Integrative analysis of transcriptome complexity in pig granulosa cells by long-read isoform sequencing

Shuxin Li, Jiarui Wang, Jiale Li, Meihong Yue, Chuncheng Liu, Libing Ma and Ying Liu

School of Life Science and Technology, Inner Mongolia University of Science & Technology, Baotou, Inner Mongolia, China

ABSTRACT

Background. In intensive and large-scale farms, abnormal estradiol levels in sows can cause reproductive disorders. The high incidence rate of reproductive disturbance will induce the elimination of productive sows in large quantities, and the poor management will bring great losses to the pig farms. The change in estradiol level has an important effect on follicular development and estrus of sows. To solve this practical problem and improve the productive capacity of sows, it is significant to further clarify the regulatory mechanism of estradiol synthesis in porcine granulosa cells (GCs). The most important function of granulosa cells is to synthesize estradiol. Thus, the studies about the complex transcriptome in porcine GCs are significant. As for precursor-messenger RNAs (pre-mRNAs), their post-transcriptional modification, such as alternative polyadenylation (APA) and alternative splicing (AS), together with long non-coding RNAs (lncRNAs), may regulate the functions of granulosa cells. However, the above modification events and their function are unclear within pig granulosa cells.

Methods. Combined PacBio long-read isoform sequencing (Iso-Seq) was conducted in this work for generating porcine granulosa cells' transcriptomic data. We discovered new transcripts and possible gene loci via comparison against reference genome. Later, combined Iso-Seq data were adopted to uncover those post-transcriptional modifications such as APA or AS, together with lncRNA within porcine granulosa cells. For confirming that the Iso-Seq data were reliable, we chose four AS genes and analyzed them through RT-PCR.

Results. The present article illustrated that pig GCs had a complex transcriptome, which gave rise to 8,793 APA, 3,465 AS events, 703 candidate new gene loci, as well as 92 lncRNAs. The results of this study revealed the complex transcriptome in pig GCs. It provided a basis for the interpretation of the molecular mechanism in GCs.

Subjects Agricultural Science, Biochemistry, Cell Biology, Molecular Biology
Keywords Pig, Granulosa cells, Iso-Seq, LncRNAs, Alternative splicing

INTRODUCTION

In mammals, granulosa cells (GCs) have very important function in oocyte development (*Hoffmann & Maser, 2007*; *Park et al., 2004*; *Robinson et al., 2012*; *Yamochi, Hashimoto & Morimoto, 2021*). Every follicle includes one GCs-surrounded oocyte, while GCs can function in supporting the oocyte. GCs can regulate follicular development and

secondary sexual characteristics of female animals through hormone secretion (*McGee & Hsueh, 2000*); meanwhile, GCs participate in the growth and atresia of follicles through proliferation and apoptosis, which play a very important role in the reproductive process of female animals (*Manabe et al., 2004*; *Quirk et al., 2004*). Therefore, it is significant to investigate the regulatory signaling pathways and the transcriptional/post-transcriptional mechanisms in GCs apoptosis and hormone synthesis. So far, pork is still the main variety of meat consumption of Chinese residents. So pig is an important species in animal husbandry. Improving the fecundity of sows is of great significance to the benefits of pig breeding. The studies about the transcriptional events in pig GCs can provide experimental basis to improve the sow fertility.

The gene function and mechanism of GCs have been illustrated through a variety of research methods. Among them, genome sequencing is the basis of genomic research, and it has been widely used in animal research. The second-generation sequencing is a frequently used method to analyze the difference in gene expression. However, the long reading length of the sequence in second-generation measurement is only 50-500 bp. The short read techniques are associated with some innate drawbacks like difficult discrimination of paralogous sequences, GC bias, difficult repetitive element mapping, as well as difficult allele phasing (*Ardui et al., 2018*). The long-read mappability accounts for a key factor to confirm repetitive elements, gene fusions and gene isoforms. Fortunately, third-generation sequencing has solved the problems encountered by second-generation technique without repeating element limitation, and its long reading length is more than 20 kb (*Eid et al., 2009*; *Korlach et al., 2010*). Second-generation sequencing has been widely used in revealing the regulation of gene expression in GCs (*Du et al., 2021*; *Li et al., 2021b*; *Toms et al., 2021*; *Xu et al., 2022*), but third-generation sequencing is rarely used. Third generation sequencing can better reveal the phenomenon of complex gene transcription regulation, and provide a basis for analyzing the function of gene regulation in GCs.

Numerous articles are conducted to explore GCs-related gene functions as well as the precise mechanisms. Previous studies have suggested that miRNAs (*Tian et al., 2020*; *Toms et al., 2021*), lncRNAs (*Hu et al., 2021*; *Meng et al., 2021*; *Ruszkowska et al., 2018*; *Zhang et al., 2019*) and mRNAs (*Kulus et al., 2021*; *LaVoie, 2017*; *Shi & Sirard, 2021*) have critical effects on modulating GCs function along with follicular development in pigs. In addition, for precursor-messenger RNAs (pre-mRNAs), their post-transcriptional modifications like alternative polyadenylation (APA) and alternative splicing (AS) enrich the proteome diversity and have critical effect in different tissues and different developmental stages (*Baralle & Giudice, 2017*; *Di Giammartino, Nishida & Manley, 2011*; *Gruber & Zavolan, 2019*). In pigs, previous studies report that AS plays an essential role in estrum of sows and in male fertility (*Tang et al., 2018*; *Zhang et al., 2018*). In pig GCs, cellular-FLICE like inhibitory protein (cFLIP) and estrogen receptor beta (ER beta) have two alternative splicing isoforms and participate in follicular development (*LaVoie et al., 2002*; *Matsuda-Minehata et al., 2005*). APA is associated with the rapid and slow porcine muscle development (*Deng et al., 2020*). In a word, there are few reports on the occurrence and function of gene AS and APA in GCs and the follicular development of pigs.

The PacBio Isoform sequencing technology (Iso-Seq) is an effective method for the integrative analysis of transcriptome complexity for revealing gene functions as well as mechanisms. This study employed PacBio third-generation sequencing for illustrating post-transcriptional modifications of Yorkshire pig GCs. Besides, reverse transcription PCR (RT-PCR) was adopted in combination to investigate different spliceosomes of genes. According to our results, GCs of pig showed transcriptome complexity, which could serve as the basis to reveal the mechanism of follicular development and the candidate DNA marker in porcine marker-assisted selection (MAS).

MATERIALS & METHODS

Ovary collection

Each animal experiment was conducted according to China Council on Animal Care guidelines. Our study protocols gained approval from NMGKJDX Laboratory Animal Management Committee (NMGKJDX-2019-10). We collected and kept adult Yorkshire pig ovaries within sterile normal saline (0.9% NaCl, 1% penicillin-streptomycin) under 37 °C at the local slaughterhouse. Then the ovaries were delivered to the laboratory as soon as possible.

GCs cell culture

To culture cells, we used a 20 ml syringe to aspirate follicular fluid in 3–5 mm follicles, which was later subject to 5 min centrifugation at 1,000 g prior to GCs collection. We later rinsed GCs thrice by serum-free DMEM/F12 medium (Invitrogen, Waltham, MA, USA), dispersed and inoculated them within the six cm plates at 1×10^6/well with DMEM/F12 (5 ml) that contained 10% fetal bovine serum (Invitrogen, Waltham, MA, USA). Thereafter, we cultivated cells under 37 °C and 5% CO_2 conditions for 72 h period. We replaced the medium at 24 h intervals (*Yu et al., 2016*).

