Base editing in pigs for precision breeding

Ruigao SONG1,2, Yu WANG3, Yanfang WANG (✉1), Jianguo ZHAO (✉1,2)

1 State Key Laboratory of Stem Cell and Reproductive Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China
2 Medical School, University of Chinese Academy of Sciences, Beijing 100049, China
3 State Key Laboratory of Animal Nutrition, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100081, China

Abstract  Pigs are one of the most important domesticated animals and have great value in agriculture and biomedicine. Single nucleotide polymorphisms (SNPs) are a dominant type of genetic variation among individual pigs and contribute to the formation of traits. Precision single base substitution provides a strategy for accurate genetic improvement in pig production with the characterization of functional SNPs and genetic variants in pigs. Base editing has recently been developed as the latest gene-editing tool that can directly make changes in single nucleotides without introducing double-stranded DNA breaks (DSBs), providing a promising solution for precise genetic modification in large animals. This review summarizes gene-editing developments and highlights recent genetic dissection related to SNPs in major economic traits which may have the potential to be modified using SNP-editing applications. In addition, limitations and future directions of base editing in pig breeding are discussed.

Keywords  base editing, genetic improvement, pigs, single nucleotide polymorphisms

1 Introduction

Pigs (Sus scrofa) are very important livestock animals that provide large quantities of meat worldwide[1]. About 730 pig breeds have been developed globally since domestication via natural and artificial selection[2,3]. Cosmopolitan lean breeds have primarily been raised in the pig industry over the past few decades, focusing on maximizing productivity and production efficiency and their traits have been significantly improved by established selection and breeding practices[4]. Notably, conventional selection and breeding still have two major vulnerabilities, namely slow genetic progress and difficulty in separating desired from undesirable traits[5].

Gene-editing technologies provide a promising platform for accelerating the breeding process in pigs in these circumstances. Early in the 1980s, the gene fragment of recombinant human growth hormone factor was originally introduced to the porcine genome by pronuclear microinjection in fertilized eggs and further expressed to improve livestock growth performance[6]. However, pronuclear injection was inefficient and the random integration of foreign gene fragments in the genome often results in unexpected consequences. In 1997, the emergence of somatic cell nuclear transfer (SCNT) technology allowed the generation of gene-targeted pigs using in vitro cultured fibroblasts that were genetically modified by homologous recombination (HR)[7,8]. Although HR was commonly used for genetic modification in model organisms the efficiency was quite low (around 10−3[9]) in somatic cells for generating precisely modified gene-targeted pigs until the development of novel gene editing tools. First, the emergence of zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs) has significantly increased the efficiency of gene editing in many species including pigs[10,11]. Using ZFNs, transgenic GFP alleles were first deleted and then a heterozygous mutation in the Ppar-γ gene was induced in pigs at a percentage of 4%–5%(6,12). Myostatin (MSTN) gene was deleted via ZFN in pigs to improve the quality of meat[13]. Shortly afterwards, clustered regularly interspaced short palindromic repeat (CRISPR)-associated (Cas) nucleases were developed and quickly applied in larger animals[14–16]. Due to scalability, affordability, and engineering flexibility, the innovative genome editing tool CRISPR/Cas9 system has created a paradigm shift in genetic modifications in large animals. Whitworth et al. prepared CD163-knockout pigs by injecting Cas9 mRNA and a single guide RNA (sgRNA) into pronuclear fertilized eggs, generating pigs that were resistant to porcine reproductive and respiratory syndrome virus (PRRSV) infection[17,18]. Xiang et al. prepared a genetically modified pig with increased growth rates by pronuclear injection of Cas9 nickase mRNA and a pair of sgRNAs to target...
intron3 of the *IGF2* gene\[^{19}\]. Using a CRISPR/Cas9-mediated HR-independent approach, Zheng et al. successfully knocked in the *mUCP1* gene into the porcine genome and the resulting pigs exhibited decreased fat deposition and improved thermoregulation during acute cold exposure\[^{20}\]. Harnessing RNAi technology, Xie et al. inserted an antiviral small hairpin RNA (shRNA) expression sequence at the porcine Rosa26 site using CRISPR/Cas9 to confer resistance to classical swine fever to pigs\[^{21}\]. These studies have dramatically improved the spectrum for making genetic modifications in pigs and successfully enhanced desired traits in pigs. However, precision gene editing at single base level is still challenging and there are advantages to simulating natural point mutations over knockout or transgenic strategies, which is urgently needed for possible incorporation into pig breeding.

