ZOOLOGICAL RESEARCH

Genome-wide identification of imprinted genes in pigs and their different imprinting status compared with other mammals

DEAR EDITOR,

Genomic imprinting often results in parent-of-origin specific differential expression of maternally and paternally inherited alleles and plays an essential role in mammalian development and growth. Mammalian genomic imprinting has primarily been studied in mice and humans, with only limited information available for pigs. To systematically characterize this phenomenon and evaluate imprinting status between different species, we investigated imprinted genes on a genome-wide scale in pig brain tissues. Specifically, we performed bioinformatics analysis of high-throughput sequencing results from parental genomes and offspring transcriptomes of hybrid crosses between Duroc and Diannan small-ear pigs. We identified 11 paternally and five maternally expressed imprinted genes in pigs with highly stringent selection criteria. Additionally, we found that the *KCNQ1* and *IGF2R* genes, which are related to development, displayed a different imprinting status in pigs compared with that in mice and humans. This comprehensive research should help improve our knowledge on genomic imprinting in pigs and highlight the potential use of imprinted genes in the pig breeding field.

Genomic imprinting is a parent-of-origin-dependent phenomenon whereby only one of the two alleles originating from parents is expressed (McGrath & Solter, 1984; Surani et al., 1984). Genomic imprinting is regulated through epigenetic mechanisms, including DNA methylation, histone modifications, and non-coding RNAs (Grandjean et al., 2001; Inoue et al., 2017; Li et al., 1993; Sleutels et al., 2002). Interestingly, genomic imprinting exhibits unique species-specific expression patterns (Kalscheuer et al., 1993). In mice, for example, *IGF2R* (insulin-like growth factor 2 receptor) is regulated by a maternal differentially methylated region (DMR) (Stöger et al., 1993). The DMR can be inherited by the next generation and cause maternal allele expression, which influences fetal development and metabolic regulation (Stöger et al., 1993; Wutz et al., 1997). In humans, *IGF2R* is reported to be biallelically expressed (Kalscheuer et al., 1993). Pigs are an important domestic species and widely applied large animal model in medical research (Rubin et al., 2012; Yan et al., 2018). A paternally expressed *IGF2* gene in pigs is known to affect muscle growth, fat deposition, and heart size (Van Laere et al., 2003). However, to the best of our knowledge, few studies have applied next-generation sequencing to detect genomic imprinting in pigs at the genome-wide scale (Ahn et al., 2019; Oczkowicz et al., 2018). Most previous studies on pigs have surveyed the imprinting status of known imprinted genes identified in mice and used for genetic manipulation of pig embryos (Bischoff et al., 2009; Park et al., 2011). Genome-wide surveys for novel imprinted genes in pigs remain poorly studied. Furthermore, the similarities and differences in imprinting status between pigs and other mammals are unclear.

To analyze imprinted genes in pigs, we selected two distantly related pig strains to generate initial crosses (Duroc pig (male)×Diannan small-ear pig (female)) and reciprocal crosses (Duroc pig (female)×Diannan small-ear pig (male)). Experiments were approved by the Institutional Animal Care and Use Committee at the Kunming Institute of Zoology, Chinese Academy of Sciences (approval ID No.: SMKX-2017023). The identification of imprinted genes was described in detail in the Supplementary Materials and Methods. Ear tissue samples were collected from the parent animals and were used to extract DNA with a TIANamp Genomic DNA Kit

Received: 09 April 2020; Accepted: 03 August 2020; Online: 04 August 2020

Foundation items: This work was supported by the Ministry of Agriculture of China (2016ZX08009003-006), National Key R&D Program of China (2019YFA0110700), and Science & Technology Department of Yunnan Province (2019HA003), Animal Branch of the Germplasm Bank of Wild Species, Chinese Academy of Sciences (Large Research Infrastructure Funding)