Transcriptome library preparation and sequencing

We utilized TRIzol reagent (Invitrogen, Waltham, MA, USA) to extract total cellular RNA in line with specific protocols. After purification, we adopted DNase I (Qiagen, Hilden, Germany) for enzymatic digestion. Later, we utilized the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) to assess RNA purity and content. Meanwhile, we adopted SMARTer PCR cDNA Synthesis Kit (Takara Biotechnology, Shiga, Japan) to generate cDNA and constructed libraries with SMRTbell[TM] Template Prep Kit 1.0 (Pacific Biosciences, Menlo Park, CA, USA). Subsequently, we adopted Sequel[TM] Sequencing Kit 2.0 (Pacific Biosciences) for library sequencing (Shanghai Personalbio Technology Co., Ltd., Shanghai, China).

Iso-Seq data processing

Figure 1 presented bioinformatic analysis pipeline. We employed SMRT Link v8.0 (https://www.pacb.com/wp-content/uploads/SMRT-Link-User-Guide-v8.0.pdf) to preprocess those Iso-Seq raw reads. Then, we obtained subreads through removing polymerase read adapters. With the parameters below, we acquired circular consensus sequencing
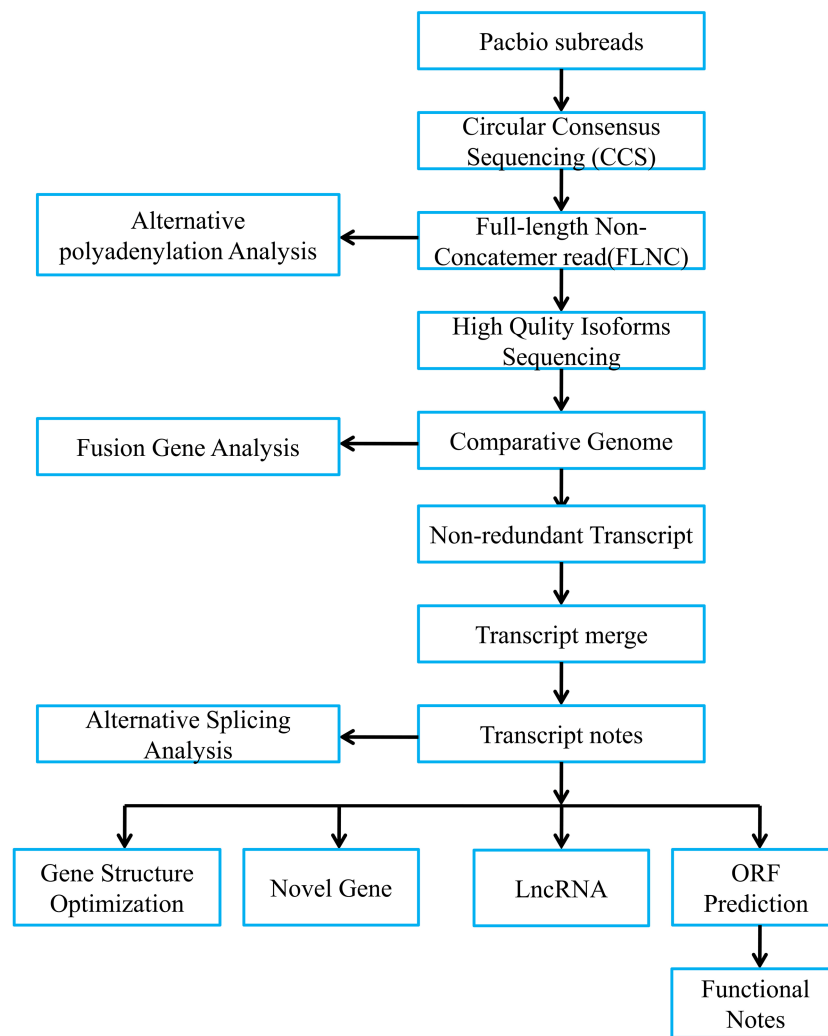
Figure 1 Overview of bioinformatic pipeline.

Full-size ⬚ DOI: 10.7717/peerj.13446/fig-1

(CCS) reads based on those subreads, which were maximum/minimum subread length = 15,000/50, minimum prediction accuracy = 0.99 and minimum pass number = 3. Afterwards, we adopted lima software to divide CCS as full-length reads. Typically, full-length CCS reads that contained the 3′ and 5′ polyA and cDNA primers were deemed as full-length non-chimeric reads (FLNCs). Using default parameters, we conducted FLNC correction of high-quality reads through LoRDEC v0.9 software. Iso-seq data were examined using SQANTI2 (*Tardaguila et al., 2018*). We finally utilized GMAP to align FLNCs into pig reference genome (https://asia.ensembl.org/Sus_scrofa_largewhite/Location/Genome?db=Core).

Table 1 Summarized information of Iso sequencing (Iso-Seq) data.

Sample	A1
Reads of Insert	385,343
Read Bases of Insert	1,027,326,158
Mean Read Length of Insert	2,666
Mean Read Quality of Insert	1
Mean Number of Passes	39.68

Identification of lncRNAs

This study discovered possible lncRNAs from porcine GCs according to our prior work (*Yang et al., 2017*). To be specific, we first eliminated short (<200 nt) transcripts, single-exons, as well as transcripts overlapped with those constructed gene models of consensus transcriptome. Secondly, we adopted 4 programs, namely, CPC (v0.9-r2) (*Kang et al., 2017*), CNCI (v2) (*Sun et al., 2013*), PLEK v1.2 (minlength−200) (*Li, Zhang & Zhou, 2014*) as well as CPAT v1.2.4 (default parameters)(*Wang et al., 2013*) for predicting new isoforms-based lncRNAs.

Isolation of RNA and reverse transcription PCR (RT-PCR)

We treated 1 µg total RNA after purification to be the template to synthesize cDNA by adopting Moloney murine leukemia virus (M-MLV, Promega, USA). Non-template controls were included in each RT-PCR. We adopted Bio-rad T100 Thermo Cycler (Bio-rad, Hercules, CA, USA) to conduct RT-PCR with a Taq PCR Mix (TaKaRa, Japan) in a 30 µl reaction system. The reaction conditions were shown below, 5 min under 95 °C; 35 s under 95 °C, 30 s under 58 °C, 40 s under 72 °C for 35 cycles; and final 40 s under 72 °C. Primers utilized in the present work were prepared by Primer 5.0, as shown in Table S1.

RESULTS

PacBio Iso-Seq and bioinformatic analysis

After quality control, 385,343 CCS reads (mean depth, 39.68 passes) and 1,027,326,158 filtered subreads (average length, 2,666 bp) were generated (Table 1). After insertion fragment recognition, sequences were classified as three types: non-full length sequence (nFL_reads), full-length chimeric sequences (FL_chimera_reads) and full-length non-chimeric sequences (FL_non-chimera_reads, FLNCs). The latter included full-length non-chimeric sequences with PolyA (FLNC_withPolyA_reads) and full-length non-chimeric sequences without PolyA (FLNC_withoutPolyA_reads). The read classification results were shown in Table 2. Finally, there were only 301,082 FLNCs used for subsequent analysis. FLNCs sequence was clustered and corrected by isoseq3 cluster software to obtain the high quality (HQ) isoform (Table 3). Subsequent analysis was conducted based on the HQ isoform.

