicted amino acid sequence, we determined that these domains were also present in equine IGF2, with the IlGF domain formed by the amino acids from position 30 to 84, and the IGF2_C domain formed by the amino acids from position 112 to 166 (Fig. 2B). The IlGF domain is common in the family of proteins that includes insulin, relaxin, and IGFs, which represents evolutionarily related active peptides (Pfam). As shown in Fig. 3A, the species-specific amino acid variations in the IlGF domain were smaller than those in the IGF2_C domain, indicating that the amino acid sequence in the IlGF domain was well-conserved among the different species; as such, we would expect that the equine IGF2 possesses homologous physiological functions. During fetal development, IGF2 plays critical roles in fetal and placental growth, in the exchange of material between the fetus and the mother via the placenta, and in the regulation of hematopoiesis in the liver [30–32]. However, these critical physiological functions have been elucidated in only a limited number of species, such as mice, pigs, and humans; for horses, only one study has shown that *IGF2* expression increased in the preovulatory follicles of adult mares [33].

To the best of our knowledge, our study represents the first attempt at determining *IGF2* mRNA expression levels in multiple fetal tissues of horses. Expression of *IGF2* in liver, neck, umbilical cord, and leg tissues was significantly greater than that in other tissues, which was consistent with the expression pattern observed in mice fetal tissues [34]. As with mice, prenatal *IGF2* expression was considerably greater than postnatal expression in skeletal muscle and liver tissues of pigs and cattle [35–37]. Such results may reflect the indispensability of IGF2 expression for fetal viability and normal body growth and development, given the vital roles IGF2 plays in fetal and hematopoietic stem cell development [30, 38]. The transcription level of *Igf2* during fetal development in mice was shown to affect mRNA expression levels of various angiogenic factors, including *Vegf*, *Flt1*, *Flt4*, *Flk1*, *Ang1*, *Ang2*, *Tie1*, and *Tie2* [39]. Thus, along with structural homology, the expression pattern in the fetal tissues indicated functional significance of *IGF2* in the fetal development of horses, as well as in other species. Although it is very difficult to collect sufficient horse fetus samples of the same embryonic days, analyses involving larger sample sizes would provide more precise and in-depth information about the dynamics of *IGF2* expression in various tissues, as in the case of mice.

Polymorphism was detected within the IGF2_C domain. Commonly, IGF2 is initially synthesized as a proIGF2 containing the IGF2_C domain. This precursor, proIGF2, undergoes post-translational processing by proprotein convertase enzymes, resulting in variants lacking the IGF2_C domain [40]. The cleavage sites of proprotein convertases on IGF2 are highly conserved among species [41], and were also determined in the present study (Fig. 3A), suggesting that similar IGF2 variants in which the IGF2_C domain is present or absent also exist in horses. Although the IGF2 variants, including proIGF2, have been shown to circulate in fetal and neonatal rats, their specific biological roles in fetal development are not entirely clear. [41]. To further explore the significance of the polymorphism identified here, it will be necessary to clarify the details of the variants present in the equine fetus and assess the physiological functions of proIGF2 containing the IGF2_C domain.