DOI: 10.24272/j.issn.2095-8137.2020.072

(Tiangen Biotech, China). RNA from offspring brain tissue samples was isolated using a TaKaRa MiniBEST Universal RNA Extraction Kit (TaKaRa, China). Total RNA and genomic DNA quality was analyzed using a NanoDrop 2000 as well as agarose gel electrophoresis. The standard Illumina protocols were applied to construct libraries and sequences for DNA-seq and RNA-seq on the Illumina platform. To remove the influence of mapping bias, we generated 1 907 M paired-end DNA-seq parent reads from seven Duroc pig samples and 10 Diannan small-ear pig samples with an average sequencing depth of 8.91× to 13.16× (Supplementary Table S1). In total, 40 648 348 single nucleotide polymorphisms (SNPs) with at least one read supported between Diannan small-ear pigs and Duroc pigs were detected (Figure 1A). After low-quality SNP filtering using the Genome Analysis Toolkit (GATK) hard filter module (McKenna et al., 2010), 32 942 732 high-quality SNPs

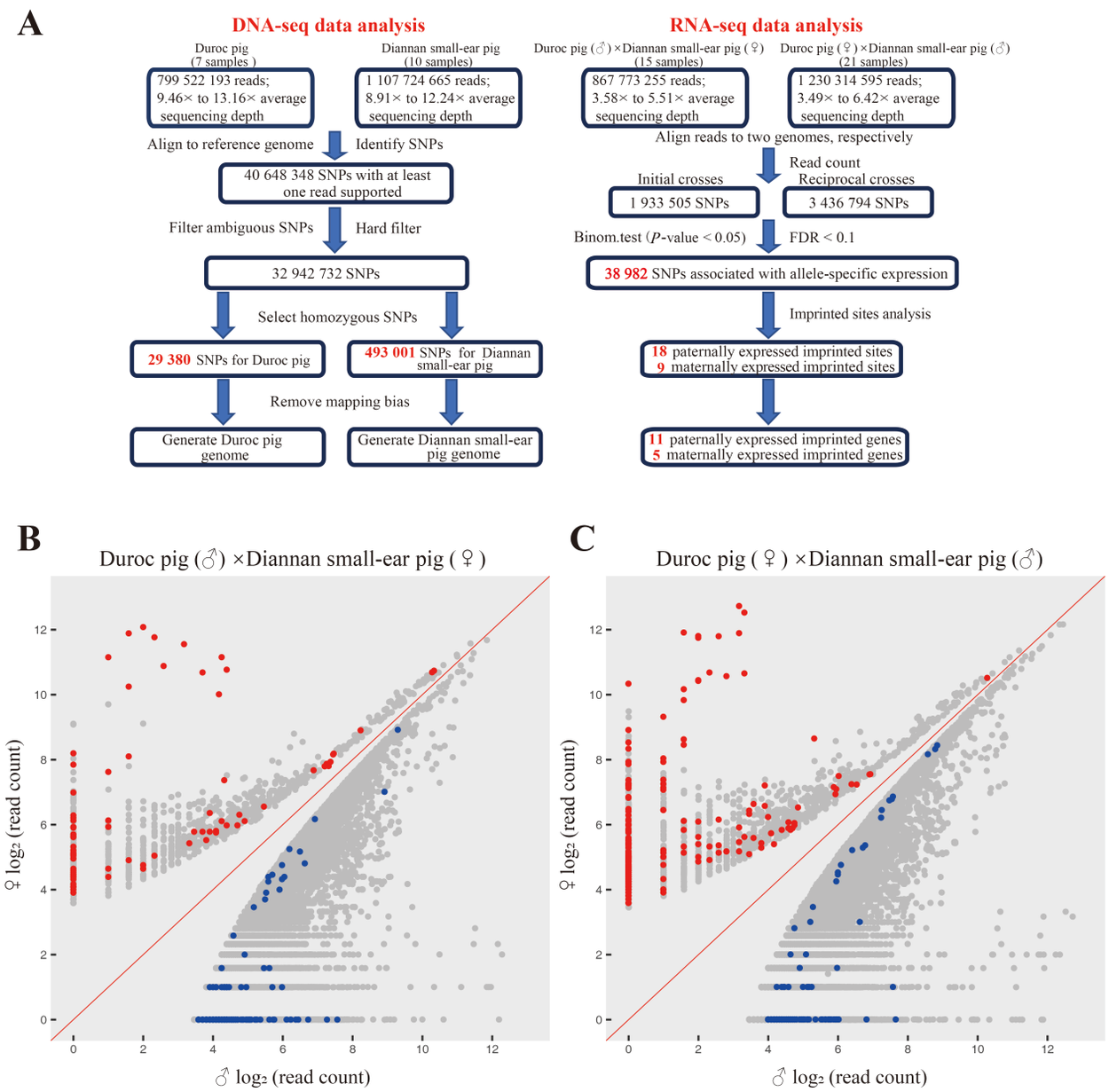

Figure 1 Genome-wide identification of imprinted genes in pigs

A: Pipeline for identification of pig imprinted genes. B, C: SNP sites associated with allele-specific expression in initial crosses (B) and reciprocal crosses (C). Gray dots are SNP sites associated with allele-specific expression. Red dots represent paternally imprinted sites. Blue dots represent maternally imprinted sites. Red lines denote 1:1 expression ratio between two alleles.

were retained (Figure 1A). We selected SNPs that were homozygous in each parent but differed between male and female parents as informative SNPs to distinguish the origin of SNPs. If the SNP site was heterozygous in one sample, the site was removed in subsequent analysis. If the genotype was different in one breed, the site was also excluded. Finally, 493 001 and 29 380 unique homozygous SNPs were used for generating the Diannan small-ear and Duroc pig genomes, respectively (Figure 1A).

Using 36 F1 offspring samples from the two types of hybrid crosses, we generated 2 098 M paired-end RNA-seq reads with an average sequencing depth of 3.49× to 6.42× (Supplementary Table S1), which were then aligned to the Diannan small-ear and Duroc pig genomes, respectively (Figure 1A). The correlations among RNA-seq samples were evaluated using Pearson correlation coefficients, which were calculated using multiBamSummary and plotCorrelation in deepTools (Ramírez et al., 2016) (Supplementary Figure S1). In total, 522 381 unique SNPs were detected in the parent DNA-seq data, which were then used to calculate the number of reads for each allele. Finally, 384 791 SNPs had more than one read supported