SINE Insertion in 5’ Flanking of IGFBP3 May Associated With Gene Expression and Phenotypic Variations

Xiaoyan Wang
Yangzhou University https://orcid.org/0000-0001-8521-2974
Yalong An
Yangzhou University
Eduard Murani
Leibniz Institute for Farm Animal Biology: Leibniz-Institut fur Nutztierbiologie
Enrico D'alessandro
University of Messina
Chengling Chi
Yangzhou University
Cai Chen
Yangzhou University
Ali Shoaib Moawad
Yangzhou University
Kui Li
Chinese Academy of Agricultural Sciences
Klaus Wimmers
Leibniz Institute for Farm Animal Biology: Leibniz-Institut fur Nutztierbiologie
Chengyi Song (cysong@yzu.edu.cn)
Yangzhou University https://orcid.org/0000-0002-0488-4718

Research

Keywords: Pig, IGFBPs gene, RIPs, Economic traits, Enhancer

Posted Date: July 13th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-693177/v1

Abstract

Background: Insulin-like growth factor binding proteins (*IGFBPs*), specifically binding to *IGF1* and *IGF2*, play an important role in regulating physiological functions of insulin-like growth factors (*IGFs*). *IGFBPs* have been considered important candidate genes for economic traits due to their involvement in physiological processes related to growth and development. However, most of the current studies on genetic markers of *IGFBPs* have focused on SNPs, and large fragment insertion mutations such as retrotransposons have rarely been considered. In this paper, we screened the porcine *IGFBP* genes (*IGFBP1*-8) for retrotransposon insertion polymorphisms (RIPs) using bioinformatics prediction combined with the PCR-based amplification. Furthermore, for two linked RIPs their population distribution and impact on promoter activity and phenotype were further evaluated.

Results: Screening of *IGFBP*s identified RIPs in *IGFBP1-5* and *IGFBP7*. In total twelve predicted RIPs were confirmed by PCR. These RIPs were detected in different breeds with an uneven distribution among them. By linkage genetic analysis and PCR verification, *IGFBP3*-1-RIP and *IGFBP3*-2-RIP are completely linked, showing only three genotypes, $SINE^{+/+}/LINE^{-/-}$, $SINE^{-/-}/LINE^{+/+}$ and $SINE^{+/-}/LINE^{-/+}$. The age of 100 kg body weight and longissimus muscle thickness of Large white individuals of $SINE^{+/+}/LINE^{-/-}$genotype were significantly (P<0.05) higher than those of $SINE^{+/-}/LINE^{+/-}$ genotype and $SINE^{-/-}/LINE^{+/+}$ genotype. However, the longissimus muscle thickness and corrected backfat thickness of $SINE^{+/+}/LINE^{-/-}$ individuals were significantly (P<0.05) thinner than those of $SINE^{+/+}/LINE^{-/-}$ genotype. The expression of *the IGFBP3* gene in liver, leg muscles and backfat of 30-day Sujiang piglets with $SINE^{+/+}/LINE^{-/-}$ genotype were significantly higher (P<0.05) than those with $SINE^{-/-}/LINE^{+/+}$ genotype by the real-time quantitative polymerase chain reaction (qPCR). Further study was conducted to confirm the effect of SINE and LINE insertion on the promoter activity of *IGFBP3*. First, the core promoter region of the *IGFBP3* gene was identified locating within 482bp upstream of ATG by using the dual-luciferase activity assay. Then SINE and LINE were combined with 482bp fragment to construct a recombinant vector respectively based on the PGL3-Promoter-Enhancer. The recombinant vector was transfected into C2C12, 3T3-L1, and Hela cells. The detection of the dual-luciferase reporter gene revealed that only SINE insertion was significantly increased *(P<0.05*) promoter activity of the *IGFBP3* gene, indicating that the SINE may act as an enhancer to regulate the promoter activity of the *IGFBP3* gene.

Conclusions: Overall, this study identified 12 RIPs in *IGFBP* gene clusters, which provided useful markers for genetic analysis of pig populations. Furthermore, based on the dual-luciferase activity assay in cells and association analysis, the linked genetic variations generated by SINE and LINE insertions in 5' flanking of *IGFBP3* may associated with variations of gene expression and phenotype.

Introduction

The insulin-like growth factor (IGF) system, which comprises ligands of IGF-I/II, IGF receptors (IGFR), IGF binding proteins (IGFBPs), and IGFBP hydrolases, is widely involved in cell proliferation and differentiation through regulating DNA transcription. It plays multiple roles in the regulation of fetal and

postnatal growth and development, and dysregulation may lead to disease [1]. *IGFBPs* comprise a family of at least six proteins that specifically bind to *IGF1* and *IGF2* with higher affinity compared to *IGF* receptors [2]. More than acting as simple carrier proteins, IGFBPs in serum could regulate the endocrine actions of IGFs by regulating the amount of IGF available to bind to IGF-I receptors. Furthermore, IGFBPs function independently of the IGFs in vitro and in vivo as growth modulators [3]. In humans, genetic variation in *IGFBPs* was associated with height by GWAS [4, 5]. For the domestic animals, polymorphisms in *IGFBPs* were associated with meat quality [6, 7], energy balance [8], and other economically-relevant traits [9, 10]. However, most of the current research on molecular markers of *IGFBPs* has focused on SNPs, and large fragment insertion mutations such as retrotransposons have rarely been considered.

Retrotransposon can freely jump, copy or translocate in the genome, which not only creates rich genetic diversity within the species, but also plays an important role in functional genomic variation and gene evolution, and may affect gene expression[11, 12]. Retrotransposons are the dominant transposons in mammalian genomes. They can be classified into ERV (Endogenous Retrovirus, ERV), LINE (Long Interspersed Nuclear Elements), and SINE (Short Interspersed Nuclear Elements) [13]. New molecular markers (retrotransposon insertion polymorphisms, RIPs) methods based on retrotransposon have become important tools for studying the genome evolution, population genetics, and molecular breeding [14, 15, 16] .

Due to the important functional and regulatory diversity of *the IGFBP* family, RIPs in porcine *IGFBP* genes, the breed distribution of these RIPs, and association with pig economic traits were investigated in this study. In addition, the genetic effect of two linkage RIPs in the 5’ flanking region of *IGFBP3* was further evaluated. These findings will help to understand the role of RIPs in the pig genomic and genetic variation, and the *IGFBP3* RIPs may be useful for growth selection in pig breeding.

Material And Methods

Animals and DNA isolation

Pool DNA of 12 breeds, including Duroc, Landrace, Large white, Sujiang, Diannan small ear, Erhualian, Wuzhishan, Bama, Tibetan, Meishan, Fengjing and wild boar pigs, were selected for RIPs identification, 6 individual for each breed. And 24 individuals of 9 breeds were selected including Duroc, Landrace, Largewhite, Sujiang, Meishan, Erhualian, Fengjing, Jiangquhai and Tibetan for population genetic diversity analysis. The origins of these breeds were shown in TableS1. In these breeds, Duroc, Landrace and Yorkshire are three lean type commercial breeds, Meishan, Erhualian, Fengjing, and Jiangquhai are four fat type indigenous breeds, and Tibetan is a miniature indigenous breed. Sujiang is a hybrid breed with 62.5% Duroc, 18.75% Jiangquhai, and 18.75% Fengjing blood. Totally, fifteen 30-day old Sujiang piglets were used for characterize the tissue expression. 863 Large white individuals with growth performance records were used for correlation analysis. TIANamp Genomic DNA Kit (TIANGEN Biotech Co. Ltd. Beijing, China) were used for DNA isolation from the samples of each individuals. The quality and concentration of DNA was detected using NanoPhotometer (Implen, Germany).

RIP screen

Eight *IGFBP* genes (*IGFBP1-7 and IGFBPL1*) including 5’ (5 kb) and 3’ (3 kb) flanking regions, respectively, were obtained from one reference genome (Duroc) and fifteen assembled non-reference genomes (Landrace, Yorkshire, Pietrain, Berkshire, Hampshire, Cross-breed of Yorkshire/Landrace/Duroc, Wuzhishan, Tibetan, Rongchang, Meishan, Bamei, Bama, Jinhua, Goettingen, and Ellegaard Gottingen minipigs) deposited in the NCBI database (https://www.ncbi.nlm.nih.gov/). The information of these genomic sequences was listed in Table S1. The retrieved sequences were aligned using the ClustalX program to screen for large structural variations (more than 50 bp). Retrotransposon (SINE, LINE, and ERV) insertions were annotated by using the program of RepeatMasker (http://www.repeatmasker.org/) with a customer library constructed previously (ref: mobile dna). Promoters were predicted by multiple online programs, including Promoter 2.0 (http://www.cbs.dtu.dk/services/Promoter/), EPD (https://epd.epfl.ch//index.php), and BDGP (https://www.fruitfly.org/seq_tools/promoter.html). The predicted large structural variations (more than 50 bp) overlapping with retrotransposon (SINE, LINE, and ERV) insertions were designated as RIPs. These RIPs were further evaluated in DNA pool by PCR amplification. For each breed, two pooled DNA samples, each containing three animal samples, were used. The primers used for RIP evaluation were listed in Table S2. All obtained RIPs were further verified by TA cloning following the manufacturer's instructions (Tiangen, Beijing, China) and sequenced in Tsingke Biotechnology Co., Ltd.