Gene loci and isoform detection

FLNC sequence was corrected to analyze the alignment position as well as sequencing errors. Based on FLNC alignment position, each gene locus and isoform was detected.

Table 2 Reads of insert (ROI) classification statistics.

Classification	A1
Total_ccs	385,343
nFL_reads	83,203
FL_chimera_reads	1,058
FLNCs	301,082
FLNC_withoutPolyA_reads	570
FLNC_withPolyA_reads	300,512

Table 3 HQ isoforms by isoseq3.

Sample	A1
High quality isoforms	21,863
High quality isoforms total bases	59,311,425
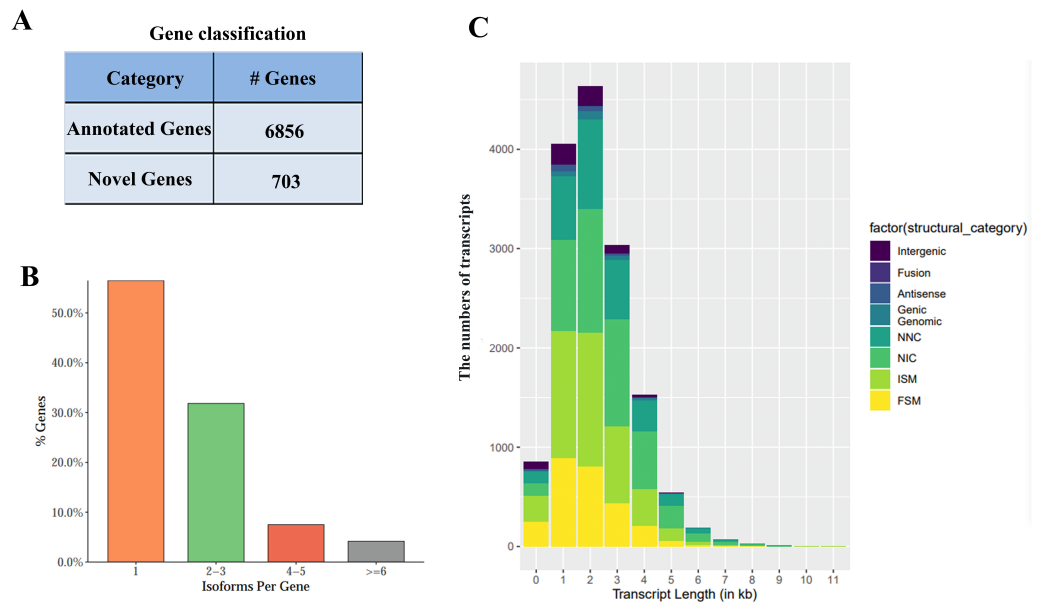
Figure 2 Iso-sequencing (Iso-Seq) data-based identification of genes together with isoforms. (A) Gene classification. (B) Transcript distribution. (C) Classifications by isoforms length.

Full-size ⬚ DOI: 10.7717/peerj.13446/fig-2

Altogether 7,559 gene loci (including 703 possible new and 6,856 identified ones) were discovered using Iso-Seq (Fig. 2A). About half of the genes had more than one splice (Fig. 2B). Those new gene loci as well as new isoforms were added into the reference annotation file (Table S2).

Isoforms were compared with the information of known transcripts according to the position and cutting information relative to the genome using SQANTI2. Isoforms were divided into FSM (full splice match), ISM (incomplete splice match), NIC (Novel in
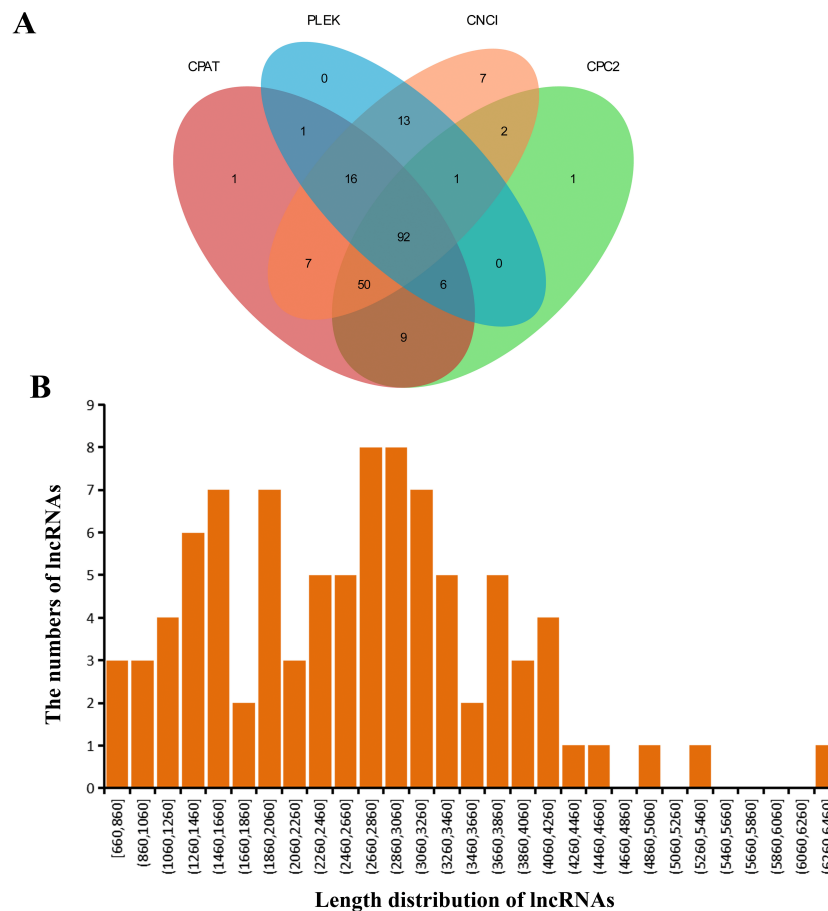
Figure 3 **Detected lncRNAs.** (A) Venn plot showing those lncRNAs detected *via* four software applications. (B) Predicted lncRNA lengths.

Full-size ◰ DOI: 10.7717/peerj.13446/fig-3

Catalog), NNC (Novel Not in Catalog), Antisense, Genic, Genomic, Fusion, and Intergenic subtypes. The length distribution of different types of isoforms was shown in Fig. 2C.

LncRNA prediction

Four softwares were adopted for lncRNA prediction in the GCs transcriptome. At last, we estimated 92 lncRNAs in total (Fig. 3A). Table S3 listed the predicted lncRNA sequences, whose lengths were 660–6,460 bp (average, 2,623 bp) (Fig. 3B).

Alternative polyadenylation events detection

This study detected APA events within porcine GCs using TAPIS pipeline. Among those genes identified 8,793 showed one or more poly(A) site (Fig. 4, Table S4), including 3,412 containing one single poly(A) site (Fig. 4), whereas 5,381 (61.20%) containing at least two poly(A) sites (Fig. 4).