by the RNA-seq data and were annotated, with 257 356 SNPs covering 9 871 genes (data not shown). The other 127 435 SNPs were located in intergenic regions (data not shown). Allele-specific expression was assayed, with significant deviation observed from the 1:1 expression ratio between the read count of two alleles. The Binom.test and false discovery rate (FDR) were used for F1 offspring RNA-seq data from the 15 initial crosses and 21 reciprocal crosses (Figure 1A). After filtering based on $P<0.05$ and FDR<0.1, 13 761 allele-specific expression sites in the initial crosses and 25 221 allele-specific expression sites in the reciprocal crosses (located in 1 775 genes) were detected in the 36 F1 offspring samples (Figure 1A–C; Supplementary Table S2).

To detect high-confidence imprinted genes, we required all allele-specific expression sites to show the same parent-biased expression direction in both the initial and reciprocal crosses. To remove the influence of random expression, the imprinted sites were required to have more than two supported samples in both the initial and reciprocal crosses. In total, 18 paternally expressed imprinted sites (covering 11 genes) and nine maternally expressed imprinted sites (covering five genes) were detected (Figure 1A; Table 1 and Supplementary Table S3). Interestingly, of the 16 imprinted genes detected, most have not been reported in any species in previous genomic imprinting studies. The known imprinted genes included *IGF2R* (Barlow et al., 1991), *GNAS* (Hayward et al., 1998), *NNA*T (Kagitani et al., 1997), and *KCNQ1* (Lee et al., 1997). *IGF2R* was one of the first imprinted genes identified in mice, and plays an important role in biological functions such as fetal growth and placental function (Barlow et al., 1991; Owens, 1991), with *IGF2R* knockout mice found to exhibit fetal overgrowth or late gestational lethality (Lau et al., 1994). In addition, *KCNQ1* is an important maternally expressed imprinted gene in mice and humans and is involved in fetal development, as well as type 2 diabetes susceptibility (Gould & Pfeifer, 1998; Yasuda et al., 2008). The newly identified imprinted genes included *KBTBD6*, *ZNF791*, *ZNF709*, *JPH3* and *NOB1* et al. (Table 1; Supplementary Table S3). *KBTBD6* (KELCH repeat and BTB domain containing 6) is known to interact with the human GABARAP subfamily of ATG8 family members in a LC3-interacting region (LIR)-dependent manner (Genau et al., 2015). Current research indicates that Zinc Finger Protein 791 (*ZNF791*) plays a critical role in female mitotic phase fetal germ cells (Li et al., 2017). *ZNF709* is a member of the zinc finger family and