### 2 Genetic basis of economic traits in pigs

SNPs are the richest and most abundant form of genomic polymorphisms, providing highly favorable markers for genetic map construction and whole genome-wide association studies (GWAS) to understand the genetic architecture of pig economic traits. GWAS analyses were conducted with a large scale SNP data set to dissect important genetic factors controlling traits of interest\[^{22}\].

At present, SNPs have been used to investigate the domestication and evolution of pigs\[^{23–25}\] and to identify functional SNP in genes related to various economic traits such as meat quality and growth traits\[^{26}\], reproduction\[^{27}\] and virus resistance\[^{28}\]. These efforts make it possible to manipulate the pig SNPs for pig breeding improvement at precise level by current gene editing tools and also to study the genetic mechanisms of economic traits in pigs\[^{29,30}\].

#### 2.1 SNPs are responsible for meat quality and growth

Reducing backfat thickness and increasing lean meat content is an important goal in the pig breeding process. However, excessive reduction in backfat thickness also leads to a decrease in intramuscular fat (IMF) and this does not satisfy consumer demand for high-quality meat\[^{31,32}\]. Scientists have taken a special interest in identifying the genes responsible for the formation of IMF and meat quality to optimize the pigs breeding to address the dilemma between meat production and quality. Genes associated with IMF have been identified and provide opportunities for genomic selection in pigs\[^{33}\]. The SNPAY183428 c.265T \(>\) C in fatty acid synthase (FASN) gene has given the most consistent results affecting backfat fatty acid composition (FAC) of large white pigs\[^{34}\]. Also SNPs in the RXRB gene have been identified to have the strongest association with oleic and monounsaturated fatty acid contents which have a major impact on fat composition in Iberian pigs\[^{35}\]. Moreover, two non-synonymous variants (I199V and T30N) in the *PRKAG3* gene have been associated with 24-h pH (pH24), drip loss (DL), protein content (PRO), cook yield (CY) \((P < 0.004)\), juiciness, tenderness (TEN) and shear force \((P < 0.004)\)\[^{36}\]. In addition, the polymorphism *IGF2* intron3-g.3072G \(>\) A has been reported to be the causal mutation for stimulating muscle growth which has a key role in the regulation of *IGF2* gene expression and FAC in the adipose tissue of pigs\[^{37,38}\]. Deep sequencing of PHKG1 revealed a point mutation \((C > A)\) in a splice acceptor site causing low meat quality in pig skeletal muscle\[^{39}\]. A c.892G \(>\) A mutation in MC4R has been associated with fatness and feed intake in the pig, and this mutation was also evaluated as a selection target for daily gain in Hampshire, Duroc, Landrace and Yorkshire pigs\[^{40}\]. In addition, the HAL-1843 (C1843T) mutation responsible for meat quality has been eliminated by most pig genetics companies from their herds and can determine the predisposition to porcine stress syndrome (PSS) in pigs\[^{41}\].