RIP Genotyping

After gel electrophoresis of PCR products using designed primers, three genotypes were identified including a short single band of homozygous type without retrotransposon insertion, a large single band of homozygous type with retrotransposon insertion, and a double bands with short one and large one of heterozygote type.

Expression analysis

30-day old Sujiang piglets were genotyped and 8 individuals of $SINE^{+/+}/LINE^{-/-}$ genotype, 4 individuals of $SINE^{+/-}$ and $LINE^{-/+}$genotype, and 3 individuals for $SINE^{-/-}/LINE^{+/+}$ genotype were selected. Following slaughter, tissue samples including liver, leg muscles, longissimus, and backfat were collected. Then total RNA were extracted using Trizol (Tiangen, Beijing, China) and cDNA was prepared according to the manufacturer’s protocol by using TAKARA Kit (Takara, Tokyo, Japan). The mRNA expression of *IGFBP3* was evaluated by quantitative real-time PCR (qPCR) using the 7900HT Fast Real-Time PCR System (Applied Biosystems, New York, American) in a total volume of 20 μl containing SYBR mix (10 μl), primers (4 ng), and cDNA sample (50 ng) according to the manufacturer’s instruction (Takara, Tokyo, Japan) (primers were listed in Table S2). Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) was used as an endogenous control. to normalize the target gene expression in four different tissues. Gene expression was quantified using the $2^{-\Delta\Delta Ct}$ method. PCR products were run on 1.5% ethidium bromide-stained agarose gels and confirmed using melting curve analyses to assess PCR product identity.

Dual-luciferase reporter assay

According to the predicted promoter and evolutionarily conserved sequences among different species combined with mVISTA (http://genome.lbl.gov/vista/mvista/submit.shtml) in the 5' flanking region of *IGFBP3*, 482bp-, 1452bp- and 3212bp-long fragments upstream from ATG of *IGFBP3* respectively were cloned from Sujiang genomic DNA and sequenced (primers were listed in Table S2). These fragments were used to construct the vectors designated pGL3-482-basic, pGL3-1452-basic, and pGL3-3212-basic. A total of 2×10^5 PK15 cells were plated on a 6-well plate. At 70–80% confluency, the cells were transfected with either pGL3-482-basic, pGL3-1452-basic, or pGL3-3212-basic vectors using Lipofectamine 3000 reagent (Invitrogen, Carlsbad, American), parallel to pGL3-basic and pGL3-cont serving as negative and positive control, respectively (Promega, Madison, American). After 48 hours, the cells were collected and luciferase activity measurement using the dual-luciferase reporter system (Promega, Madison, American) was used to determine the core promoter region. Subsequently, amplify the retrotransposon insertion, retrotransposon-free and core promoter fragments respectively, and combine the transposon-inserted and non-inserted fragments with the core promoter through the second round of overlapping PCR. and the combined sequences were cloned into the pGL3-basic vector to construct $SINE^+$-Promoter-Luc (En), SINE–Promoter-Luc (En), $LINE^+$-Promoter-Luc (En), and LINE–Promoter-Luc (En) vectors to verify the effect of transposon fragments on the activity of *IGFBP3* core promoter (primers were listed in Table S2). A total of 2×10^5 C2C12, 3T3-L1, and Hela cells were plated on a 6-well plate. Transfection and dual luciferase activity measurement was performed as described above.

Statistical analysis

The genotype and the allele frequencies were calculated, and Hardy-Weinberg equilibrium were tested using chi-square test Popgene X^2 [17]. Polymorphism information content (PIC) was calculated according to the formula presented by [18]. SHEsis (http://analysis.bio-x.cn/myAnalysis.php) was used to analyze locus linkage[19]. Finally, for the production performance correlation of 863 Large white pigs of different genotypes, and All data were expressed as mean ± SD. The expression of different genotypes of Sujiang pigs and the results of all cell transfection experiments were processed by SPSS25.0 (SPSS, Chicago, USA). Tukey's post-hoc test was used for one-way analysis of variance. T-test with a paired two-tailed distribution and ANOVA were used to analyze the associations between genotype and phenotype of 863 Large white individuals.

Animal welfare

All treatments and protocols involving animals in this study were strictly done in accordance with the guidelines of the Animal Experiment Ethics Committee of Yangzhou University (approval number: YZUDWSY2018-12).

Results

Twelve RIPs generated by retrotransposon insertions in the pig *IGFBPs* gene cluster

Eight *IGFBP* genes and their flanking sequences from sixteen assembled pig genomes, representing lean type pigs (Cross-breed of Yorkshire/Landrace/Duroc, Duroc, Landrace, Yorkshire, Pietrain, Berkshire, and Hampshire), fat type pigs (Rongchang, Meishan, Bamei, and Jinhua), and miniature pigs (Bama, Wuzhishan, Tibetan, Goettingen, and Ellegaard Gottingen) were used to screen for structural variation by alignment with ClustalX program [20]. We identified 130 large SVs (Structural Variations, more than 50 bp long), and 49 of them were predicted as RIPs, including 29 SINE, 6 ERV, and 14 LINE RIPs, which were summarized in Additional file 1 (Table S3). Subsequently, all these predicted RIPs were further experimentally validated by PCR amplification and fragment analysis using specific primer pairs spanning the insertions in pooled DNA samples. Twelve RIPs, including seven SINE RIPs, two ERV RIPs, and three Line RIPs, were found segregating in 11 domesticated pig breeds and wild boar samples (Fig. 1A). All these RIPs were further confirmed by TA cloning and sequencing. Details of these RIPs, including retrotransposon types according to our previous retrotransposon annotation in pig [20], direction, insertion loci, and sequence length were shown in Fig. 1B and 1C. We named those insertions as *IGFBP1*-1-RIP, *IGFBP1*-2-RIP, *IGFBP1*-3-RIP, *IGFBP2*-1-RIP, *IGFBP3*-1-RIP, *IGFBP3*-2-RIP, *IGFBP3*-3-RIP, *IGFBP3*-4-RIP, *IGFBP4*-1-RIP, *IGFBP5*-1-RIP, *IGFBP7*-1-RIP, *IGFBP7*-2-RIP. Three RIPs, 165 bp L1D14, 327 bp ERV14, and 2243 bp L1D7, were located at the 3' flanking of *IGFBP1* for *IGFBP1*-1-RIP, *IGFBP1*-2-RIP, *IGFBP1*-3-RIP. 176 bp SINEA9 was located at the intron 3 of *IGFBP2* for *IGFBP2*-1-RIP. 265 bp SINEA2 and 159 bp L1D14 were located at 5' flanking and 202 bp ERV14, 141 bp SINEA3 was located at 3' flanking of *IGFBP3* for *IGFBP3*-1-RIP, *IGFBP3*-2-RIP, *IGFBP3*-3-RIP, and *IGFBP3*-4-RIP. 85 bp SINEA6 and 274 bp SINEA1 were located at the intron 2 and intron 3 of *IGFBP4*, *IGFBP5* for *IGFBP4*-1-RIP, and *IGFBP5*-1-RIP. Two SINEA1 (301 bp and 288 bp) were located at 5' flanking and intron 4 respectively of IGFBP7 for *IGFBP7*-1-RIP and *IGFBP7*-2-RIP.

RIP distribution in different pig breeds

A collection of one commercial pig breed, a range of native Chinese breeds, and crossbreed in Jiangsu province and Chinese miniature pig breeds were selected to further conduct population genetic analysis for these RIPs. The number of pigs used, breed, genotype, allele frequency, Hardy-Weinberg equilibrium test, and the polymorphism information content (PIC) was shown in Table 1. All RIPs of *IGFBPs* were polymorphic across these breeds and two alleles (RIP+/-) were detected by PCR. Differences in allele frequency were found between the commercial breed and native pig breeds for *IGFBP1*-1-RIP, *IGFBP3*-1-RIP, *IGFBP3*-2-RIP, *IGFBP4*-1-RIP, *IGFBP5*-1-RIP, *IGFBP7*-2-RIP, while RIP^+ was dominant in Chinese native pig for *IGFBP1*-1-RIP, *IGFBP1*-3-RIP, *IGFBP3*-2-RIP, *IGFBP5*-1-RIP, and RIP^- was dominant for *IGFBP3*-1-RIP, *IGFBP3*-4-RIP, *IGFBP4*-1-RIP. RIP^+ for *IGFBP1*-3-RIP and *IGFBP2*-1-RIP was only found in Chinese native pig breeds and RIP^+ for *IGFBP3*-3 was only found in commercial pig breeds. In most cases the RIPs were in Hardy-Weinberg equilibrium ($P > 0.05$), except *IGFBP1*-1-RIP, *IGFBP2*-1-RIP, *IGFBP3*-1-RIP, *IGFBP3*-2-RIP in Sujiang and Jiangquhai which is one of origination of Sujiang, *IGFBP1*-2-RIP, *IGFBP1*-3-RIP, *IGFBP2*-1-RIP,

IGFBP3-4-RIP, *IGFBP7*-2-RIP in Erhualian, *IGFBP1*-2-RIP, *IGFBP1*-3-RIP, *IGFBP3*-4-RIP in Fengjing, *IGFBP1*-3-RIP in Meishan and *IGFBP1*-2-RIP, *IGFBP7*-1-RIP in Tibetan, which deviated from the Hardy-Weinberg equilibrium ($P < 0.05$). RIPs in all breeds display moderate to low PIC (ranging from 0.375 for *IGFBP1*-3-RIP, *IGFBP3*-4-RIP in Meishan to 0.034 in *IGFBP3*-1-RIP, *IGFBP4*-1-RIP in Fengjing).