Table 1 Details on 16 imprinted genes detected in pigs

Ensembl ID	Gene symbol	Expressed allele in mammals		
		Human	Mouse	Pig
ENSSSCG00000039556	*KCNQ1*	Maternally	Maternally	Paternally
ENSSSCG00000004044	*IGF2R*	Biallelically	Maternally	Paternally
ENSSSCG00000007336	*NNAT*	Paternally	Paternally	Paternally
ENSSSCG00000007520	*GNAS*	Maternally	Maternally	Maternally
ENSSSCG00000031378	*KBTBD6*	N/A	N/A	Paternally
ENSSSCG00000029347	*ZNF791*	N/A	N/A	Paternally
ENSSSCG00000002753	*NOB1*	N/A	N/A	Maternally
ENSSSCG00000013717	*ZNF709*	N/A	N/A	Paternally
ENSSSCG00000014838	*PGM2L1*	N/A	N/A	Maternally
ENSSSCG00000002653	*JPH3*	N/A	N/A	Paternally
ENSSSCG00000022177	*DIS3L2*	N/A	N/A	Paternally
ENSSSCG00000036033	*THRB*	N/A	N/A	Paternally
ENSSSCG00000037530	*TACC2*	N/A	N/A	Paternally
ENSSSCG00000025243	*SGIP1*	N/A	N/A	Paternally
ENSSSCG00000048719	N/A	N/A	N/A	Maternally
ENSSSCG00000051274	N/A	N/A	N/A	Maternally

N/A: Not available.

its knockdown in human cells leads to increased expression of p53 (Yan et al., 2016). *JPH3* (Junctophilin 3) is a novel tumor suppressor gene methylated in colorectal and gastric tumors, promoting mitochondrial-mediated apoptosis, and is also a potential metastasis and survival biomarker for digestive cancers (Hu et al., 2017). Taken together, our method reliably identified imprinted genes on a genome-wide scale. Further studies and experimental validation of these genes should provide new information on genomic imprinting in pigs. In addition, imprinted genes could be a new class of gene for application in pig breeding.

In general, imprinting status is constant within a species and is conserved among different species (Thorvaldsen & Bartolomei, 2007). Interestingly, *KCNQ1* is a maternally expressed imprinted gene in mice and humans (Gould & Pfeifer, 1998), but was paternally expressed in pigs in our data. Previous research has shown the detection of *KCNQ1* genomic imprinting to be non-informative in pigs (Bischoff et al., 2009). Our study is the first to report on *KCNQ1* as a paternally expressed imprinted gene in pigs. Specifically, for *KCNQ1*, all eight allele-specific expression sites showed paternally expressed imprinting status in the offspring of the initial and reciprocal crosses (Supplementary Table S2 and S3). In addition, *IGF2R* was paternally expressed in 19 brain tissue samples at a precise allele-specific expression site (Supplementary Table S2 and S3), with imprinting status differing from that reported in previous studies on pigs (Bischoff et al., 2009; Braunschweig, 2012; Killian et al., 2001; Shen et al., 2012). The paternally expressed imprinting status of *IGF2R* in pigs also differed from that found in mice and humans (Table 1). Thus, further studies are needed to analyze the biological significance of the different imprinting statuses between different species.

In total, we identified 11 paternally and five maternally expressed imprinted genes in the pig genome, which is currently the most comprehensive analysis of imprinted genes in pigs. We also found that *KCNQ1* and *IGF2R* displayed a different imprinting status in pigs compared to that in mice and humans. This study highlights the potential use of imprinted genes within the pig breeding field.

DATA AVAILABILITY

The DNA-seq and RNA-seq datasets used in this study were submitted to the Genome Sequence Archive (GSA) with ID CRA001638.

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

Y.P.Z. and Z.Y.Z. initiated the project. Z.Y.Z. designed the study. Y.Q.W. performed data analysis and interpretation. H.Z., Y.J.L., and H.W. collected the samples. Z.Y.Z. and Y.Q.W. wrote the manuscript. S.K. revised the manuscript. All authors read and approved the final version of the manuscript.