Alternative splicing events detection

Based on Iso-Seq data, we identified 3,465 AS events in total. There were four types of AS events (Fig. 5), among which, exon skipping (1,162, 33.54%) showed the highest prevalence,
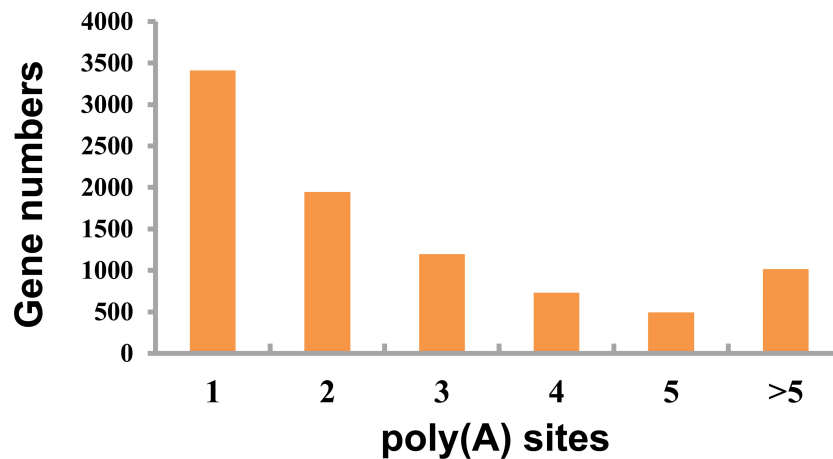
Figure 4 Detected APA events. Distribution of different numbers of poly(A) sites.
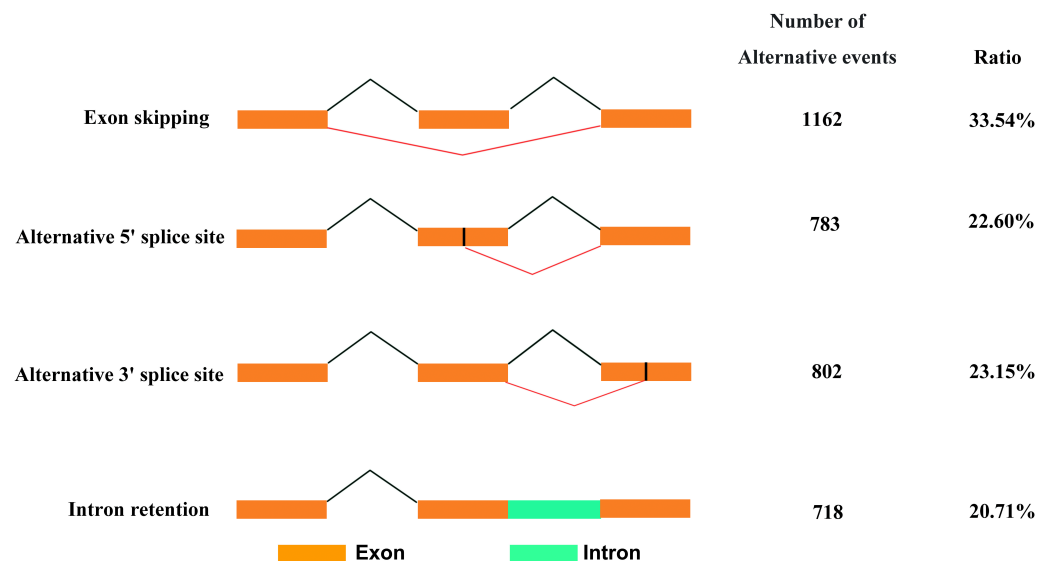
Figure 5 Classification of AS events.

while alternative 5′ (783, 22.60%), alternative 3′ (802, 23.15%) and intronretaining (718, 20.71%) ranked the second to fourth places, respectively.

Validation of alternative splicing genes

For confirming whether the Iso-Seq results were reliable, we screened four AS genes within porcine GCs through RT-PCR (Fig. 6). RT-PCR came to similar results to sequencing results, which indicated that the Iso-Seq data were reliable.
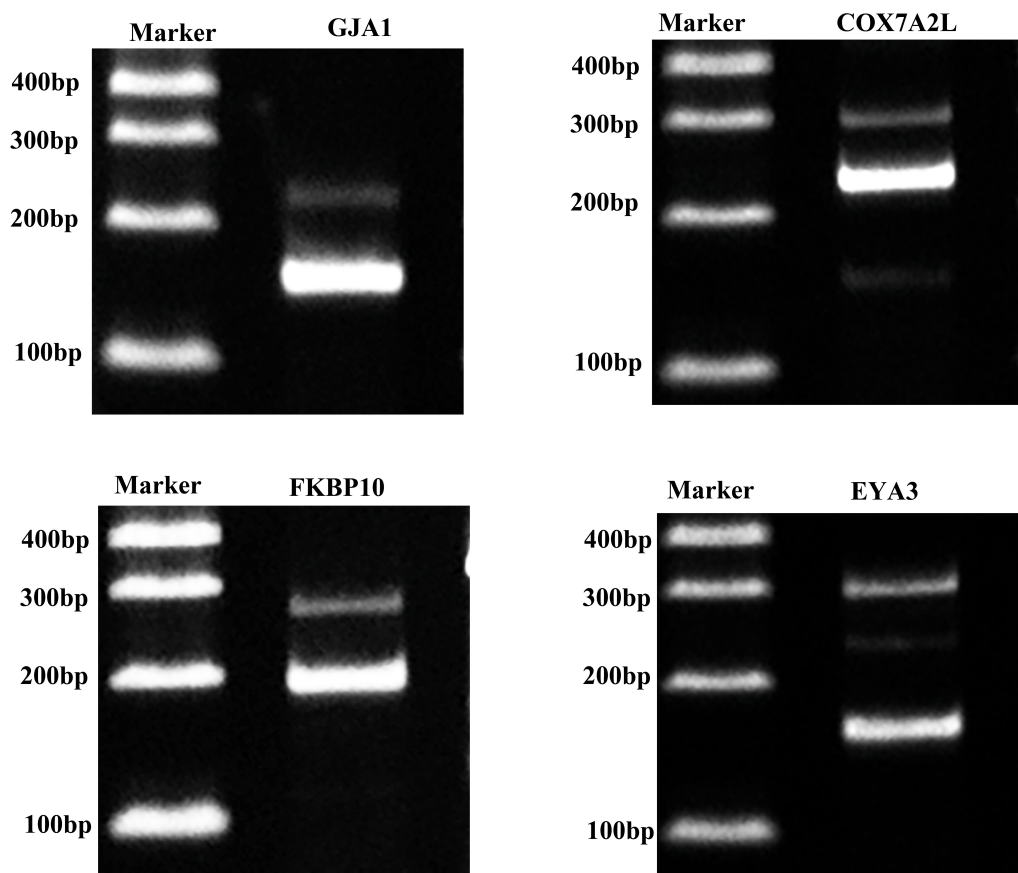
Fusion gene analysis

Gene fusion refers to the process where all or part of the sequences of two genes fuse with each other to form a new gene. The production of fusion gene may be the consequence of chromosome translocation, intermediate deletion or chromosome inversion, which is usually tumorigenic. Gene fusion is a common feature of tumor, which enhances cancer genesis and progression, and is adopted to be the diagnostic and therapeutic target against tumors. Using Fusion Finder, multiple genes were extracted in the alignment results, and the sequences with certain reads support were the candidate fusion genes. Later, 147 candidate fusion genes were annotated. The annotation results were displayed in Table 4 and Table S5.

Cluster analysis of isomers by GO

In order to further clarify the biological function of isoforms from Iso-Seq, cluster analysis was carried out by using GO. GO databases showed isoforms participated in 24 biological processes (Fig. 7, Table S6). Among them, 712 isoforms were related to reproductive

Table 4 Fusion gene annotation results (example).