Insertion of SINE and LINE in the 5’ flanking of IGFBP3 were linked loci and closely associated with growth traits of Large white pigs

In this study, the association of RIPs in the 5’ flanking of *IGFBPs* including *IGFBP3*-1-RIP and *IGFBP3*-2-RIPwith growth traits of Large white was analyzed. Both, the SINE and LINE insertions, were at the 5’ flanking of *IGFBP3* for *IGFBP3*-1-RIP and *IGFBP3*-2-RIP with a very close genome distance (about 400 bp far away). These reside 6463 bp for SINE, and 6058 bp for LINE upstream of the ATG of *IGFBP3*, respectively (Fig. 2b). Three genotypes for SINE and LINE insertion were identified by PCR separately; and also three genotypes for the combination of the two RIPs were found, which is each homozygote for SINE, homozygote for LINE, and heterozygote for joint SINE and LINE insertion (Fig. 2a). Linkage disequilibrium tests by SHEsis shown they were in a full linkage status in all detected breeds ($r^2 = 1$) (Fig. 2c and Table 1). These two-linkage loci were named *IGFBP3*-SINE-LINE-RIP. Three combined genotype, $SINE^{+/+}/LINE^{-/-}$, $SINE^{+/-}/LINE^{-/+}$, $SINE^{-/-}/LINE^{+/+}$, were represented homozygote for SINE, homozygote for LINE and heterozygote for joint SINE and LINE insertion respectively. Association analysis revealed (Table 2) that individuals with the $SINE^{-/-}/LINE^{+/+}$ genotype had a better growth rate and larger longissimus area, and thicker corrected back fat ($P < 0.05$) compared with those of $SINE^{+/+}/LINE^{-/-}$ genotype. Individuals with the $SINE^{+/-}/LINE^{-/+}$ genotype also had better growth rate and bigger area of longissimus ($P < 0.05$) but thicker correcting back fat thickness compared with those of $SINE^{+/+}/LINE^{-/-}$.

SINE and LINE linked RIPs associated with the mRNA expression changes of pig IGFBP3

To determine whether the SINE and LINE insertions influence the expression of *IGFBP3*, we investigated the mRNA expressions of *IGFBP3* by qPCR in the liver, longissimus muscles, and backfat for different genotypes in 30-day old Sujiang piglets. The qPCR results revealed that the expression of *IGFBP3* in the liver, $SINE^{+/+}/LINE^{-/-}$ homozygotes was higher than those of of $SINE^{-/-}/LINE^{+/+}$ homozygotes ($P < 0.05$). There was a significant difference in the *IGFBP3* expression in backfat between those of $SINE^{+/+}/LINE^{-/-}$ genotype and $SINE^{+/-}/LINE^{-/+}$ genotype or $SINE^{-/-}/LINE^{+/+}$ genotype ($P < 0.05$) (Fig. 3). The results indicated that the linked SINE and LINE RIPs in the 5’ flanking of *IGFBP3* are associated with changes in mRNA expression level of *IGFBP3*.

265bp SINE may act as an enhancer to increasing the activity of the IGFBP3 promoter

Both SINE and LINE were inserted into the 5′ flanking of the *IGFBP3*. To illustrate the biological roles of the SINE or LINE in the regulation of promoter to increase the *IGFBP3* expression, we first analyzed the structure of the 5′ flanking of *IGFBP3* and predicted the core promoter region upstream of the translation start site (ATG) of the *IGFBP3* gene (Fig. 4A). Three fragments each spanning 482bp, 1452bp, 3212bp upstream to the translation start site (ATG) respectively, and relatively evolutionarily conserved among pig, human, cattle, sheep, dog, rat, and mouse (Fig. 4B), were cloned separately into a luciferase reporter vector (pGL3-basic) (Fig. 4C) and then submitted for luciferase activity evaluation. The dual-luciferase activity assay revealed that the three genomic fragments displayed obvious promoter activity compared with the pGL3-basic ($P < 0.01$) in the PK15 cell (Fig. 4D). The promoter activity driven by the 482bp fragment was almost two times higher than that driven by the 1452bp fragment, but equal to the longest 3212bp fragment. By Promoter 2.0 (http://www.cbs.dtu.dk/cgi-bin/webface2.fcgi?jobid=60545EBB000070EE2561249E&wait=20) prediction, there was one putative promoter (2700bp upstream to the first exon) with high scores (> 1) (”▯” in Fig. 4A and B). In the EPD website, the human and mouse *IGFBP3* gene promoter sequences have been confirmed to have a high degree of homology with the pig *IGFBP3* gene in the 121-181bp region upstream of the ATG. Moreover, also on the BDGP website it was predicted that this region harbors a highly active promoter (= 1) (Fig. 4A and B”▯”). Together, these data suggested that the 482bp fragment may contain the core promoter of the porcine *IGFBP3* and is sufficient to derive its expression.

To further examine whether the SINE or LINE insertion regulated activity of the promoter of *IGFBP3*, we evaluated the enhancer activity of the SINE and LINE insertion by cloning it into the luciferase reporter vector containing *IGFBP3*-promoter (482bp). The schematic diagram of vectors was shown in Fig. 5a. The luciferase activity assay revealed that the SINE insertion significantly improves *IGFBP3* gene promoter activities in C2C12 (Fig. 5b), 3T3-L1 (Fig. 5c), and HeLa cell lines (Fig. 5d) ($P < 0.05$). However, there was no obvious effect for LINE insertion. These data support that the SINE insertion may act as an enhancer and be involved in the regulation of *IGFBP3*.

Discussion

IGFs are key growth-promoting peptides that act in endocrine, autocrine, and paracrine fashion to promote different cellular functions, growth, proliferation, differentiation and survival. IGF signaling is regulated by IGFBPs which can bind IGFs with equal or greater affinity than the IGF1R in both circulation and local tissues[2, 21]. Different studies focused on the potential association of SNPs in *IGFBPs* gene with growth traits in domestic animals[22, 23, 24]. Few studies concerning retrotransposons insertion polymorphism were reported for the *IGFBPs* gene. In this study, 12 RIPs, including 7 SINE (ranged from 85bp to 301bp), 3 LINE (ranged from 159bp to 2243bp), and 2 ERV (ranged from 202bp to 327bp), in six pig *IGFBPs* gene were identified by comparative genomics and PCR. It is known that the repetitive portion of mammal genomes is dominated by SINEs and LINEs, followed by ERV [25]. Our lab also proved this finding for the pig genome[20]. Surveys about the human genome showed that transposable elements are not randomly distributed[25, 26, 27]. In our previous analysis, 2.84%, 0.63%, 35.1%, 1.38% of retrotransposons in the pig genome occupied the 5′ flanking, exon, intron, and 3′ flanking of protein-

coding genes[20]. These 12 retrotransposons in *IGFBPs* were mostly inserted into 5' or 3' flanking, only 4 retrotransposons inserted into introns and none into exons. Retrotransposons in introns can affect splicing of genes and form multiple transcripts [28]. Thus, these insertions in 5' flanks and introns may have function consequences for *IGFBPs*.