Genome-wide comparisons of horse and human chromosomes show

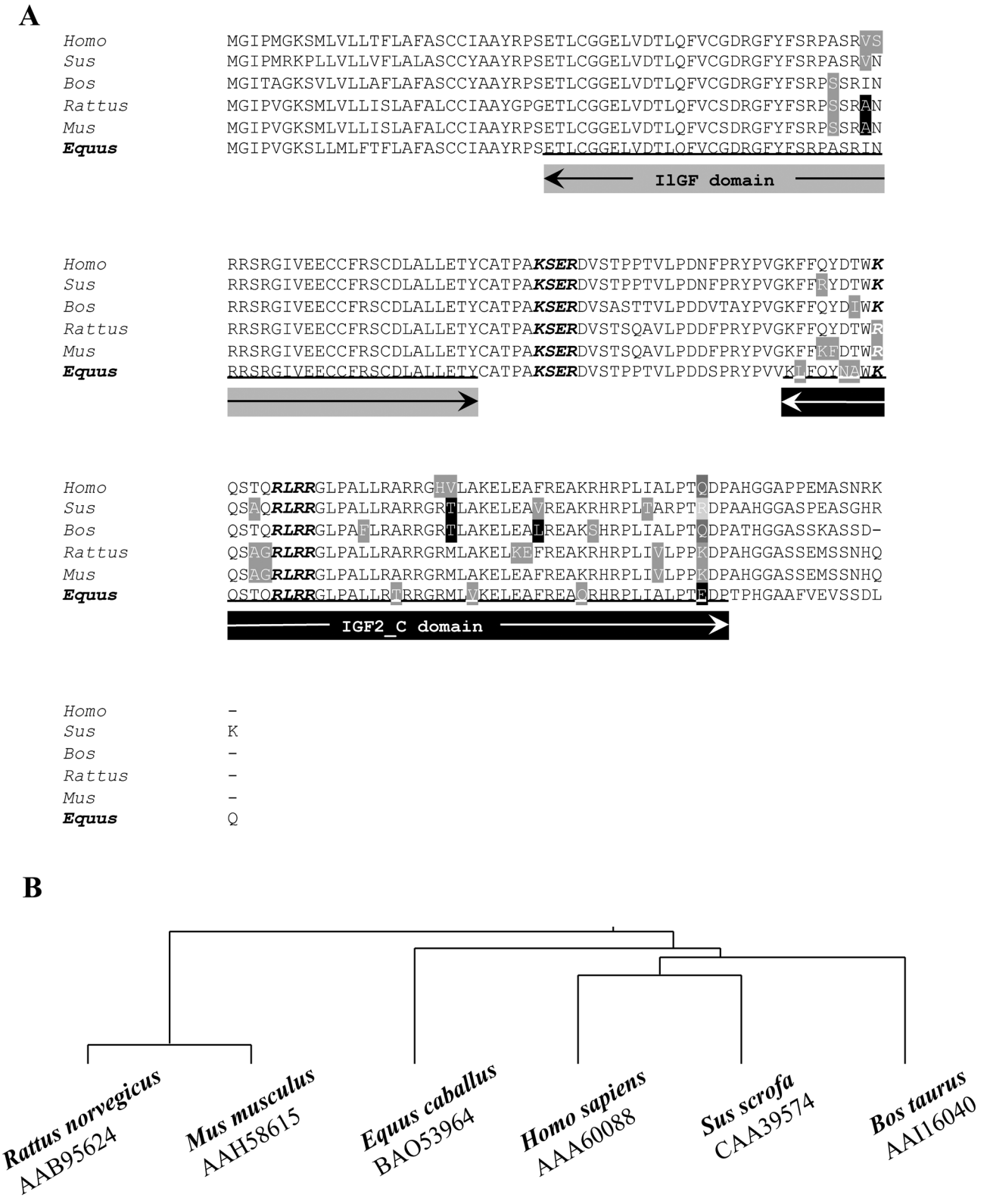


Fig. 3. Comparison of IGF2 amino acid sequences among different species. A: Predicted IGF2 amino acid sequences in *Homo sapiens* (*Homo*), *Sus scrofa* (*Sus*), *Bos taurus* (*Bos*), *Mus musculus* (*Mus*), *Rattus norvegicus* (*Rattus*), and *Equus caballus* (*Equus*). All sequences contained the same cleavage sites for proprotein convertases [41], which are shown in italic and bold font. B: Comparison of the IGF2 amino acid sequences in humans, mice, rats, cattle, and pigs using the ClustalW algorithm.