Isoform	Position	Part1_transid	Part2_transid
PBfusio n.64	LUXX01048117.1:2433701-24 33871:-|LUXX01048117.1:298 9215-2989242:-	ENSSSCG00025040040_ENSSSC G00025039201	ENSSSCG00025040040
PBfusio n.113	LUXX01039762.1:1602380-16 02480:-|LUXX01052712.1:101 603-101739:-	ENSSSCG00025022795	ENSSSCG00025056743
PBfusio n.52	LUXX01021592.1:400524-400 768:+|LUXX01021592.1:72958 8-729843:+	ENSSSCG00025024992	ENSSSCG00025024992
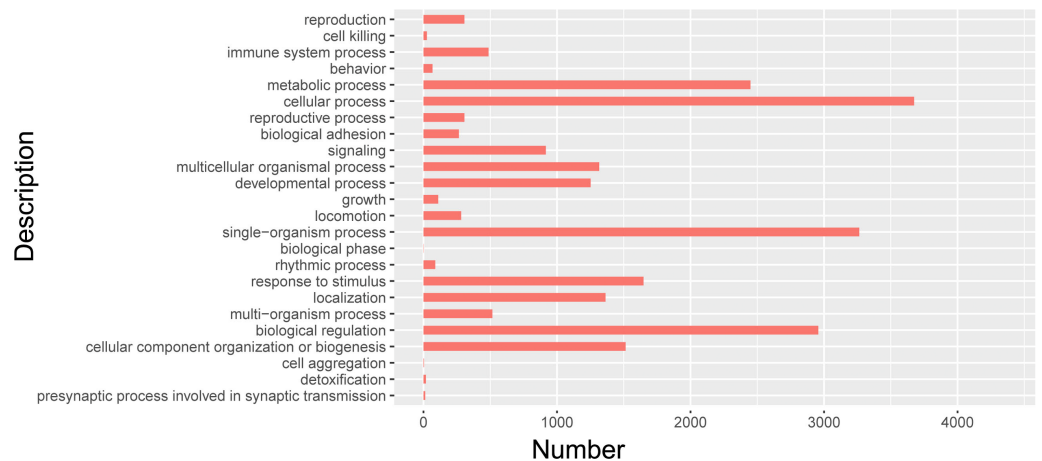
Figure 7 GO biological analysis of isoforms in pig GCs.

Full-size ☑ DOI: 10.7717/peerj.13446/fig-7

process. The most isoforms were enriched in cellular process. Detailed information of other biological processes was shown in Table S6.

DISCUSSION

The third generation sequencing technology is under rapid development in recent years, which has been applied to numerous research fields, such as genome resequencing (*Li et al., 2016*), *de-novo* genome assembly (*Jenjaroenpun et al., 2018*; *Morabito et al., 2020*), transcriptome research (*Mannarapu, Dariya & Bandapalli, 2021*; *Yuan et al., 2021*), methylation detection (*Liu et al., 2019a*; *Pillai, Gopalan & Lam, 2017*), disease-related structural variation detection (*Cretu Stancu et al., 2017*; *Roberts et al., 2021*), and virus analysis in epidemiology.

Previously, the second-generation sequencing was the most commonly used method to analyze gene molecular mechanism in pig (*Liu et al., 2019b*; *Tang et al., 2017*). Recently, third-generation sequencing has been increasingly employed in studying the functional genomics and epigenetics in pig (*Beiki et al., 2019*; *Ma et al., 2021*; *Zhang et al., 2020*). However, the AS and APA events in the pig ovary are rarely reported. In the present study, we revealed the transcriptomic complexity in porcine GCs using PacBio long-read

sequencing, resulting in 8,793 APA, 3,465 AS events, 703 candidate new gene loci, together with 92 lncRNAs.

LncRNAs can modulate gene levels within different biological events. Some studies have conducted to identify lncRNA functions within porcine GCs and ovary. *Liu et al. (2018)* identified altogether 2076 lncRNAs (including 714 novel and 1,362 known ones) from the libraries established based on Duroc ovaries. *Ruszkowska et al. (2018)* discovered 1,666 lncRNAs within porcine GCs. *Knapczyk-Stwora et al. (2019)* found altogether 5,592 RNA sequences to be lncRNAs, among which, 136 (at 53 long non-coding loci) were annotated within relevant databases among 11-day-old piglet ovaries. Meanwhile, the same authors discovered altogether 4,669 RNA sequences to be lncRNAs, among which, 1,236 (at 355 long non-coding loci) were annotated within relevant data-bases in pigs by illumina sequencing (*Knapczyk-Stwora et al., 2020*). In the current study, we found 92 lncRNAs from porcine GCs by PacBio sequencing, and such number was significantly lower than previously reported. We speculated that the reason might be that the third-generation sequencing method was more accurate in analyzing lncRNAs than the second-generation sequencing.

APA has certain influence on 3′-untranslated region (3′-UTR) composition and length, which also modulates the translation or stability of mRNA for affecting vital biological events. Wu and colleagues discovered the dynamic alterations in 3′-UTR landscape when oocytes became mature, which might be related to the modulation of porcine oocyte meiosis (*Wu et al., 2021*). As reported by *Deng et al. (2020)* APA had a certain effect on the rapid and slow muscle growth regulated by miRNAs as well as RNA binding proteins (RBPs). Polyadenylation site (PAS) is possibly related to immune response as well as androstenone contents within pigs (*Wang et al., 2016*). In this study, result showed that Npr3 (ENSSSCG00025035065) had 5 APA sites and could generate different mRNA isoforms in pig GCs. Previous reports detected luteinizing hormone could regulate the expression of Npr3 in mouse and bovine GCs (*Dos Santos et al., 2018*; *Lee et al., 2013*). However the function of different mRNA isoforms of Npr3 is needed to further research. The genes which were similar to Npr3 with multiple APA sites in porcine GCs, we found in total 5,381 genes. The APA sites of these genes greatly enriched the diversity of proteins in porcine GCs and provided favorable conditions for the function of GCs.

RNA splicing is a form of post-transcriptional modification, and gene AS is common in eukaryotes. With the application of new generation sequencing technology, recent studies on AS of porcine ovary suggest that 94.4–95.5% of the expressed genes in pig ovary are selectively spliced, and the frequency of AS events is similar in estrous and intersexual periods. Functional analysis of genes with AS events finds that many genes related to hormone metabolism and gonadal development are under different levels of modulation (*Tang et al., 2018*). In addition, the relationship between steroid hormones and AS events has gradually attracted the attention of researchers. Studies have shown that estradiol modulates class B scavenger receptor (sr-b) AS within liver cells by acting on the AS factors (*Zhang et al., 2007*). AS is closely related to emotional regulation in postmenopausal women (*Hou et al., 2018*). Different splice isomers of ERβ are linked with estradiol and RBP (*Shults et al., 2018*). Studies have shown that the splice isomer of

estrogen receptor ER-α 36 could promote the activity of tamoxifen agonist in glioblastoma cells (*Qu et al., 2019*). Previous studies showed that GJA1 was expressed in GCs and related to the proliferation of GCs in rat, human and bovine (*Chen et al., 2013*; *Lin et al., 2021*; *Rovani et al., 2014*). In this study the result showed that GJA1 had two different isoforms in pig GCs. However, whether their functions in the pig were similar with other species needed further research. FKBP10 could promote proliferation of glioma cells (*Cai et al., 2021*). The function and expression of FKBP10 had not been reported in GCs. The results of this study demonstrated that it expressed in GCs and had different splicing isomers. C1QTNF3 was revealed to have different transcripts in pig GCs in this study and also was reported to promote proliferation of granulosa cells and protect of granulosa cells from apoptosis in mouse and human (*Gershon & Dekel, 2020*). The previous reports almost are about AS, estradiol and related diseases or the function of genes. However, there are rarely reports regarding AS occurrence as well as the regulatory and action mechanisms of different splices in porcine GCs. Our results found altogether 3,465 AS events by Iso-Seq data. Such results were more accurate than the sequencing results of ovarian tissue, which could provide a basis for revealing the physiological functions such as steroid hormone synthesis in GCs.