Because of artificial and natural selection, there are major differences in growth and body size among pig breeds that originated in China and Europe[29, 30]. Therefore, we investigated the population distribution of RIPs of *IGFBPs* in different pig breeds. An interesting distribution was found for *IGFBP1*-3-RIP, a LINE insertion of 2243bp, which appeared in the Chinese local pig breeds and crossbreeds of Sujiang (Duroc×Jiangquhai×Fengjing), while *IGFBP3*-3-RIP appeared in the European pig breeds. Selective forces can explain why some transposon elements can reach fixation in certain genomic locations[27]. Maybe these RIPs can be used for tracing and authentication purposes for specific breeds. Some loci (*IGFBP4*-1-RIP, *IGFBP5*-1-RIP) are in HW equilibrium, while other loci (*IGFBP1*-2-RIP, *IGFBP1*-3-RIP, *IGFBP2*-1-RIP) for more than three pig breeds deviate from HW disequilibrium, indicating that these loci may have experienced strong selection and play biological roles in gene regulation and phenotype variation. Sujiang and its cross parent Jiangquhai are in genetic disequilibrium in *IGFBP1*-1-RIP, *IGFBP2*-1-RIP, *IGFBP3*-SINE-LINE-RIP, indicating that this locus may have experienced strong purification selection considering the retrotransposon insertion playing positive roles in growth. The insertion alleles were dominated in Chinese local breeds for *IGFBP1*-1-RIP, *IGFBP1*-3-RIP, and *IGFBP3*-2-RIP, while deletion alleles were dominated for *IGFBP3*-1-RIP and *IGFBP3*-4-RIP. Whether these RIPs affect the growth between Chinese local breeds and European pig breeds should be studied further.

Previous studies have shown that the variation in *the IGFBP3* gene was closely associated with the growth traits of domestic animals. Pigs with the TAT/TAT genotype in intron 2 of *IGFBP3* had higher B-point and C-point back-fat thickness than pigs with the GGC/GGC genotype and the allelic frequency of TAT in Chinese breeds was over 50% [31]. SNP in *the IGFBP3* gene was connected closely with the energy metabolism of horse[32] and have a significant association with body weight in the Sirohi breed of the goat at different ages [33]. When we investigated the association of *IGFBP3*-SINE-LINE-RIP in 5' flanking with growth traits of Large white, we found that the *IGFBP3*-SINE-LINE-RIP were significantly associated with age at 100 kg body weight, corrected backfat thickness, longissimus area of Large white. There are some important regulatory elements in the 5' flanking of protein-coding genes and insertion polymorphism by retrotransposons can modify the transcriptional activity of target genes by changing promoter activity, and further influence the biological function of target genes. *The IGFBP3* expression was investigated for different genotypes by qPCR in 30-day old Sujiang piglets. The expression of *IGFBP3* in $SINE^{+/+}/LINE^{-/-}$ individuals in the liver, backfat was significantly higher compared to $SINE^{-/-}/LINE^{+/+}$ individuals ($P < 0.05$). It has been shown previously that IGFBP3 was the most dominant IGFBP in postnatal serum, binding to IGFs with the highest affinity to promote or inhibit growth both *in vivo* and *in vitro*. Large amounts of IGFBP3 caused a reduction in the free IGF levels while small amounts of IGFBP3 protected IGFs, intensifying their effects [34]. IGFBP3 knock-down in SMMC-7721 cells can promote cell proliferation rather than affect cell apoptosis[35]. Decreased IGFBP3 expression was related

to the over-activated IGF pathway and accelerated HCC cells proliferation[36]. In line with this, homozygotes with the SINE insertion, with a higher level of *IGFBP3* in the liver and leg muscle, showed slower growth and a smaller longissimus muscle area compared with homozygote with the LINE insertion, since there was no significant difference between two genotypes for IGFBP3 level in longissimus muscle of 30-day Sujiang piglets. IGFBP3 levels in serum are age-dependent, being low at birth and increasing during childhood to reach a peak during puberty, after which serum concentration of IGFBP3 starts to decrease[34]. The changing pattern was confirmed In EY pigs (F1 crossbreds from Erhualian boars ×Yorkshire sows) that *IGFBP3* levels reached their low peak at the perinatal period and high peak value at around puberty respectively [37]. The underlying mechanism for the effect of the spatial-temporal expression of this gene with the SINE insertion on growth traits needs to be revealed by the in-depth study.

Retrotransposons may harbor the primary types of regulatory sequences for their expression and when they are inserted upstream of protein-coding gene, they may contribute enhancer elements in a lineage-specific fashion[38, 39]. Thus, the SINE and LINE in 5' flanking of pig *IGFBP3* were furthered studied here to evaluate whether they acted as an enhancer in gene regulation. The results showed that the SINE insertion could enhance the activity of *IGFBP3* promoter in cell lines C2C12, 3T3-L1 and HeLa originated from muscle cells, adipose cells, and cancer cells while LINE showed no significant effect. According to the analysis of the age distribution, the SINEA2 subfamily was one of the youngest SINE families[20] and the sequence harbors a tRNA head which is required for RNA polymerase III dependent transcriptions. L1D14 was a relatively older subfamily compared to the younger L1D1-L1D7. Multiple studies have shown in different ways that SINEs contribute more considerably to regulatory divergence compared to LINEs[38]. This might be the reason that the expression of *IGFBP3* of liver and muscle of individuals of $SINE^{+/+}/LINE/^{-/-}$ was higher than that of individuals of $SINE^{-/-}/LINE^{+/+}$. Therefore, compared with homozygote of LINE insertion, grew faster, had a larger longissimus muscle area and thicker backfat, heterozygote of SINE and LINE insertion had a thinner backfat while a faster growth, a larger longissimus muscle area, which was more adaptive to be selected in commercial breeding.

Conclusion

In summary, twelve RIPs were identified in *IGFBP1-5* and *IGFBP7* that were detected in different breeds with an uneven distribution among them. Among them, *IGFBP3*-1-RIP and *IGFBP3*-2-RIP are completely linked, showing only three genotypes ($SINE^{+/+}/LINE^{-/-}$, $SINE^{-/-}/LINE^{+/+}$ and $SINE^{+/-}/LINE^{-/+}$). These two RIPs were associated with the age of 100 kg body weight, longissimus muscle thickness, and the corrected backfat thickness of Large white. The expression of *the IGFBP3* gene in liver, backfat, and leg muscle of 30-day Sujiang piglets with $SINE^{+/+}/LINE^{-/-}$ genotype were significantly higher ($P < 0.05$) than those with $SINE^{-/-}/LINE^{+/+}$ genotype. Only SINE insertion was significantly increased *($P < 0.05$*) promoter activity of the *IGFBP3* gene, indicating that the SINE may act as an enhancer to regulate the promoter activity of the *IGFBP3* gene. Based on the possible inhibitory effect of high expression of *IGFBP3* gene on growth and the characteristics of fast growth, thin backfat, and thick longissimus muscles exhibited by

$SINE^{+/-}/LINE^{-/+}$ heterozygotes, $SINE^{+/-}/LINE^{-/+}$ individuals should be used for commercial pigs, and it is worthy to comprehensively evaluate the influence of SINE and LINE inserting in 5' flanking of *IGFBP3* gene on pig's economic traits comprehensively.

Abbreviations

ERV	Endogenous Retrovirus
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
IGFBPs	Insulin-like growth factor-binding proteins
IGF	Insulin-like growth factor
LINE	Long interspersed nuclear element
PIC	Polymorphism information content
qPCR	Quantitative polymerase chain reaction
RIPs	Retrotransposon insertion polymorphisms
SINE	Short interspersed nuclear element
SVs	Structural Variations
TSS	Transcription start site

Declarations

Acknowledgments

Not applicable.

Authors' contributions

Xiaoyan Wang, Eduard Murani, Enrico D'alessandro, Kui Li, Klaus Wimmers and Chengyi Songdesigned research. Yalong An, Xiaoyan Wang, Chengling Chi, Cai Chen performed, analyzed, and interpreted all experiments. Xiaoyan Wang and Yalong An wrote the paper. Eduard Murani, Ali Shoaib Moawad and Chengyi Song finally approved the manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (31872977, 32002146), China Postdoctoral Science Foundation (2020M671630), and the Jiangsu Agriculture Science and Technology Innovation Fund [CX(19)2016]; and the Priority Academic Program Development of Jiangsu Higher Education Institutions and the High-end Talent Support Program of Yangzhou University.

Availability of data and materials

All data needed to evaluate the conclusions in this paper are present either in the main text or the Supporting information.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interest.

References

1. Song F, Zhou XX, Hu Y, Li G, Wang Y. The Roles of Insulin-Like Growth Factor Binding Protein Family in Development and Diseases. Adv. Ther. 2021;38(2):885-903.
2. Jones JI, Clemmons DR. Insulin-like growth factors and their binding proteins: Biological actions. Endocr Rev. 1995;16(1):3-34.
3. Mohan S, Baylink DJ. IGF-binding proteins are multifunctional and act via IGF-dependent and -independent mechanisms. J Endocrinol. 2002;175(1):19-31.
4. Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet. 2014;46(11):1173-1186.
5. Bartell E, Fujimoto M, Khoury JC, Khoury PR, Vedantam S, Astley CM, et al. Protein QTL analysis of IGF-I and its binding proteins provides insights into growth biology. Hum Mol Genet. 2020;29(15):2625-2636.
6. Wang W, Ouyang K, Shangguan X, Xu M. Association of Porcine IGF Binding Protein-5 Gene with Meat Quality. Biochem Genet. 2010;48(3-4):257-265.
7. Verardo LL, Sevón-Aimonen ML, Serenius T, Hietakangas V, Uimari P. Whole-genome association analysis of pork meat pH revealed three significant regions and several potential genes in Finnish Yorkshire pigs. BMC Genet. 2017;18(1):13.
8. Grilz-Seger G, Neuditschko M, Ricard A, Velie B, Lindgren G, Mesarič M, et al. Genome-wide homozygosity patterns and evidence for selection in a set of european and near eastern horse breeds. Genes (Basel). 2019;10(7):491.
9. Maciulla JH, Zhang HM, DeNise SK. A novel polymorphism in the bovine insulin-like growth factor binding protein-3 (IGFBP3) gene. Anim. Genet. 1997;28(5):375.