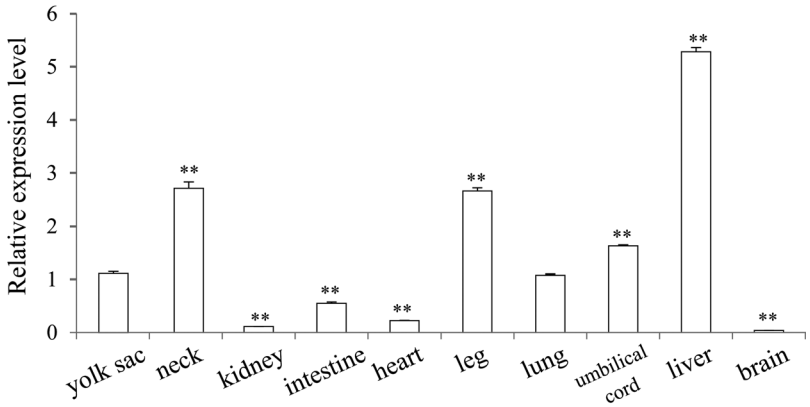


Fig. 4. Graphical representation of the equine *IGF2* mRNA expression in fetal tissues. Expression levels of equine *IGF2* mRNA in 10 different tissues (yolk sac, neck, kidney, intestine, heart, leg, lung, umbilical cord, liver, and brain) analyzed by qPCR. The values represent the levels of expression relative to that of the internal control gene (*GAPDH*) and are expressed as mean ± SEM. Asterisks indicate differences in expression levels relative to that in the yolk sac ($P < 0.01$).

strong synteny between these species, and thus equines can serve as model systems for numerous human ailments relating to infertility, inflammatory diseases, and muscle disorders [42, 43]. However, interspecies comparison of the predicted amino acid sequences of IGF2 revealed that the equine sequence was more dissimilar to the human sequence than it was to the sequences of other species (Fig. 3B). Likewise, comparison of nucleic acids containing the UTR sequences indicated that the equine sequence was not closely homologous with the human sequence (Supplementary Table 1). This might be due to lower rates of conservation in the UTR resulting from evolutionary constraints. The length of the 5′-UTR plays important roles in the regulation of *IGF2* at the translation level [44–47]. Moreover, the 3′-UTR sequence could also influence mRNA stability and translatability in the regulation of *IGF2* for controlling the protein synthesis in some instances via interactions with microRNAs [20, 48]. Although transcriptional and translational regulation of the equine *IGF2* gene has yet to be fully elucidated, knowledge of the genetic information contained in the UTR sequence might prove useful for gaining a better understanding of mRNA translatability of the equine *IGF2* gene.

In conclusion, the complete mRNA sequence of the equine *IGF2* containing the 5′- and 3′-UTRs was analyzed, and was identified as a new transcript variant. In addition, the predicted IGF2 protein possessed two major protein domains, IlGF2 and IGF2_C. Comparison of the equine IGF2 amino acid sequence with those of other species revealed that the equine *IGF2* gene was highly conserved among the species included in the analysis. Furthermore, the expression pattern of the *IGF2* mRNA in horse fetal tissues was determined for the first time, and was found to be consistent with the expression patterns in several other mammalian species. Taken together, our results suggested that the transcriptional and translational products of *IGF2* have similar functions in the fetal development of mammals, including horses.

Acknowledgement

Financial support for this research was provided in part by a Grant-in-Aid for Challenging Exploratory Research to MK (No. 16K14587) from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References