#### 2.2 SNPs responsible for reproduction

Reproductive traits are closely associated with production efficiency and economic profits. A number studies have identified candidate genes related to reproductive traits\[^{42}\]. Currently, a total of 2412 QTLs have been found on different pig chromosomes for endocrine, litter trait, reproductive organ and reproductive traits\[^{43}\]. From these QTLs, large numbers of SNPs and genes were identified associated with reproductive performance in pigs. It has been reported that 14 genes (BHLHA15, OCM2, IL1B2, GCK, SMAD2, HABP2, PAQR5, GRB10, PRELID2, DMKN, GPI, GPHBPI, ADCY2, and ACVR2B) were identified to be important in swine reproductive traits but still need further investigation\[^{44}\]. A study also found that the non-synonymous mutations in the AHR gene were associated with increased litter size in multiple European commercial lines\[^{45}\]. One study reported that sows homozygous with the A/A genotype in *SOD1* conceived three piglets more than sows with the A/T genotype on average, making this SNP a possible marker for increasing the litter size\[^{46}\].

#### 2.3 SNPs responsible for disease resistance

Genomic prediction of porcine response to different diseases would be very valuable to the pig industry. An SNP responsible for viral load (VL) and weight gain (WG) was discovered in 2014 this SNP, WUR10001215 (WUR), was shown to capture 99.3% of the genetic variance (GV) found in infection trial data of pigs infected with PRRSV (NVSL 97-7985)\[^{47–49}\]. Furthermore, the WUR SNP was shown to be associated with VL for two PRRSV isolates, NVSL-97-7895 (NVSL) and KS-2006-72109 (KS06)\[^{28}\]. Between the two isolates, genetic correlations for WG and
VL were both estimated at 0.86, indicating a high possibility of accurate genomic prediction\textsuperscript{[28]}. Serão et al.\textsuperscript{[50]} showed that moderate prediction accuracies for PRRSV antibody response were obtained using the SNPs located within the two major QTL on the \textit{Sus scrofa} chromosome 7 (SSC7). SRCR domain 5 of CD163 was found to be essential for successful infection with PRRSV\textsuperscript{[51]} and precision editing this domain conferred on the pigs the ability to resist PRRSV\textsuperscript{[18]}. In addition, a previous study identified two SNPs (rs55618716, ST) that were associated with fecal egg count (FEC) ($P < 0.01$), indicating resistance to \textit{Trichurus} in pigs\textsuperscript{[52]}.

We discuss only briefly genetic variation such as SNPs in three major traits in pigs as discussed above (Table 1). Most SNPs also induce only minor changes in phenotypic, physiological and biochemical characteristics. Thus, the identification of functional single base polymorphisms in genes with large effects on the phenotype which can be used for precise breeding still need further verification. Reverse genetic strategy is therefore necessary to measure or confirm the function of SNPs. Gene editing tools are therefore expected to play a key role in both genetic improvement by targeted genetic variation and also in the study of pig genome annotation.