10. Liu H, Liu C, Yang G, Li H, Dai J, Cong Y, et al. DNA polymorphism of insulin-like growth factor-binding protein-3 gene and its association with cashmere traits in cashmere goats. Asian-Australasian J Anim Sci. 2012;25(11):1515-1520.
11. Platt RN, Vandewege MW, Ray DA. Mammalian transposable elements and their impacts on genome evolution. Chromosom Res. Chromosome Research; 2018;26:25–43.
12. Garcia-Perez JL, Widmann TJ, Adams IR. The impact of transposable elements on mammalian development. Dev. 2016;143(22):4101-4114.
13. Chuong EB, Elde NC, Feschotte C. Regulatory activities of transposable elements: From conflicts to benefits. Nat. Rev. Genet. 2017;18(2):71-86.
14. Kalendar R. The use of retrotransposon-based molecular markers to analyze genetic diversity. Ratar i Povrt. 2011;48:261–274.
15. Kalendar R, Amenov A, Daniyarov A. Use of retrotransposon-derived genetic markers to analyse genomic variability in plants. Funct. Plant Biol. 2019;46(1):15-29..
16. Payer LM, Burns KH. Transposable elements in human genetic disease. Nat. Rev. Genet. 2019;20(12):760-772..
17. Yeh F, Yang R, Boyle T. POPGENE Version 1.32 Microsoft Windows-based freeware for populations genetic analysis. University of Alberta, Edmonton. 1999;
18. Botstein D, White RL, Skolnick M, Davis RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. 1980;32(3):314-331..
19. Li Z, Zhang Z, He Z, Tang W, Li T, Zeng Z, et al. A partition-ligation-combination-subdivision em algorithm for haplotype inference with multiallelic markers: Update of the SHEsis (http://analysis.bio-x.cn). Cell Res. 2009;19(4):519-523.
20. Chen C, Wang W, Wang X, Shen D, Wang S, Wang Y, et al. Retrotransposons evolution and impact on lncRNA and protein coding genes in pigs. Mob DNA. 2019;10:19.
21. Allard JB, Duan C. IGF-binding proteins: Why do they exist and why are there so many? Front Endocrinol (Lausanne). 2018;9:1–12.
22. Prasongsook S, Choi I, Bates RO, Raney NE, Ernst CW, Tumwasorn S. Association of Insulin-like growth factor binding protein 2 genotypes with growth, carcass and meat quality traits in pigs. J Anim Sci Technol. 2015;57:31.
23. Leng L, Wang S, Li Z, Wang Q, Li H. A polymorphism in the 3'-flanking region of insulin-like growth factor binding protein 2 gene associated with abdominal fat in chickens. Poult Sci. 2009;88(5):938-942.
24. Hosnedlova B, Vernerova K, Kizek R, Bozzi R, Kadlec J, Curn V, et al. Associations between IGF1, IGFBP2 and tgfß3 genes polymorphisms and growth performance of broiler chicken lines. Animals. 2020;10(5):800.
25. Platt RN, Vandewege MW, Ray DA. Mammalian transposable elements and their impacts on genome evolution. Chromosom. Res. 2018;26(1-2):25-43.

26. Sundaram V, Wysocka J. Transposable elements as a potent source of diverse cis-regulatory sequences in mammalian genomes. Philos. Trans. R. Soc. B Biol. Sci. 2020;375(1795):20190347.
27. Bourque G, Burns KH, Gehring M, Gorbunova V, Seluanov A, Hammell M, et al. Ten things you should know about transposable elements. Genome Biol. 2018 Nov 19;19(1):199.
28. Göke J, Ng HH. CTRL + INSERT : retrotransposons and their contribution to regulation and innovation of the transcriptome . EMBO Rep. 2016;17:1131–1144.
29. Wilkinson S, Lu ZH, Megens HJ, Archibald AL, Haley C, Jackson IJ, et al. Signatures of Diversifying Selection in European Pig Breeds. PLoS Genet. 2013;9(4):e1003453.
30. Yang S, Li X, Li K, Fan B, Tang Z. A genome-wide scan for signatures of selection in Chinese indigenous and commercial pig breeds. BMC Genet. 2014 15;15:7.
31. Wang W, Meng Q, Hu X, Li N. Genetic variation and association of insulin-like growth factor binding protein-3 with performance in swine. Biochem Genet. 2009;47(3-4):315-321.
32. Yu H, Cheng Y, Shi H, Jia H, Xia J, Hao L, et al. Linkage disequilibrium and functional analysis of PRE1 insertion together with SNPs in the promoter region of IGFBP7 gene in different pig breeds. J Appl Genet. Journal of Applied Genetics; 2018;59:231–241.
33. Sharma A, Dutt G, Sivalingam J, Singh MK, Pathodiya OP, Khadda BS, et al. Novel SNPs in IGF1, GHR and IGFBP-3 genes reveal significant association with growth traits in Indian goat breeds. Small Rumin Res. 2013, 115(1-3):7-14.
34. Collett-Solberg PF, Cohen P. Genetics, chemistry, and function of the IGF/IGFBP system. Endocrine. 2000;12(2):121-136.
35. Yin J, Ding J, Huang L, Tian X, Shi X, Zhi L, et al. SND1 Affects Proliferation of Hepatocellular Carcinoma Cell Line SMMC-7721 by Regulating IGFBP3 Expression. Anat Rec. 2013;296(10):1568-1575.
36. Huynh H, Chow PKH, Ooi LLP, Soo KC. A possible role for insulin-like growth factor-binding protein-3 autocrine/paracrine loops in controlling hepatocellular carcinoma cell proliferation. Cell Growth Differ. 2002;13(3):115-122.
37. Pan Z, Zhang J, Zhang J, Zhou B, Chen J, Jiang Z, et al. Expression profiles of the insulin-like growth factor system components in liver tissue during embryonic and postnatal growth of Erhualian and Yorkshire reciprocal cross F1 pigs. Asian-Australasian J Anim Sci. 2012;25(7):903-912.
38. Ali A, Han K, Liang P. Role of transposable elements in gene regulation in the human genome. Life. 2021;11(2):118.
39. Ye M, Goudot C, Hoyler T, Lemoine B, Amigorena S, Zueva E. Specific subfamilies of transposable elements contribute to different domains of T lymphocyte enhancers. Proc Natl Acad Sci U S A. 2020;117(14):7905-7916.

Tables

Table 1
Analysis of RIP Distribution in Different Breeds

Loci	Breeds	N	Genotype			Allele frequency		P value	PIC
			+/+	+/-	-/-	+	-		
IGFBP1-1-RIP	Large white	24	0.00	33.33	66.67	16.67	83.33	0.327	0.239
	Sujiang	24	8.33	12.50	79.17	14.58	85.42	0.015	0.218
	Meishan	24	37.50	41.67	20.83	58.33	41.67	0.484	0.367
	Erhualian	24	91.67	8.33	0.00	95.83	4.17	0.831	0.077
	Fengjing	24	83.33	16.67	0.00	91.67	8.33	0.656	0.141
	Jiangquhai	24	58.33	16.67	25.00	66.67	33.33	0.002	0.346
IGFBP1-2-RIP	Large white	24	0.00	33.33	66.67	16.67	83.33	0.327	0.239
	Sujiang	24	8.33	12.50	79.17	14.58	85.42	0.015	0.218
	Meishan	24	16.67	45.83	37.50	39.58	60.42	0.838	0.364
	Erhualian	24	4.17	4.17	91.67	6.25	93.75	0.002	0.110
	Fengjing	24	58.33	0.00	41.37	58.33	41.67	9.63E-07	0.368
	Tibet pig	24	58.33	0.00	41.37	58.33	41.67	9.63E-07	0.368
	Wuzhishan	24	8.33	41.67	50.00	29.17	70.83	0.967	0.328
IGFBP1-3-RIP	Sujiang	24	0.00	16.67	83.33	8.33	91.67	0.656	0.141
	Meishan	24	41.67	20.83	37.50	52.08	47.92	0.004	0.375
	Erhualian	24	83.33	4.17	12.50	85.42	14.58	4.51E-05	0.218
	Fengjing	24	91.67	4.17	4.17	93.75	6.25	0.002	0.110
IGFBP2-1-RIP	Sujiang	24	25.00	4.17	70.83	27.08	72.92	1.18E-05	0.317
	Meishan	24	50.00	50.00	0.00	75.00	25.00	0.103	0.305
	Erhualian	24	37.50	0.00	62.50	37.50	62.50	9.63E-07	0.359
	Fengjing	24	0.00	20.83	79.17	10.42	89.58	0.569	0.169
	Jiangquhai	24	16.67	12.50	70.83	22.92	77.08	0.002	0.291
IGFBP3-1-RIP	Large white	24	62.50	37.50	0.00	81.25	18.75	0.258	0.258
	Sujiang	24	79.17	12.50	8.33	85.42	14.58	0.015	0.218