Gene fusion can modify epigenetically, and the resultant products-encoded new proteins are related to carcinogenesis. But little research is conducted to analyze chimeric genes. Li and colleagues examined chimeric genes for their biological effects as well as related mechanisms by weighted co-expression network analysis (*Li, Li & Ma, 2021a*). *Liu et al. (2021)* firstly reported that the chimeric RNAs related mechanism in modulating skeletal muscle development in pigs. This work analyzed altogether 147 fusion genes in porcine GCs. This study offered a new direction for exploring chimeric genes' effect on GCs.

CONCLUSIONS

We employed PacBio Iso-Seq in this work to generate the integrative transcriptomic data for porcine GCs, which gave rise to 8,793 APA, 3,465 AS events, 703 possible new gene loci, together with 92 lncRNAs. This work illustrated that porcine GCs had a complex transcriptome and discovered numerous possible transcripts, thus facilitating to understand follicular development.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

Competing Interests

The authors declare there are no competing interests.

Author Contributions

- Shuxin Li performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Jiarui Wang performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Jiale Li performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Meihong Yue performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Chuncheng Liu analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Libing Ma analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Ying Liu conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.

Animal Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

the Institutional Animal Care and Use Committee of Inner Mongolia University of Science & Technology, School of Life Science and Technology (NMGKJDX-2019-10)

Data Availability

The following information was supplied regarding data availability:

The sequences are available at NCBI Sequence Read Archive: SRR17818022.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.13446#supplemental-information.

REFERENCES

Ardui S, Ameur A, Vermeesch JR, Hestand MS. 2018. Single molecule real-time (SMRT) sequencing comes of age: applications and utilities for medical diagnostics. *Nucleic Acids Research* **46**:2159–2168 DOI 10.1093/nar/gky066.

Baralle FE, Giudice J. 2017. Alternative splicing as a regulator of development and tissue identity. *Nature Reviews Molecular Cell Biology* **18**:437–451 DOI 10.1038/nrm.2017.27.

Beiki H, Liu H, Huang J, Manchanda N, Nonneman D, Smith TPL, Reecy JM, Tuggle CK. 2019. Improved annotation of the domestic pig genome through integration of Iso-Seq and RNA-seq data. *BMC Genomics* **20**:344.

Cai HQ, Zhang MJ, Cheng ZJ, Yu J, Yuan Q, Zhang J, Cai Y, Yang LY, Zhang Y. 2021. FKBP10 promotes proliferation of glioma cells via activating AKT-CREB-PCNA axis. *Journal of Biomedical Science* **28**:13 DOI 10.1186/s12929-020-00705-3.

Chen H, Zhao L, Chu G, Kito G, Yamauchi N, Shigeyoshi Y, Hashimoto S, Hattori MA. 2013. FSH induces the development of circadian clockwork in rat granulosa cells via a gap junction protein Cx43-dependent pathway. *American Journal of Physiology, Endocrinology and Metabolism* **304**:E566–E575 DOI 10.1152/ajpendo.00432.2012.

Cretu Stancu M, Van Roosmalen MJ, Renkens I, Nieboer MM, Middelkamp S, De Ligt J, Pregno G, Giachino D, Mandrile G. 2017. Mapping and phasing of structural variation in patient genomes using nanopore sequencing. *Nature Communications* **8**:1326 DOI 10.1038/s41467-017-01343-4.

Deng L, Li L, Zou C, Fang C, Li C. 2020. Characterization and functional analysis of polyadenylation sites in fast and slow muscles. *BioMed Research International* **2020**:2626584.

Di Giammartino DC, Nishida K, Manley JL. 2011. Mechanisms and consequences of alternative polyadenylation. *Molecular Cell* **43**:853–866 DOI 10.1016/j.molcel.2011.08.017.

Dos Santos JT, De Cesaro MP, Ferst JG, Pereira Dau AM, Da Rosa PRA, Pasqual BM, Antoniazzi AQ, Gasperin BG, Bordignon V, Goncalves PBD. 2018. Luteinizing hormone upregulates NPPC and downregulates NPR3 mRNA abundance in bovine granulosa cells through activation of the EGF receptor. *Theriogenology* **119**:28–34 DOI 10.1016/j.theriogenology.2018.06.012.

Du X, Li Q, Yang L, Zeng Q, Wang S, Li Q. 2021. Transcriptomic data analyses reveal that sow fertility-related lincRNA NORFA is essential for the normal states and functions of granulosa cells. *Frontiers in Cell and Developmental Biology* **9**:610553 DOI 10.3389/fcell.2021.610553.

Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P. 2009. Real-time DNA sequencing from single polymerase molecules. *Science* **323**:133–138 DOI 10.1126/science.1162986.

Gershon E, Dekel N. 2020. Newly identified regulators of ovarian folliculogenesis and ovulation. *International Journal of Molecular Sciences* **21**:4565.

Gruber AJ, Zavolan M. 2019. Alternative cleavage and polyadenylation in health and disease. *Nature Reviews Genetics* **20**:599–614 DOI 10.1038/s41576-019-0145-z.

Hoffmann F, Maser E. 2007. Carbonyl reductases and pluripotent hydroxysteroid dehydrogenases of the short-chain dehydrogenase/reductase superfamily. *Drug Metabolism Reviews* **39**:87–144 DOI 10.1080/03602530600969440.

Hou X, Adeosun SO, Zhao X, Hill R, Zheng B, Reddy R, Su X, Meyer J, Mosley T, Wang JM. 2018. ERbeta agonist alters RNA splicing factor expression and has a longer window of antidepressant effectiveness than estradiol after long-term ovariectomy. *Journal of Psychiatry and Neuroscience* **43**:170199.

Hu H, Fu Y, Zhou B, Li Z, Liu Z, Jia Q. 2021. Long non-coding RNA TCONS_00814106 regulates porcine granulosa cell proliferation and apoptosis by sponging miR-1343. *Molecular and Cellular Endocrinology* **520**:111064 DOI 10.1016/j.mce.2020.111064.

Jenjaroenpun P, Wongsurawat T, Pereira R, Patumcharoenpol P, Ussery DW, Nielsen J, Nookaew I. 2018. Complete genomic and transcriptional landscape analysis using third-generation sequencing: a case study of Saccharomyces cerevisiae CEN.PK113-7D. *Nucleic Acids Research* **46**:e38 DOI 10.1093/nar/gky014.

Kang YJ, Yang DC, Kong L, Hou M, Meng YQ, Wei L, Gao G. 2017. CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Research* **45**:W12–W16 DOI 10.1093/nar/gkx428.

Knapczyk-Stwora K, Nynca A, Ciereszko RE, Paukszto L, Jastrzebski JP, Czaja E, Witek P, Koziorowski M, Slomczynska M. 2019. Flutamide-induced alterations in transcriptional profiling of neonatal porcine ovaries. *Journal of Animal Science and Biotechnology* **10**:35 DOI 10.1186/s40104-019-0340-y.