<table>
<thead>
<tr>
<th>Functional SNP</th>
<th>Gene</th>
<th>Economic trait</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.265T &gt; C\textsuperscript{[34]}</td>
<td>FASN</td>
<td>Backfat fatty acid composition</td>
</tr>
<tr>
<td>c.2573T &gt; C\textsuperscript{[53]}</td>
<td>MTTP</td>
<td>Backfat fatty acid composition</td>
</tr>
<tr>
<td>g.G3072A\textsuperscript{[84]}</td>
<td>IGFB2</td>
<td>Backfat thickness</td>
</tr>
<tr>
<td>c.555C &gt; T\textsuperscript{[55]}</td>
<td>CTNNBL1</td>
<td>Backfat traits</td>
</tr>
<tr>
<td>c.892G &gt; A\textsuperscript{[40]}</td>
<td>MC4R</td>
<td>Fatness</td>
</tr>
<tr>
<td>c.205G &gt; A\textsuperscript{[56]}</td>
<td>SLC39A7</td>
<td>Carcass traits</td>
</tr>
<tr>
<td>g.15G &gt; A\textsuperscript{[87]}</td>
<td>CTSK</td>
<td>Backfat thickness</td>
</tr>
<tr>
<td>g.233C &gt; T\textsuperscript{[58]}</td>
<td>CRH</td>
<td>Growth and body composition</td>
</tr>
<tr>
<td>c.131T &gt; A\textsuperscript{[59]}</td>
<td>APOA2</td>
<td>Fatty acid composition</td>
</tr>
<tr>
<td>c.1199V\textsuperscript{[36]}</td>
<td>PRKAG3</td>
<td>Meat quality</td>
</tr>
<tr>
<td>c.720N\textsuperscript{[86]}</td>
<td>PRKAG3</td>
<td>Meat quality</td>
</tr>
<tr>
<td>g.8227C &gt; G\textsuperscript{[60]}</td>
<td>MUC4</td>
<td>Production traits</td>
</tr>
<tr>
<td>c.C2604T\textsuperscript{[61]}</td>
<td>PIK3C3</td>
<td>Production traits</td>
</tr>
<tr>
<td>g.A53G\textsuperscript{[62]}</td>
<td>IGFBP3</td>
<td>Litter size</td>
</tr>
<tr>
<td>g.35547A &gt; G\textsuperscript{[63]}</td>
<td>ESR2</td>
<td>Sperm motility</td>
</tr>
<tr>
<td>g.158 A &gt; G\textsuperscript{[84]}</td>
<td>PLCz</td>
<td>Sperm concentration</td>
</tr>
<tr>
<td>g.358A &gt; T\textsuperscript{[65]}</td>
<td>CD9</td>
<td>Sperm motility</td>
</tr>
<tr>
<td>g.C7462G\textsuperscript{[66]}</td>
<td>CXCL12</td>
<td>Pseudorabies virus disease resistance</td>
</tr>
<tr>
<td>c.12164 + 79G &gt; A\textsuperscript{[66]}</td>
<td>BAT2</td>
<td>Immunological traits</td>
</tr>
<tr>
<td>c.86172 + 140C &gt; T\textsuperscript{[86]}</td>
<td>Mx1</td>
<td>Immunological traits</td>
</tr>
<tr>
<td>g.G443A\textsuperscript{[67]}</td>
<td>TAP1</td>
<td>PRRSV resistance</td>
</tr>
<tr>
<td>c.933A &gt; G\textsuperscript{[68]}</td>
<td>TLR3</td>
<td>PRRSV susceptibility</td>
</tr>
<tr>
<td>c.761A &gt; G\textsuperscript{[69]}</td>
<td>IRF7</td>
<td>Health and immunity</td>
</tr>
<tr>
<td>g.2115T &gt; G\textsuperscript{[70]}</td>
<td>LMP2</td>
<td>Haematological traits</td>
</tr>
<tr>
<td>g.1232C &gt; G\textsuperscript{[70]}</td>
<td>LMP7</td>
<td>Haematological traits</td>
</tr>
<tr>
<td>c.C522T\textsuperscript{[71]}</td>
<td>BPI</td>
<td>Disease resistance</td>
</tr>
<tr>
<td>c.A1060G\textsuperscript{[73]}</td>
<td>BPI</td>
<td>Disease resistance</td>
</tr>
<tr>
<td>c.C1027A\textsuperscript{[72]}</td>
<td>TLR4</td>
<td>\textit{Salmonella} shedding</td>
</tr>
<tr>
<td>c.8C &gt; G\textsuperscript{[73]}</td>
<td>PSMB6</td>
<td>Immunological traits</td>
</tr>
<tr>
<td>c.144T &gt; C\textsuperscript{[74]}</td>
<td>BCL10</td>
<td>Immunological traits</td>
</tr>
</tbody>
</table>
3 Base editing

3.1 Conventional base editing by homology-directed repair is time-consuming and has low efficiency

Many strategies and tools have been tried to develop novel and efficient methods for single base induction or substitution in large animals over recent decades. Of these, CRISPR/Cas9 is an efficient and convenient gene editing technology that induces double-stranded DNA breaks (DSBs) for base editing. DSBs may be repaired by cellular homology-directed repair (HDR) that uses a donor DNA template such as introduced single-stranded donor oligonucleotides (ssODNs) or a double-stranded DNA that encodes the target-point nucleotide flanked by sequences homologous to the regions upstream and downstream of the DSB. The repair results in the knock-in of specific point mutations. Although CRISPR/Cas9 is used extensively to make precise insertions, deletions or any point mutation of interest, a number of limitations are attributed to HDR editing. HDR remains inefficient (typically ~ 0.1%–5%) because editing is restricted to the G2 and S phases of the cell cycle and is often accompanied by additional small insertions or deletions (indels), thus impeding the use of HDR for precise genome editing.