Loci	Breeds	N	Genotype			Allele frequency		P value	PIC
	Meishan	24	16.67	45.83	37.50	39.58	60.42	0.838	0.364
	Erhualian	24	0.00	8.33	91.67	4.17	95.83	0.831	0.077
	Fengjing	24	0.00	4.17	95.83	2.08	97.92	0.917	0.040
	Jiangquhai	24	33.33	4.17	62.50	35.42	64.58	8.48E-06	0.353
	Tibetan pig	24	0.00	4.17	95.83	2.08	97.92	0.917	0.040
IGFBP3-2-RIP	Large white	24	0.00	37.50	62.50	18.75	81.25	0.258	0.258
	Sujiang	24	8.33	12.50	79.17	14.58	85.42	0.015	0.218
	Meishan	24	37.50	45.83	16.67	60.42	39.58	0.838	0.364
	Erhualian	24	91.67	8.33	0.00	95.83	4.17	0.831	0.077
	Fengjing	24	95.83	4.17	0.00	97.92	2.08	0.917	0.040
	Jiangquhai	24	62.50	4.17	33.33	64.58	35.42	8.48E-06	0.353
	Tibetan pig	24	95.83	4.17	0.00	97.92	2.08	0.917	0.040
IGFBP3-3-RIP	Large white	24	0.00	25.00	75.00	12.50	87.50	0.484	0.195
IGFBP3-4-RIP	Large white	24	58.33	41.67	0.00	79.17	20.83	0.197	0.275
	Sujiang	24	0.00	8.33	91.67	4.17	95.83	0.831	0.077
	Meishan	24	20.83	54.17	25.00	47.92	52.08	0.676	0.375
	Erhualian	24	4.17	75.00	20.83	41.67	58.33	0.008	0.368
	Fengjing	24	8.33	0.00	91.67	8.33	91.67	9.63E-07	0.141
	Jiangquhai	24	0.00	33.33	66.67	16.67	83.33	0.327	0.239
IGFBP4-1-RIP	Large white	24	4.17	62.50	33.33	35.42	64.58	0.073	0.353
	Sujiang	24	0.00	29.17	70.83	14.58	85.42	0.403	0.218
	Meishan	24	0.00	12.50	87.50	6.25	93.75	0.744	0.110
	Erhualian	24	0.00	29.17	70.83	14.58	85.42	0.403	0.218
	Fengjing	24	0.00	4.17	95.83	2.08	97.92	0.917	0.040
	Jiangquhai	24	0.00	2.08	79.17	10.42	89.58	0.569	0.169

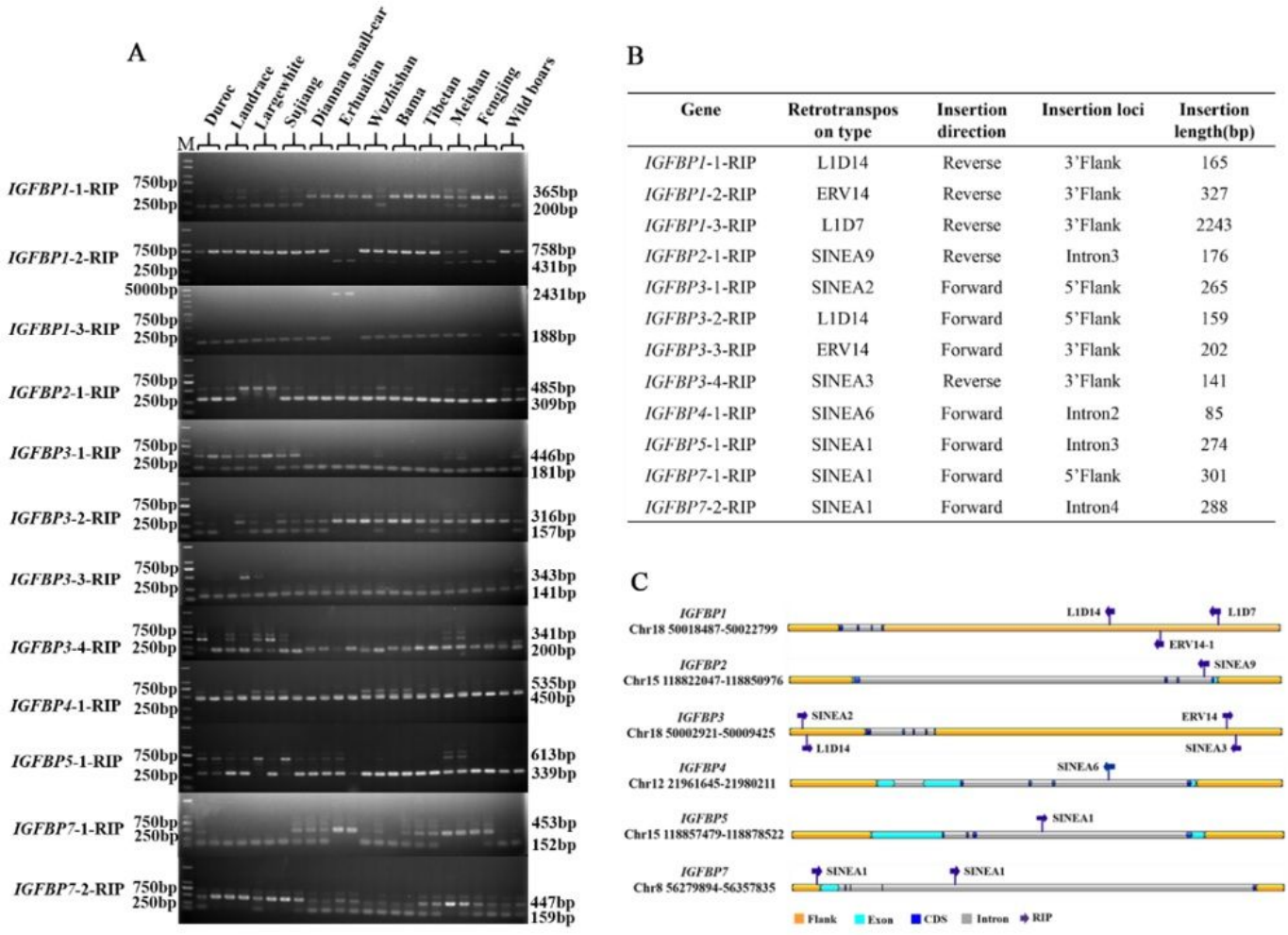
Loci	Breeds	N	Genotype			Allele frequency		P value	PIC
IGFBP5-1-RIP	Large white	24	4.17	62.50	33.33	35.42	64.58	0.073	0.353
	Sujiang	24	29.17	50.00	20.83	54.17	45.83	0.973	0.373
	Meishan	24	58.33	41.67	0.00	79.17	20.83	0.197	0.275
	Erhualian	24	45.83	33.33	20.83	62.50	37.50	0.157	0.359
	Jiangquhai	24	16.67	41.67	41.67	37.50	62.50	0.586	0.359
IGFBP7-1-RIP	Large white	24	0.00	50.00	50.00	25.00	75.00	0.102	0.305
	Sujiang	24	8.33	41.67	50.00	29.17	70.83	0.967	0.328
	Erhualian	24	91.67	8.33	0.00	95.83	4.17	0.831	0.077
	Fengjing	24	91.67	8.33	0.00	95.83	4.17	0.831	0.077
	Tibetan pig	24	16.67	79.17	4.17	56.25	43.75	0.003	0.371
	Wuzhishan	24	0.00	8.33	91.67	4.17	95.83	0.831	0.077
IGFBP7-2-RIP	Large white	24	91.67	8.33	0.00	95.83	4.17	0.831	0.077
	Sujiang	24	62.50	33.33	4.17	79.17	20.83	0.959	0.275
	Meishan	24	83.33	16.67	0.00	91.67	8.33	0.656	0.141
	Erhualian	24	41.67	58.33	0.00	70.83	29.17	0.048	0.327
	Fengjing	24	33.33	54.17	12.50	60.42	39.58	0.516	0.364
	Jiangquhai	24	58.33	33.33	8.33	75.00	25.00	0.586	0.305

Table 2
Association analysis between *IGFBP3*-SINE-LINE-RIP and growth traits of Large white