Yin-Qiao Wu[1,2,#], Heng Zhao[3,#], Ying-Ju Li[1,2], Saber Khederzadeh[1,2], Hong-Jiang Wei[3], Zhong-Yin Zhou[1,*], Ya-Ping Zhang[1,4,*]

[1] *State Key Laboratory of Genetic Resources and Evolution*, *and Yunnan Laboratory of Molecular Biology of Domestic Animals*, *Kunming Institute of Zoology*, *Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650223, *China*

[2] *Kunming College of Life Science*, *University of Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650204, *China*

[3] *Key Laboratory of Animal Gene Editing and Animal Cloning in Yunnan Province*, *Yunnan Agricultural University*, *Kunming*, *Yunnan* 650201, *China*

[4] *Center for Excellence in Animal Evolution and Genetics*, *Chinese Academy of Sciences*, *Kunming*, *Yunnan* 650223, *China*

[#]Authors contributed equally to this work

*Corresponding authors, E-mail: zhouzhongyin@mail.kiz.ac.cn; zhangyp@mail.kiz.ac.cn

REFERENCES

Ahn B, Choi MK, Yum J, Cho IC, Kim JH, Park C. 2019. Analysis of allele-specific expression using RNA-seq of the Korean native pig and Landrace reciprocal cross. *Asian-Australasian Journal of Animal Sciences*, **32**(12): 1816−1825.

Barlow DP, Stöger R, Herrmann BG, Saito K, Schweifer N. 1991. The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the *Tme* locus. *Nature*, **349**(6304): 84−87.

Bischoff SR, Tsai S, Hardison N, Motsinger-Reif AA, Freking BA, Nonneman D, et al. 2009. Characterization of conserved and nonconserved imprinted genes in swine. *Biology of Reproduction*, **81**(5): 906−920.

Braunschweig MH. 2012. Biallelic transcription of the porcine *IGF2R* gene. *Gene*, **500**(2): 181−185.

Genau HM, Huber J, Baschieri F, Akutsu M, Dötsch V, Farhan H, et al. 2015. CUL3-KBTBD6/KBTBD7 ubiquitin ligase cooperates with GABARAP proteins to spatially restrict TIAM1-RAC1 signaling. *Molecular Cell*, **57**(6): 995−1010.

Gould TD, Pfeifer K. 1998. Imprinting of mouse *Kvlqt1* is developmentally regulated. *Human Molecular Genetics*, **7**(3): 483−487.

Grandjean V, O'Neill L, Sado T, Turner B, Ferguson-Smith A. 2001. Relationship between DNA methylation, histone H4 acetylation and gene expression in the mouse imprinted *Igf2*‐*H19* domain. *FEBS Letters*, **488**(3): 165−169.

Hayward BE, Moran V, Strain L, Bonthron DT. 1998. Bidirectional imprinting of a single gene: *GNAS1* encodes maternally, paternally, and biallelically derived proteins. *Proceedings of the National Academy of Sciences of the United States of America*, **95**(26): 15475−15480.

Hu XT, Kuang YY, Li LL, Tang HM, Shi QL, Shu XS, et al. 2017. Epigenomic and functional characterization of junctophilin 3 (JPH3) as a

novel tumor suppressor being frequently inactivated by promoter CpG methylation in digestive cancers. *Theranostics*, **7**(7): 2150−2163.

Inoue A, Jiang L, Lu FL, Suzuki T, Zhang Y. 2017. Maternal H3K27me3 controls DNA methylation-independent imprinting. *Nature*, **547**(7664): 419−424.

Kagitani F, Kuroiwa Y, Wakana S, Shiroishi T, Miyoshi N, Kobayashi S, et al. 1997. *Peg5*/*Neuronatin* is an imprinted gene located on sub-distal chromosome 2 in the mouse. *Nucleic Acids Research*, **25**(17): 3428−3432.

Kalscheuer VM, Mariman ECM, Schepens MT, Rehder H, Ropers HH. 1993. The insulin-like growth factor type-2 receptor gene is imprinted in the mouse but not in humans. *Nature Genetics*, **5**(1): 74−78.

Killian JK, Nolan CM, Wylie AA, Li T, Vu TH, Hoffman AR, et al. 2001. Divergent evolution in *M6P/IGF2R* imprinting from the Jurassic to the Quaternary. *Human Molecular Genetics*, **10**(17): 1721−1728.