In addition, DSBs created by CRISPR/Cas9 often result in translocations, indels and rearrangements, and this impacts the efficiency of single base editing. These factors prevent the widespread use of CRISPR/Cas9 for livestock breeding for the introduction of SNP mutations.

3.2 Base editors enable direct base substitution without DSB

Base editors developed by David Liu’s group are a breakthrough in gene editing and enable direct generation of precise point mutations in genomic DNA without generating DSBs or requiring a donor template. Base editors are composed of fusion proteins that include catalytically impaired Cas nucleases, laboratory-evolved nucleobase deaminases, base-modified deaminases that operate only on single-stranded DNA, and other proteins such as uracil glycosylase inhibitor (UGI) that help to preserve the resulting single-nucleotide change. Two types of base editor tools are currently available, cytidine base editors (CBEs) that convert the C-G base pair into T-A and adenine base editors (ABEs) that convert A-T to G-C. These editors can collectively mediate the targeted installation of all four transition mutations (C-to-T, G-to-A, A-to-G, and T-to-C).

In 2016, Komor et al. first reported that the CBE system could efficiently convert cytidines within an editing window of about five nucleotides and correct a variety of point mutations with minimum indel formation. Many evolved base editors have recently been explored. The fourth-generation base editors (BE4 and SaBE4) with two UGI can increase the efficiency of C-G to T-A base editing while decreasing the frequency of undesired by-products compared to BE3. Further evolutions yielded BE4max and A3A-PBE which have promoted the efficiency of base editing by adding nuclear localization signal (NLS) or replacing rAPOBEC to APOBEC3A. YEE-BE3 and BE3-PAPAPAP were explored to narrow the editing window to 1–2 nt to reduce the bystander effect. However, dCas9 and nCas9 still follow the NGG principle and this restrains the editing scope.

The ABE system was also explored in David Liu’s laboratory. This system combined adenine deamination and nCas9 which can convert A-T to G-C with approximately 50% efficiency, at least 99.9% purity and no more than a 0.1% indel rate in human cells. The editing window of this ABE system is 4–9 nt. In 2019, Huang et al. broadened the targeting scope of CP1249-ABEmax to 4–12 nt. Furthermore, the PAM-modified Cas9 variants (VQR-SpCas9 (PAM:NGA), VRQR-SpCas9 (PAM:NGA), SaCas9 (PAM:NNRRT), ScCas9 (PAM:NNGN), xCas9 (PAM:NG) and SpCas9-NG (PAM:NG) were exploited to expand applications of the ABE system. With access to these two base editors, CBE and ABE are able to introduce all four transition mutations without requiring a double-strand DNA break. Most recently, a new cutting-edge technique, a catalytically impaired Cas9 fused engineered reverse transcriptase, was reported that showed high editing efficiency programmed with prime editing gRNA that both encodes the desired edit and specifies the target site, in addition to efficiency, prime editing also showed an expanded scope, greater capabilities and much less off-target byproducts than other gene editing tools.