Genotype	Age at 100 kg body weight/d		Correcting back fat thickness/mm		Loin eye muscle area/mm	
	N	Mean ± SD	N	Mean ± SD	N	Mean ± SD
$SINE^{+/+}/LINE^{-/-}$	604	162.66 ± 9.02 a	604	10.81 ± 2.58 a	340	56.29 ± 4.86 a
$SINE^{+/-}/LINE^{-/+}$	237	160.33 ± 10.25 b	237	11.04 ± 2.68 a	108	58.06 ± 5.23 b
$SINE^{-/-}/LINE^{+/+}$	22	159.24 ± 6.92 b	22	13.28 ± 3.62 b*	11	59.99 ± 4.22 b

The same letter in the same column indicates no significant difference between groups. Different letters indicated significant difference between groups ($P < 0.05$). The band * indicates that the difference is extremely significant ($P < 0.01$). The data were shown as average value ± standard error

Figures



Gene	Retrotransposon type	Insertion direction	Insertion loci	Insertion length(bp)
IGFBP1-1-RIP	L1D14	Reverse	3'Flank	165
IGFBP1-2-RIP	ERV14	Reverse	3'Flank	327
IGFBP1-3-RIP	L1D7	Reverse	3'Flank	2243
IGFBP2-1-RIP	SINEA9	Reverse	Intron3	176
IGFBP3-1-RIP	SINEA2	Forward	5'Flank	265
IGFBP3-2-RIP	L1D14	Forward	5'Flank	159
IGFBP3-3-RIP	ERV14	Forward	3'Flank	202
IGFBP3-4-RIP	SINEA3	Reverse	3'Flank	141
IGFBP4-1-RIP	SINEA6	Forward	Intron2	85
IGFBP5-1-RIP	SINEA1	Forward	Intron3	274
IGFBP7-1-RIP	SINEA1	Forward	5'Flank	301
IGFBP7-2-RIP	SINEA1	Forward	Intron4	288

Figure 1

PCR identification and characteristic of RIPs in IGFBPs. a RIPs were identified in DNA pool including eleven pig breeds and wild boars by PCR; b Characteristic of RIPs in IGFBPs including retrotransposon type, direction, insertion loci and length; c Schematic diagram of IGFBPs structure and location of RIPs in IGFBPs.

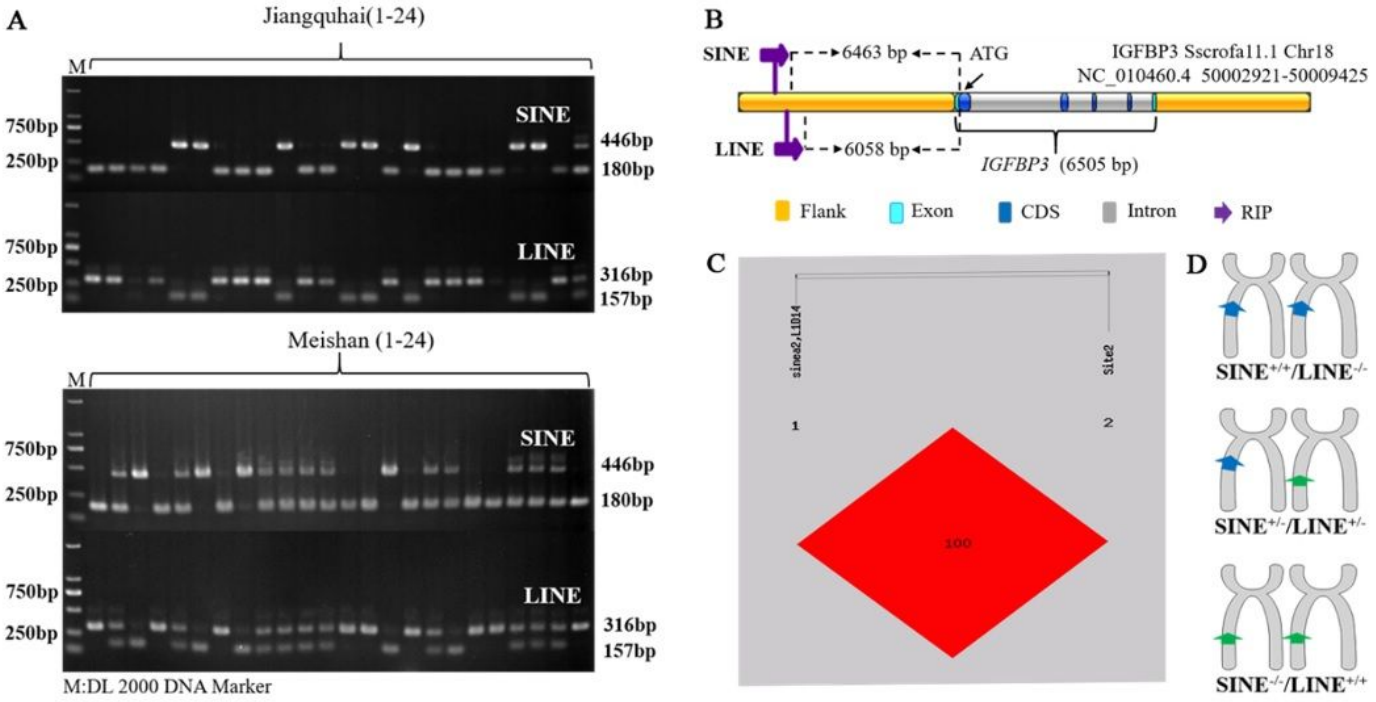


Figure 2

Insertion of SINE and LINE in the 5’ flanking of IGFBP3 have linked loci. a Polymorphism of SINE and LINE were identified in Jiangquhai and Meishan pigs by PCR using the same 24 individuals; b Schematic diagram SINE and LINE in the 5’ flanking of IGFBP3; c Linkage disequilibrium coefficients D' and r2 of the SINE and LINE insertion in IGFBP3 by SHEsis. D' value equal to 1, r2 value equal to 1; d Schematic diagram of the genotype of SINE and LINE insertion.

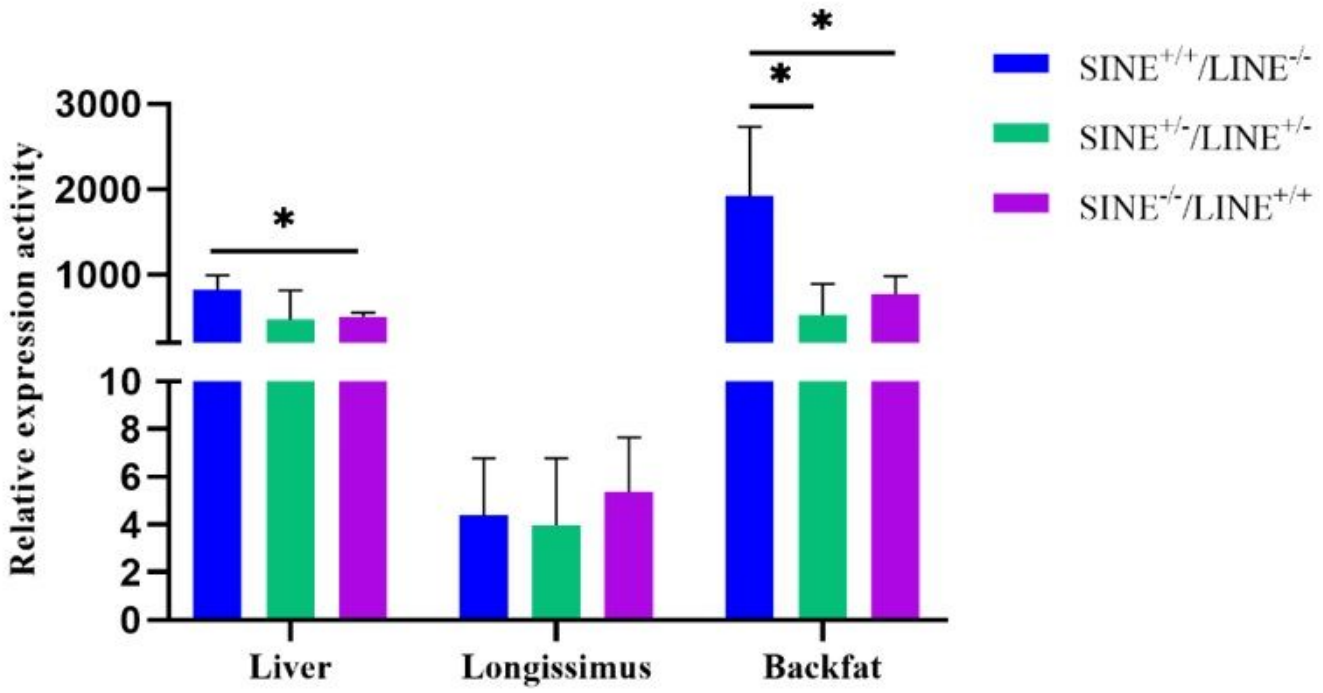


Figure 3

Effect of SINE and LINE insertion on the expression of IGFBP3 in different tissues of 30-day old piglets. 8, 4, 3 piglets for SINE+/+/LINE-/-, SINE+/-/LINE+/-, SINE-/-/LINE+/+ genotype respectively were selected for qPCR. All measurements were performed in 3 replicates for each individual. GAPDH was used to normalize the target gene expression. The values shown are mean ± SD. * showed $P<0.05$; ** showed $P<0.01$.