Lau MM, Stewart CE, Liu Z, Bhatt H, Rotwein P, Stewart CL. 1994. Loss of the imprinted IGF2/cation-independent mannose 6-phosphate receptor results in fetal overgrowth and perinatal lethality. *Genes & Development*, **8**(24): 2953−2963.

Lee MP, Hu RJH, Johnson LA, Feinberg AP. 1997. Human *KVLQT1* gene shows tissue-specific imprinting and encompasses Beckwith-Wiedemann syndrome chromosomal rearrangements. *Nature Genetics*, **15**(2): 181−185.

Li E, Beard C, Jaenisch R. 1993. Role for DNA methylation in genomic imprinting. *Nature*, **366**(6453): 362−365.

Li L, Dong J, Yan LY, Yong J, Liu XX, Hu YQ, et al. 2017. Single-Cell RNA-Seq analysis maps development of human germline cells and gonadal niche interactions. *Cell Stem Cell*, **20**(6): 891−892.

McGrath J, Solter D. 1984. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell*, **37**(1): 179−183.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky AM, et al. 2010. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**(9): 1297−1303.

Oczkowicz M, Szmatola T, Piórkowska K, Ropka-Molik K. 2018. Variant calling from RNA-seq data of the brain transcriptome of pigs and its application for allele-specific expression and imprinting analysis. *Gene*, **641**: 367−375.

Owens JA. 1991. Endocrine and substrate control of fetal growth: placental and maternal influences and insulin-like growth factors. *Reproduction, Fertility and Development*, **3**(5): 501−517.

Park CH, Uh KJ, Mulligan BP, Jeung EB, Hyun SH, Shin T, et al. 2011. Analysis of imprinted gene expression in normal fertilized and uniparental preimplantation porcine embryos. *PLoS One*, **6**(7): e22216.

Ramírez F, Ryan D, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*, **44**(W1): W160−W165.

Rubin CJ, Megens HJ, Barrio AM, Maqbool K, Sayyab S, Schwochow D, et al. 2012. Strong signatures of selection in the domestic pig genome. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(48): 19529−19536.

Shen CJ, Cheng WTK, Wu SC, Chen HL, Tsai TC, Yang SH, et al. 2012. Differential differences in methylation status of putative imprinted genes among cloned swine genomes. *PLoS One*, **7**(2): e32812.

Sleutels F, Zwart R, Barlow DP. 2002. The non-coding *Air* RNA is required for silencing autosomal imprinted genes. *Nature*, **415**(6873): 810−813.

Stöger R, Kubicka P, Liu CG, Kafri T, Razin A, Cedar H, et al. 1993. Maternal-specific methylation of the imprinted mouse *Igf2r* locus identifies the expressed locus as carrying the imprinting signal. *Cell*, **73**(1): 61−71.

Surani MAH, Barton SC, Norris ML. 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature*, **308**(5959): 548−550.

Thorvaldsen JL, Bartolomei MS. 2007. SnapShot: imprinted gene clusters. *Cell*, **130**(5): 958.e1−958.e2.

Van Laere AS, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, et al. 2003. A regulatory mutation in *IGF2* causes a major QTL effect on muscle growth in the pig. *Nature*, **425**(6960): 832−836.

Wutz A, Smrzka OW, Schweifer N, Schellander K, Wagner EF, Barlow DP. 1997. Imprinted expression of the *Igf2r* gene depends on an intronic CpG island. *Nature*, **389**(6652): 745−749.

Yan S, Tu ZC, Liu ZM, Fan NN, Yang HM, Yang S, et al. 2018. A huntingtin knockin pig model recapitulates features of selective neurodegeneration in huntington's disease. *Cell*, **173**(4): 989−1002.e13.

Yan WS, Scoumanne A, Jung YS, Xu ES, Zhang J, Zhang YH, et al. 2016. Mice deficient in poly(C)-binding protein 4 are susceptible to spontaneous tumors through increased expression of ZFP871 that targets p53 for degradation. *Genes & Development*, **30**(5): 522−534.

Yasuda K, Miyake K, Horikawa Y, Hara K, Osawa H, Furuta H, et al. 2008. Variants in *KCNQ1* are associated with susceptibility to type 2 diabetes mellitus. *Nature Genetics*, **40**(9): 1092−1097.