Knapczyk-Stwora K, Nynca A, Ciereszko RE, Paukszto L, Jastrzebski JP, Czaja E, Witek P, Koziorowski M, Slomczynska M. 2020. Transcriptomic profiles of the ovaries from piglets neonatally exposed to 4-tert-octylphenol. *Theriogenology* **153**:102–111 DOI 10.1016/j.theriogenology.2020.04.027.

Korlach J, Bjornson KP, Chaudhuri BP, Cicero RL, Flusberg BA, Gray JJ, Holden D, Saxena R, Wegener J, Turner SW. 2010. Real-time DNA sequencing from single polymerase molecules. *Methods in Enzymology* **472**:431–455 DOI 10.1016/S0076-6879(10)72001-2.

Kulus J, Kulus M, Kranc W, Jopek K, Zdun M, Jozkowiak M, Jaskowski JM, Piotrowska-Kempisty H, Bukowska D. 2021. Transcriptomic profile of new gene markers encoding proteins responsible for structure of porcine ovarian granulosa cells. *Biology (Basel)* **10**:1214.

LaVoie HA. 2017. Transcriptional control of genes mediating ovarian follicular growth, differentiation, and steroidogenesis in pigs. *Molecular Reproduction and Development* **84**:788–801 DOI 10.1002/mrd.22827.

LaVoie HA, De Simone DC, Gillio-Meina C, Hui YY. 2002. Cloning and characterization of porcine ovarian estrogen receptor beta isoforms. *Biology of Reproduction* **66**:616–623 DOI 10.1095/biolreprod66.3.616.

Lee KB, Zhang M, Sugiura K, Wigglesworth K, Uliasz T, Jaffe LA, Eppig JJ. 2013. Hormonal coordination of natriuretic peptide type C and natriuretic peptide receptor 3 expression in mouse granulosa cells. *Biology of Reproduction* **88**:42.

Li Q, Du X, Wang L, Shi K, Li Q. 2021b. TGF-beta1 controls porcine granu-losa cell states: a miRNA-mRNA network view. *Theriogenology* **160**:50–60 DOI 10.1016/j.theriogenology.2020.11.001.

Li P, Li Y, Ma L. 2021a. Potential role of chimeric genes in pathway-related gene co-expression modules. *World Journal of Surgical Oncology* **19**:149 DOI 10.1186/s12957-021-02248-9.

Li G, Shen M, Le S, Tan Y, Li M, Zhao X, Shen W, Yang Y, Wang J. 2016. Genomic analyses of multidrug resistant Pseudomonas aeruginosa PA1 resequenced by single-molecule real-time sequencing. *Bioscience Reports* **36**:e00418.

Li A, Zhang J, Zhou Z. 2014. PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC Bioinformatics* **15**:311 DOI 10.1186/1471-2105-15-311.

Lin TC, Wang KH, Chuang KH, Kao AP, Kuo TC. 2021. Downregulation of gap junctional intercellular communication and connexin 43 expression by bisphenol A in human granulosa cells. *Biotechnology and Applied Biochemistry* **68**:676–682 DOI 10.1002/bab.1979.

Liu H, Begik O, Lucas MC, Ramirez JM, Mason CE, Wiener D, Schwartz S, Mattick JS, Smith MA, Novoa EM. 2019a. Accurate detection of m(6)A RNA modifications in native RNA sequences. *Nature Communications* **10**:4079 DOI 10.1038/s41467-019-11713-9.

Liu Y, Li M, Bo X, Li T, Ma L, Zhai T, Huang T. 2018. Systematic analysis of long non-coding RNAs and mRNAs in the ovaries of duroc pigs during different follicular stages using RNA sequencing. *International Journal of Molecular Sciences* **19**:1722.

Liu D, Xia J, Yang Z, Zhao X, Li J, Hao W, Yang X. 2021. Identification of Chimeric RNAs in Pig Skeletal Muscle and Transcriptomic Analysis of Chimeric RNA TNNI2-ACTA1 V1. *Frontiers in Veterinary Science* **8**:742593 DOI 10.3389/fvets.2021.742593.

Liu Y, Yang Y, Li W, Ao H, Zhang Y, Zhou R, Li K. 2019b. Effects of melatonin on the synthesis of estradiol and gene expression in pig granulosa cells. *Journal of Pineal Research* **66**:e12546 DOI 10.1111/jpi.12546.

Ma H, Jiang J, He J, Liu H, Han L, Gong Y, Li B, Yu Z, Tang S. 2021. Long-read assembly of the Chinese indigenous Ningxiang pig genome and identification of genetic variations in fat metabolism among different breeds. *Molecular Ecology Resources*.

Manabe N, Goto Y, Matsuda-Minehata F, Inoue N, Maeda A, Sakamaki K, Miyano T. 2004. Regulation mechanism of selective atresia in porcine follicles: regulation of granulosa cell apoptosis during atresia. *Journal of Reproduction and Development* **50**:493–514 DOI 10.1262/jrd.50.493.

Mannarapu M, Dariya B, Bandapalli OR. 2021. Application of single-cell sequenc-ing technologies in pancreatic cancer. *Molecular and Cellular Biochemistry* **476**:2429–2437 DOI 10.1007/s11010-021-04095-4.

Matsuda-Minehata F, Goto Y, Inoue N, Manabe N. 2005. Changes in expression of anti-apoptotic protein, cFLIP, in granulosa cells during follicular atresia in porcine ovaries. *Molecular Reproduction and Development* **72**:145–151 DOI 10.1002/mrd.20349.

McGee EA, Hsueh AJ. 2000. Initial and cyclic recruitment of ovarian follicles. *Endocrine Reviews* **21**:200–214.

Meng L, Zhao K, Wang CC, Tao J, Wu Z, Teerds K, Zhang S. 2021. Characterization of long non-coding RNA profiles in porcine granulosa cells of healthy and atretic antral follicles: implications for a potential role in apoptosis. *International Journal of Molecular Sciences* **22**:2677.

Morabito C, Aiese Cigliano R, Marechal E, Rebeille F, Amato A. 2020. Illumina and PacBio DNA sequencing data, de novo assembly and annotation of the genome of Aurantiochytrium limacinum strain CCAP_4062/1. *Data Brief* **31**:105729 DOI 10.1016/j.dib.2020.105729.

Park JY, Su YQ, Ariga M, Law E, Jin SL, Conti M. 2004. EGF-like growth factors as mediators of LH action in the ovulatory follicle. *Science* **303**:682–684 DOI 10.1126/science.1092463.

Pillai S, Gopalan V, Lam AK. 2017. Review of sequencing platforms and their applications in phaeochromocytoma and paragangliomas. *Critical Reviews in Oncology/Hematology* **116**:58–67 DOI 10.1016/j.critrevonc.2017.05.005.

Qu C, Ma J, Zhang Y, Han C, Huang L, Shen L, Li H, Wang X, Liu J, Zou W. 2019. Estrogen receptor variant ER-alpha36 promotes tamoxifen agonist activity in glioblastoma cells. *Cancer Science* **110**:221–234 DOI 10.1111/cas.13868.

Quirk SM, Cowan RG, Harman RM, Hu CL, Porter DA. 2004. Ovarian follicular growth and atresia: the relationship between cell proliferation and survival. *Journal of Animal Science* **82 E-Suppl**:E40–E52.