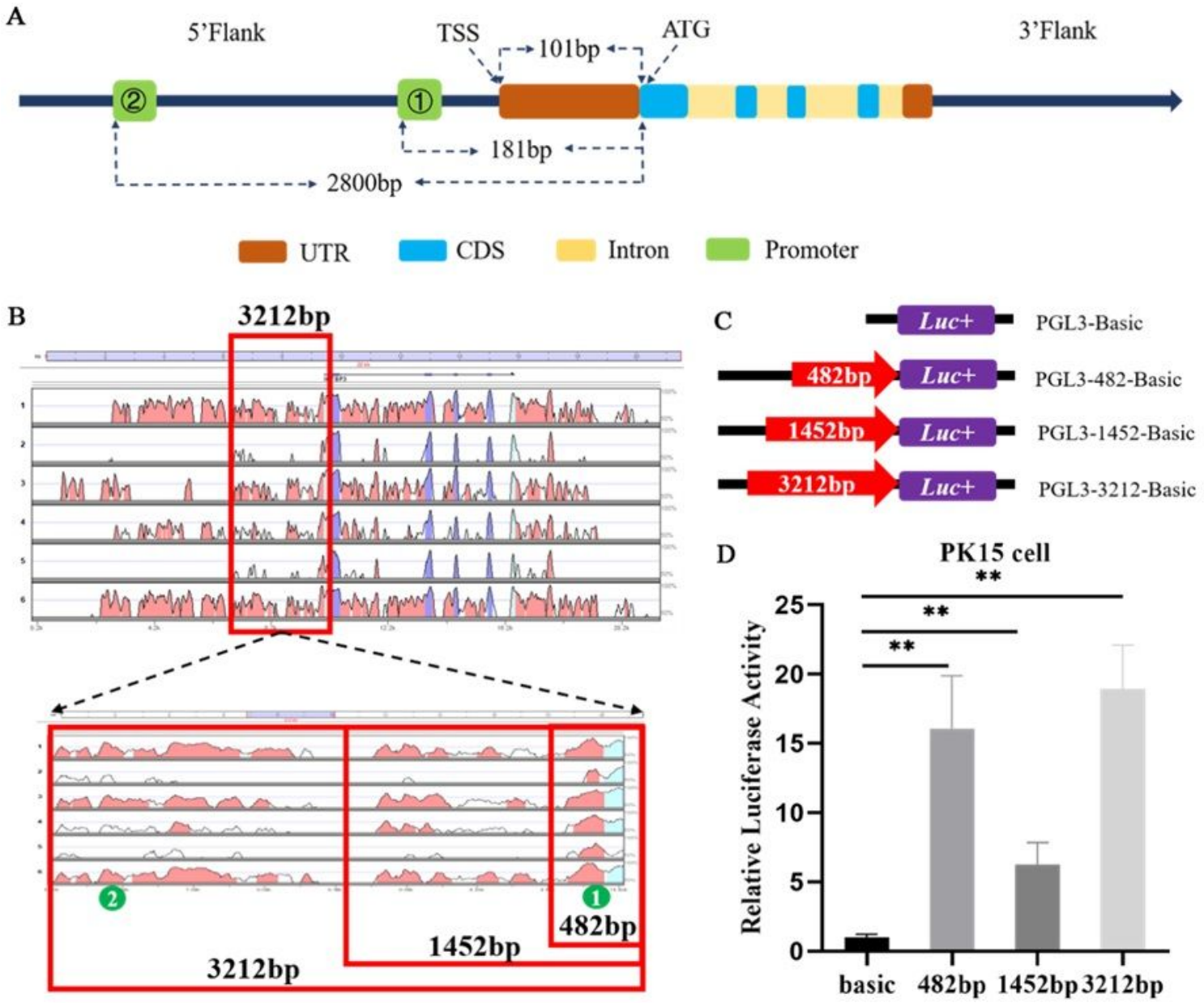


Figure 4

Identify the core promoter of IGFBP3. a Core promoter region prediction (TSS, Transcription start site; ATG, Start codon);b Sequence alignment of IGFBP3 among pig, human, cattle, rat and mouse; c Schematic diagram of the recombinant vector using pGL3-basic vector, En: SV40 enhancer; d Results on the luciferase activity assays. ** showed P<0.01 between groups (P<0.01).

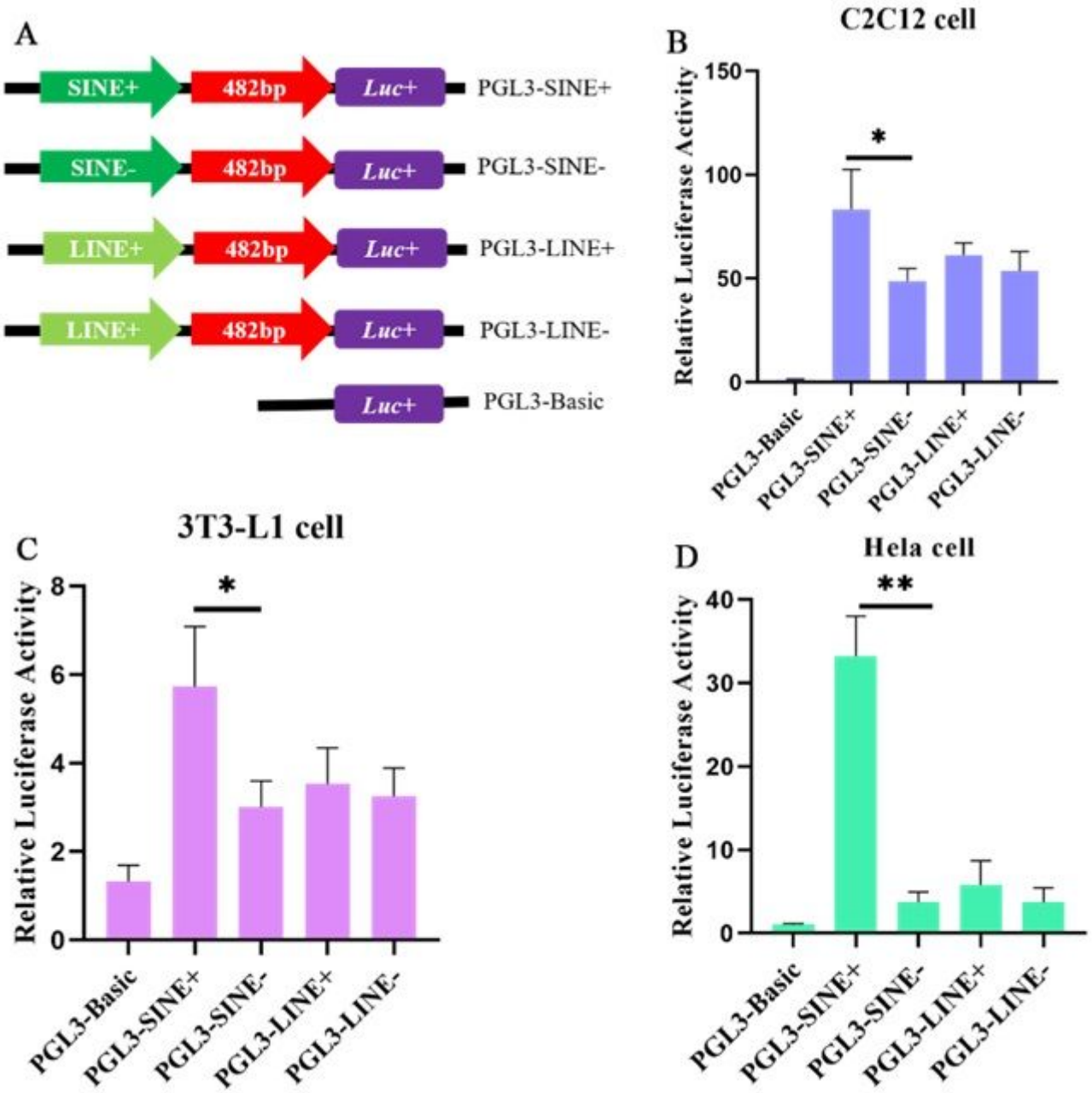


Figure 5

Effect of SINE and LINE insertions on IGFBP3 promoter activity. a Plasmid diagram of the recombinant vector using pGL3-basic vector; b Results on the luciferase activity assays of SINE and LINE on the activity of IGFBP3 Promoter in C2C12; c Results on the luciferase activity assays of SINE and LINE on the activity of IGFBP3 Promoter in 3T3-L1; d Results on the luciferase activity assays of SINE and LINE on the activity of IGFBP3 Promoter in Hela cell. * showed P<0.05; ** showed P<0.01.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryTables.docx