1. **Leighton PA, Ingram RS, Eggenschwiler J, Efstratiadis A, Tilghman SM.** Disruption of imprinting caused by deletion of the H19 gene region in mice. *Nature* 1995; **375**: 34–39. [Medline] [CrossRef]
2. **Ripoche MA, Kress C, Poirier F, Dandolo L.** Deletion of the H19 transcription unit reveals the existence of a putative imprinting control element. *Genes Dev* 1997; **11**: 1596–1604. [Medline] [CrossRef]
3. **Wang ZQ, Fung MR, Barlow DP, Wagner EF.** Regulation of embryonic growth and lysosomal targeting by the imprinted Igf2/Mpr gene. *Nature* 1994; **372**: 464–467. [Medline] [CrossRef]
4. **Jeon JT, Carlborg O, Törnsten A, Giuffra E, Amarger V, Chardon P, Andersson-Eklund L, Andersson K, Hansson I, Lundström K, Andersson L.** A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the IGF2 locus. *Nat Genet* 1999; **21**: 157–158. [Medline] [CrossRef]
5. **Nezer C, Moreau L, Brouwers B, Coppieters W, Detilleux J, Hanset R, Karim L, Kvasz A, Leroy P, Georges M.** An imprinted QTL with major effect on muscle mass and fat deposition maps to the IGF2 locus in pigs. *Nat Genet* 1999; **21**: 155–156. [Medline] [CrossRef]
6. **DeChiara TM, Robertson EJ, Efstratiadis A.** Parental imprinting of the mouse insulin-like growth factor II gene. *Cell* 1991; **64**: 849–859. [Medline] [CrossRef]
7. **Burgos C, Galve A, Moreno C, Altarriba J, Reina R, García C, López-Buesa P.** The effects of two alleles of IGF2 on fat content in pig carcasses and pork. *Meat Sci* 2012; **90**: 309–313. [Medline] [CrossRef]
8. **Huang YZ, Wang J, Zhan ZY, Cao XK, Sun YJ, Lan XY, Lei CZ, Zhang CL, Chen H.** Assessment of association between variants and haplotypes of the IGF2 gene in beef cattle. *Gene* 2013; **528**: 139–145. [Medline] [CrossRef]
9. **Clark DL, Clark DI, Beever JE, Dilger AC.** Increased prenatal IGF2 expression due to the porcine intron3-G3072A mutation may be responsible for increased muscle mass. *J Anim Sci* 2015; **93**: 2546–2558. [Medline] [CrossRef]
10. **DeChiara TM, Efstratiadis A, Robertson EJ.** A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. *Nature* 1990; **345**: 78–80. [Medline] [CrossRef]
11. **Giannoukakis N, Deal C, Paquette J, Goodyer CG, Polychronakos C.** Parental genomic imprinting of the human IGF2 gene. *Nat Genet* 1993; **4**: 98–101. [Medline] [CrossRef]
12. **Dindot SV, Farin PW, Farin CE, Romano J, Walker S, Long C, Piedrahita JA.** Epi-