Roberts HE, Lopopolo M, Pagnamenta AT, Sharma E, Parkes D, Lonie L, Freeman C, Knight SJL, Lunter G. 2021. Short and long-read genome sequencing methodologies for somatic variant detection; genomic analysis of a patient with diffuse large B-cell lymphoma. *Scientific Reports* **11**:6408 DOI 10.1038/s41598-021-85354-8.

Robinson JW, Zhang M, Shuhaibar LC, Norris RP, Geerts A, Wunder F, Eppig JJ, Potter LR, Jaffe LA. 2012. Luteinizing hormone reduces the activity of the NPR2 guanylyl cyclase in mouse ovarian follicles, contributing to the cyclic GMP decrease that promotes resumption of meiosis in oocytes. *Developmental Biology* **366**:308–316 DOI 10.1016/j.ydbio.2012.04.019.

Rovani MT, Gasperin BG, Ilha GF, Ferreira R, Bohrer RC, Duggavathi R, Bordignon V, Goncalves PB. 2014. Expression and molecular consequences of inhibition of estrogen receptors in granulosa cells of bovine follicles. *Journal of Ovarian Research* **7**:96 DOI 10.1186/s13048-014-0096-0.

Ruszkowska M, Nynca A, Paukszto L, Sadowska A, Swigonska S, Orlowska K, Molcan T, Jastrzebski JP, Ciereszko RE. 2018. Identification and characterization of long non-coding RNAs in porcine granulosa cells exposed to 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin. *Journal of Animal Science and Biotechnology* **9**:72 DOI 10.1186/s40104-018-0288-3.

Shi M, Sirard MA. 2021. Cocultured porcine granulosa cells respond to excess non-esterified fatty acids during in vitro maturation. *Journal of Ovarian Research* **14**:142 DOI 10.1186/s13048-021-00904-y.

Shults CL, Dingwall CB, Kim CK, Pinceti E, Rao YS, Pak TR. 2018. 17beta-estradiol regulates the RNA-binding protein Nova1, which then regulates the alternative splicing of estrogen receptor beta in the aging female rat brain. *Neurobiol Aging* **61**:13–22 DOI 10.1016/j.neurobiolaging.2017.09.005.

Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, Liu Y, Chen R, Zhao Y. 2013. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Research* **41(17)**:e166 DOI 10.1093/nar/gkt646.

Tang LT, Ran XQ, Mao N, Zhang FP, Niu X, Ruan YQ, Yi FL, Li S, Wang JF. 2018. Analysis of alternative splicing events by RNA sequencing in the ovaries of Xiang pig at estrous and diestrous. *Theriogenology* **119**:60–68 DOI 10.1016/j.theriogenology.2018.06.022.

Tang Z, Wu Y, Yang Y, Yang YT, Wang Z, Yuan J, Yang Y, Hua C, Fan X. 2017. Comprehensive analysis of long non-coding RNAs highlights their spatio-temporal expression patterns and evolutional conservation in Sus scrofa. *Scientific Reports* **7**:43166 DOI 10.1038/srep43166.

Tardaguila M, de la Fuente L, Marti C, Pereira C, Pardo-Palacios FJ, Del Risco H, Ferrell M, Mellado M, Macchietto M. 2018. SQANTI: extensive characterization of long-read transcript sequences for quality control in full-length transcriptome identification and quantification. *Genome Research* **28(3)**:396–411 DOI 10.1101/gr.222976.117.

Tian Y, Zhang MY, Li N, Wang JJ, Ge W, Tan SJ, Shen W, Li L. 2020. Zearalenone exposure triggered porcine granulosa cells apoptosis via microRNAs-mediated focal adhesion pathway. *Toxicology Letters* **330**:80–89 DOI 10.1016/j.toxlet.2020.05.009.

Toms D, Pan B, Bai Y, Li J. 2021. Small RNA sequencing reveals distinct nuclear microRNAs in pig granulosa cells during ovarian follicle growth. *Journal of Ovarian Research* **14**:54 DOI 10.1186/s13048-021-00802-3.

Wang H, Li R, Zhou X, Xue L, Xu X, Liu B. 2016. Genome-wide analysis and functional characterization of the polyadenylation site in pigs using RNAseq data. *Scientific Reports* **6**:36388 DOI 10.1038/srep36388.

Wang L, Park HJ, Dasari S, Wang S, Kocher JP, Li W. 2013. CPAT: coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Research* **41**:e74 DOI 10.1093/nar/gkt006.

Wu ZW, Mou Q, Fang T, Wang Y, Liang H, Wang C, Du ZQ, Yang CX. 2021. Global 3′-untranslated region landscape mediated by alternative polyadenylation during meiotic maturation of pig oocytes. *Reproduction in Domestic Animals* **57(1)**:33–44 DOI 10.1111/rda.14026.

Xu G, Hu Y, Yu D, Chen X, Li X, Duan S, Zhang N, Xu G, Hu J. 2022. Discovery of differentially expressed MicroRNAs in porcine ovaries with smaller and larger litter size. *Frontiers in Genetics* **13**:762124 DOI 10.3389/fgene.2022.762124.

Yamochi T, Hashimoto S, Morimoto Y. 2021. Mural granulosa cells support to maintain the viability of growing porcine oocytes and its developmental competence after insemination. *Journal of Assisted Reproduction and Genetics* **38**:2591–2599 DOI 10.1007/s10815-021-02212-2.

Yang Y, Zhou R, Zhu S, Li X, Li H, Yu H, Li K. 2017. Systematic identification and molecular characteristics of long noncoding RNAs in pig tissues. *BioMed Research International* **2017**:6152582 DOI 10.1155/2017/6152582.

Yu C, Li M, Wang Y, Liu Y, Yan C, Pan J, Liu J, Cui S. 2016. MiR-375 mediates CRH signaling pathway in inhibiting E2 synthesis in porcine ovary. *Reproduction* **153**:63–73 DOI 10.1530/REP-16-0323.

Yuan Z, Ge L, Sun J, Zhang W, Wang S, Cao X, Sun W. 2021. Integrative analysis of Iso-Seq and RNA-seq data reveals transcriptome complexity and differentially expressed transcripts in sheep tail fat. *PeerJ* **9**:e12454 DOI 10.7717/peerj.12454.

Zhang Y, Cui Y, Zhang X, Wang Y, Gao J, Yu T, Lv X, Pan C. 2018. Pig StAR: mRNA expression and alternative splicing in testis and Leydig cells, and association analyses with testicular morphology traits. *Theriogenology* **118**:46–56 DOI 10.1016/j.theriogenology.2018.05.031.

Zhang M, Li Z, Li J, Huang T, Peng G, Tang W, Yi G, Zhang L, Song Y. 2020. Revisiting the pig IGHC gene locus in different breeds uncovers nine distinct IGHG genes. *Journal of Immunology* **205**:2137–2145 DOI 10.4049/jimmunol.1901483.

Zhang FL, Li N, Wang H, Ma JM, Shen W, Li L. 2019. Zearalenone exposure induces the apoptosis of porcine granulosa cells and changes long noncoding RNA expression to promote antiapoptosis by activating the JAK2-STAT3 pathway. *Journal of Agricultural and Food Chemistry* **67**:12117–12128 DOI 10.1021/acs.jafc.9b05189.

Zhang X, Moor AN, Merkler KA, Liu Q, McLean MP. 2007. Regulation of alternative splicing of liver scavenger receptor class B gene by estrogen and the involved regulatory splicing factors. *Endocrinology* **148**:5295–5304 DOI 10.1210/en.2007-0376.