3.3 Base editor mediated precise genetic modifications

A number of studies showing successful base substitution in pigs using base editors illustrates their feasibility in pig breeding. A study reported that base editor mediated GGTATA, B4galNT2, and CMAH modification enables a porcine pericardium with reduced immunogenicity but comparable physical characteristics and collagen composition compared with the wild-type porcine pericardium, providing a promising alternative for bioprosthetic heart valves. Recently, Yuan et al. successfully introduced C-to-T and C-to-G mutations in GGTATA, B4galNT2, and CMAH loci in porcine blastocysts at an efficiency of 66.7%–71.4%, significantly higher than the editing efficiency of 3.7% using CRISPR/Cas9. Furthermore, a study reported that CBEs efficiently induced C-to-T...
conversions of triple genes simultaneously, including RAG1, RAG2, and IL2RG, or DMD, TYR, and LMNA. These findings will help to accelerate the generation of animal models with multiplex point mutations and studies in gene therapies of genetic diseases\cite{101}. The pig model with RAG1, RAG2, and IL2RG mutations lacked B cells, T cells, and NK cells, providing a great prospect for xenotransplantation. In addition, Li et al. were able to knock out the TWIST2 or TYR genes in pigs to simulate human ablepharon macrostomia syndrome (AMS) or oculocutaneous albinism type 1 (OCA1) disease by introducing premature stop codons via third-generation base editor BE3\cite{102}. The resulting TWIST2 mutant pigs showed the expected phenotypes with absent eyelids, microtia, macrostomia, hypotrichosis, and abnormal trotters, while TYR mutant pigs showed typical albinism phenotypes and completely lost the dark pigment in their skin and hair. These results suggest that base editors provide a more simple and efficient method for single nucleotide editing that may be used to improve traits, provide disease resistance and accelerate the breeding process in pigs. Base editors are also convenient tools that can provide advantages in gene pyramiding that may be used to more rapidly breed for multiple economic traits than conventional breeding and selection methods (Fig. 1).

4 Conclusions and future perspectives

Base-editing technology currently shows great potential in model creation and future potential for precision breeding. Functional SNPs with phenotypic effects that may be modified by base editors for the purpose of genetic improvement in pig breeding can be expected. However, the number of SNPs affecting the economic traits with potential breeding value are limited. Thus, considerable efforts are required to accelerate the deciphering of the underlying genetic mechanism of pig body composition traits and disease resistance. Although there are no such published papers so far, base editing might be used for genome-wide screening to identify novel genes that are associated with economic traits in the future.

Although CBE and ABE have already been demonstrated to be efficient and precise for making point mutations in the genome of a wide variety of species\cite{84,103}, only two of the six possible base-pair transitions can be achieved by these editors and this limits their applications. Further, CBE and ABE still account for unexpected off-target editing, making it difficult to distinguish if a point mutation has been accurately engineered\cite{94,104}. In this regard, a detailed analysis of off-target editing efficiency of base editors is needed, and
potential biological consequences of off-target mutations should be assessed. Nevertheless, unwanted editing by-products observed in the editing of model organisms might not be a crucial problem in pig breeding, and by-products can be eliminated by dilution through individual pig mating. More powerful techniques with higher editing efficiency should still therefore be explored to bypass difficulties encountered in the production of gene-edited animals. Prime editing may be a new way to promise greater precision for base-edited pigs.

In summary, the optimization of base editors for higher precision and specificity coupled with the annotation of the genetic basis for desired traits will provide solutions for more efficient and precise pig breeding. Furthermore, animals gene edited by base editor only in the endogenous genome subtly without introducing foreign DNA may be viewed and regulated differently to currently genetically modified organisms (GMOs).

Acknowledgements This work was supported by the National Natural Science Foundation of China (81671274, 31925036, 31272440, and 31801031), the National Transgenic Project of China (2016ZX08009003-006-007), and the Elite Youth Program of the Chinese Academy of Agricultural Sciences (ASTIP-IAS05).

Compliance with ethics guidelines Ruigao Song, Yu Wang, Yanfang Wang, and Jianguo Zhao declare that they have no conflicts of interest or financial conflicts to disclose.

This article is a review and does not contain any studies with human or animal subjects performed by any of the authors.

References


Neurodegeneration, 2019, 14(1): 17