genetic and genomic imprinting analysis in nuclear transfer derived Bos gaurus/Bos taurus hybrid fetuses. *Biol Reprod* 2004; **71**: 470–478. [Medline] [CrossRef]
13. **Tian XC.** Genomic imprinting in farm animals. *Annu Rev Anim Biosci* 2014; **2**: 23–40. [Medline] [CrossRef]
14. **Otte K, Choudhury D, Charalambous M, Engström W, Rozell B.** A conserved structural element in horse and mouse IGF2 genes binds a methylation sensitive factor. *Nucleic Acids Res* 1998; **26**: 1605–1612. [Medline] [CrossRef]
15. **Otte K, Engström W.** Insulin-like growth factor II in the horse: determination of a cDNA nucleotide sequence and expression in fetal and adult tissue. *Gen Comp Endocrinol* 1994; **96**: 270–275. [Medline] [CrossRef]
16. **Proudfoot NJ, Brownlee GG.** 3 non-coding region sequences in eukaryotic messenger RNA. *Nature* 1976; **263**: 211–214. [Medline] [CrossRef]
17. **Bugaut A, Balasubramanian S.** 5-UTR RNA G-quadruplexes: translation regulation and targeting. *Nucleic Acids Res* 2012; **40**: 4727–4741. [Medline] [CrossRef]
18. **Mayr C.** Regulation by 3'-untranslated regions. *Annu Rev Genet* 2017. doi: 10.1146/annurev-genet-120116-024704 [Medline] [CrossRef]
19. **Wong QW, Vaz C, Lee QY, Zhao TY, Luo R, Archer SK, Preiss T, Tanavde V, Vardy LA.** Embryonic stem cells exhibit mRNA isoform specific translational regulation. *PLoS ONE* 2016; **11**: e0143235. [Medline] [CrossRef]
20. **Saha S, Choudhury J, Ain R.** MicroRNA-1413p and miR-200a-3p regulate insulin-like growth factor 2 during mouse placental development. *Mol Cell Endocrinol* 2015; **414**: 186–193. [Medline] [CrossRef]
21. **Kaneda M, Takahashi M, Yamanaka KI, Saito K, Taniguchi M, Akagi S, Watanabe S, Nagai T.** Epigenetic analysis of bovine parthenogenetic embryonic fibroblasts. *J Reprod Dev* 2017; **63**: 365–375. [Medline] [CrossRef]
22. **Bateman A, Birney E, Cerruti L, Durbin R, Etwiller L, Eddy SR, Griffiths-Jones S, Howe KL, Marshall M, Sonnhammer EL.** The Pfam protein families database. *Nucleic Acids Res* 2002; **30**: 276–280. [Medline] [CrossRef]
23. **Letunic I, Goodstadt L, Dickens NJ, Doerks T, Schultz J, Mott R, Ciccarelli F, Copley RR, Ponting CP, Bork P.** Recent improvements to the SMART domain-based sequence annotation resource. *Nucleic Acids Res* 2002; **30**: 242–244. [Medline] [CrossRef]
24. **Thompson JD, Higgins DG, Gibson TJ.** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994; **22**: 4673–4680. [Medline] [CrossRef]
25. **Burton GJ, Fowden AL.** Review: The placenta and developmental programming: balancing fetal nutrient demands with maternal resource allocation. *Placenta* 2012; **33**(Suppl): S23–S27. [Medline] [CrossRef]
26. **Kaneko-Ishino T, Ishino F.** Retrotransposon silencing by DNA methylation contributed to the evolution of placentation and genomic imprinting in mammals. *Dev Growth Differ* 2010; **52**: 533–543. [Medline] [CrossRef]
27. **Okae H, Hiura H, Nishida Y, Funayama R, Tanaka S, Chiba H, Yaegashi N, Nakayama K, Sasaki H, Arima T.** Re-investigation and RNA sequencing-based identification of genes with placenta-specific imprinted expression. *Hum Mol Genet* 2012; **21**: 548–558. [Medline] [CrossRef]
28. **Daimon M, Johnson TR, Ilan J, Ilan J.** The third IGF-II promoter specifies transcription of three transcripts out of five in human placenta. *Mol Reprod Dev* 1992; **33**: 413–417. [Medline] [CrossRef]
29. **Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Usdin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalon DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Ketteman M, Madan A, Rodrigues S, Sanchez A, Whiting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U, Smailus DE, Schnerch A, Schein JE, Jones SJ, Marra MA. Mammalian Gene Collection Program Team.** Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci USA* 2002; **99**: 16899–16903. [Medline] [CrossRef]
30. **Wu Q, Kawahara M, Kono T.** Synergistic role of Igf2 and Dlk1 in fetal liver development and hematopoiesis in bi-maternal mice. *J Reprod Dev* 2008; **54**: 177–182. [Medline] [CrossRef]
31. **Kumaravelu P, Hook L, Morrison AM, Ure J, Zhao S, Zuyev S, Ansell J, Medvinsky A.** Quantitative developmental anatomy of definitive haematopoietic stem cells/long-term repopulating units (HSC/RUs): role of the aorta-gonad-mesonephros (AGM) region and the yolk sac in colonisation of the mouse embryonic liver. *Development* 2002; **129**: 4891–4899. [Medline]
32. **Constância M, Hemberger M, Hughes J, Dean W, Ferguson-Smith A, Fundele R, Stewart F, Kelsey G, Fowden A, Sibley C, Reik W.** Placental-specific IGF-II is a major modulator of placental and fetal growth. *Nature* 2002; **417**: 945–948. [Medline] [CrossRef]
33. **Watson ED, Bae SE, Thomassen R, Thomson SR, Woad K, Armstrong DG.** Insulin-like growth factors-I and -II and insulin-like growth factor-binding protein-2 in dominant equine follicles during spring transition and the ovulatory season. *Reproduction* 2004; **128**: 321–329. [Medline] [CrossRef]
34. **Kawahara M, Wu Q, Takahashi N, Morita S, Yamada K, Ito M, Ferguson-Smith AC, Kono T.** High-frequency generation of viable mice from engineered bi-maternal embryos. *Nat Biotechnol* 2007; **25**: 1045–1050. [Medline] [CrossRef]
35. **Blondin P, Farin PW, Crosier AE, Alexander JE, Farin CE.** In vitro production of embryos alters levels of insulin-like growth factor-II messenger ribonucleic acid in bovine fetuses 63 days after transfer. *Biol Reprod* 2000; **62**: 384–389. [Medline] [CrossRef]
36. **Farin CE, Alexander JE, Farin PW.** Expression of messenger RNAs for insulin-like growth factors and their receptors in bovine fetuses at early gestation from embryos produced in vivo or in vitro. *Theriogenology* 2010; **74**: 1288–1295. [Medline] [CrossRef]
37. **Van Laere AS, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, Archibald AL, Haley CS, Buys N, Tally M, Andersson G, Georges M, Andersson L.** A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* 2003; **425**: 832–836. [Medline] [CrossRef]
38. **Zhang CC, Lodish HF.** Insulin-like growth factor 2 expressed in a novel fetal liver cell population is a growth factor for hematopoietic stem cells. *Blood* 2004; **103**: 2513–2521. [Medline] [CrossRef]
39. **Kawahara M, Wu Q, Kono T.** Involvement of insulin-like growth factor 2 in angiogenic factor transcription in Bi-maternal mouse conceptuses. *J Reprod Dev* 2010; **56**: 79–85. [Medline] [CrossRef]
40. **Qiu Q, Basak A, Mbikay M, Tsang BK, Gruslin A.** Role of pro-IGF-II processing by proprotein convertase 4 in human placental development. *Proc Natl Acad Sci USA* 2005; **102**: 11047–11052. [Medline] [CrossRef]
41. **Qiu Q, Jiang JY, Bell M, Tsang BK, Gruslin A.** Activation of endoproteolytic processing of insulin-like growth factor-II in fetal, early postnatal, and pregnant rats and persistence of circulating levels in postnatal life. *Endocrinology* 2007; **148**: 4803–4811. [Medline] [CrossRef]
42. **Chowdhary BP, Paria N, Raudsepp T.** Potential applications of equine genomics in dissecting diseases and fertility. *Anim Reprod Sci* 2008; **107**: 208–218. [Medline] [CrossRef]
43. **Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, Imsland F, Lear TL, Adelson DL, Bailey E, Bellone RR, Blöcker H, Distl O, Edgar RC, Garber M, Leeb T, Mauceli E, MacLeod JN, Penedo MC, Raison JM, Sharpe T, Vogel J, Andersson L, Antczak DF, Biagi T, Binns MM, Chowdhary BP, Coleman SJ, Della Valle G, Fryc S, Guérin G, Hasegawa T, Hill EW, Jurka J, Kiialainen A, Lindgren G, Liu J, Magnani E, Mickelson JR, Murray J, Nergadze SG, Onofrio R, Pedroni S, Piras MF, Raudsepp T, Rocchi M, Røed KH, Ryder OA, Searle S, Skow L, Swinburne JE, Syvänen AC, Tozaki T, Valberg SJ, Vaudin M, White JR, Zody MC, Lander ES, Lindblad-Toh K. Broad Institute Genome Sequencing Platform Broad Institute Whole Genome Assembly Team.** Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science* 2009; **326**: 865–867. [Medline] [CrossRef]
44. **de Pagter-Holthuizen P, Jansen M, van der Kammen RA, van Schaik FM, Sussenbach JS.** Differential expression of the human insulin-like growth factor II gene. Characterization of the IGF-II mRNAs and an mRNA encoding a putative IGF-II-associated protein. *Biochim Biophys Acta* 1988; **950**: 282–295. [Medline] [CrossRef]
45. **Nielsen FC, Gammeltoft S, Christiansen J.** Translational discrimination of mRNAs coding for human insulin-like growth factor II. *J Biol Chem* 1990; **265**: 13431–13434. [Medline]
46. **Nielsen FC, Christiansen J.** Endonucleolysis in the turnover of insulin-like growth factor II mRNA. *J Biol Chem* 1992; **267**: 19404–19411. [Medline]
47. **Meinsma D, Holthuizen PE, Van den Brande JL, Sussenbach JS.** Specific endonucleolytic cleavage of IGF-II mRNAs. *Biochem Biophys Res Commun* 1991; **179**: 1509–1516. [Medline] [CrossRef]
48. **LeRoith D, Roberts CT Jr.** Insulin-like growth factors. *Ann NY Acad Sci* 1993; **692**: 1–9. [Medline] [CrossRef]