

Sus scrofa domesticus reveals the genetic evolution of adaptive traits during early divergence

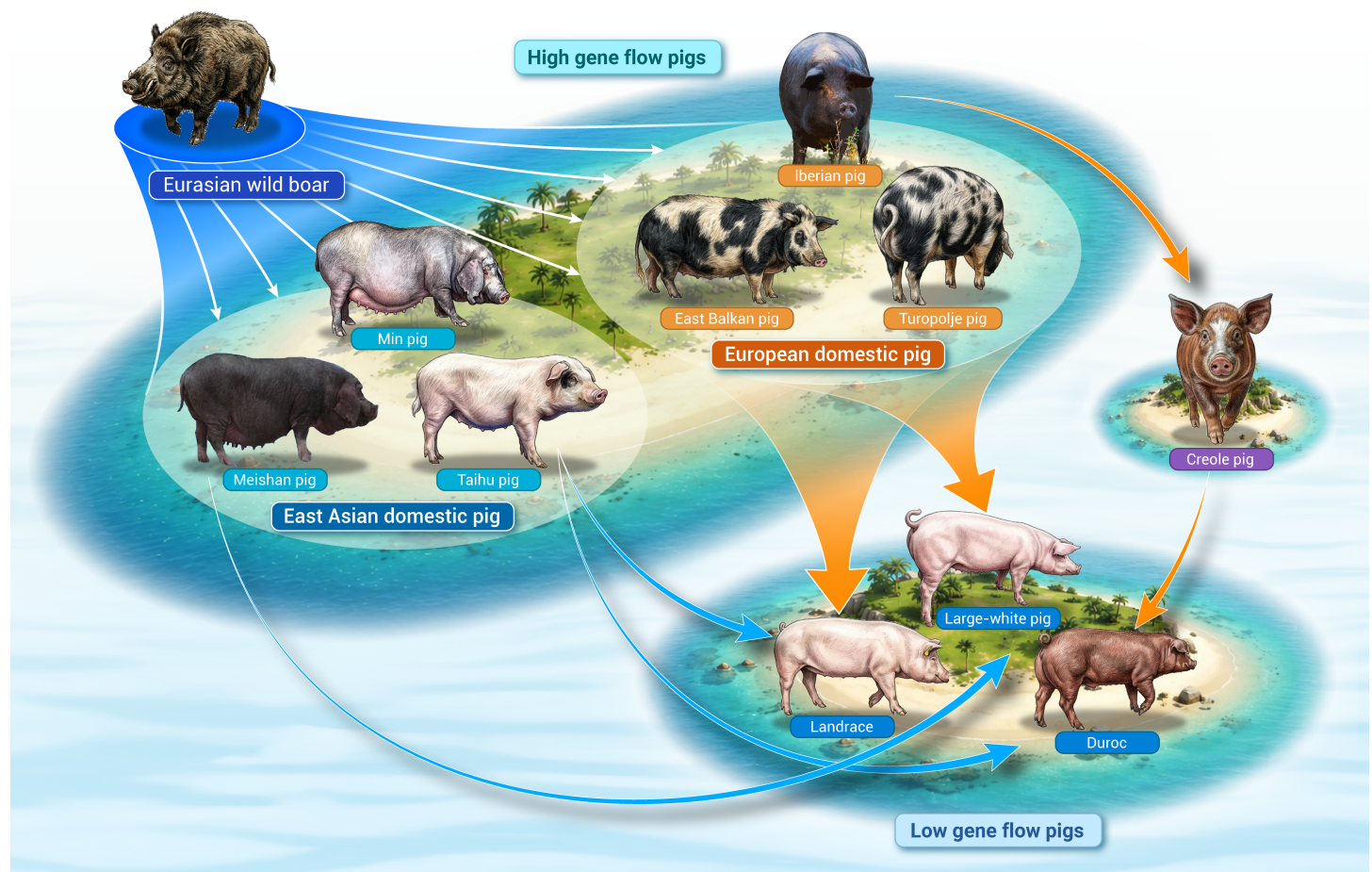
Bohan Rong,^{1,2,6,7,*} Naiqi Niu,^{2,7} Wencheng Zong,^{2,7} Run Zhang,² Xiaomei Ma,¹ Na Zhang,¹ Zhentong Shen,³ Yu Pang,^{1,6} Xu Lin,^{1,6} Di Liu,^{4,*} Yulong Yin,^{5,*} Longchao Zhang,^{2,*} and Xiuqin Yang^{1,6,*}

*Correspondence: xiuqinyang@neau.edu.cn (X.Y.); zhlichias@163.com (L.Z.); yinyulong@isa.ac.cn (Y.Y.); liudi1963@163.com (D.L.); rbh19901116@163.com (B.R.)

Received: December 15, 2025; Accepted: May 22, 2026; Published Online: May 26, 2026; <https://doi.org/10.59717/j.xinn-life.2026.100222>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

GRAPHICAL ABSTRACT



PUBLIC SUMMARY

- Domestic pigs diverge from wild boars yet remain interfertile, serving as a model for early speciation.
- Pigs split into high and low gene flow groups shaped by introgression and selection.
- The profile of bile acid shift improves glucose balance and metabolic health under energy-rich domestic diets.
- Regulatory changes precede coding alterations, showing traits evolve before reproductive isolation.

Sus scrofa domesticus reveals the genetic evolution of adaptive traits during early divergence

Bohan Rong,^{1,2,6,7,*} Naiqi Niu,^{2,7} Wencheng Zong,^{2,7} Run Zhang,² Xiaomei Ma,¹ Na Zhang,¹ Zhentong Shen,³ Yu Pang,^{1,6} Xu Lin,^{1,6} Di Liu,^{4,*} Yulong Yin,^{5,*} Longchao Zhang,^{2,*} and Xiuqin Yang^{1,6,*}

¹College of Animal Science and Technology, Northeast Agricultural University, Harbin 150030, China

²State Key Laboratory of Animal Biotech Breeding, Institute of Animal Sciences, Chinese Academy of Agricultural Sciences (CAAS), Beijing 100193, China

³College of Life Science, Nankai University, Tianjin 300071, China

⁴Institute of Animal Husbandry, Heilongjiang Academy of Agricultural Sciences, Harbin 150086, China

⁵Key Laboratory of Agro-ecological Processes in Subtropical Region, Institute of Subtropical Agriculture, Chinese Academy of Sciences, Changsha 410125, China

⁶Engineering Research Center for Intelligent Breeding and Farming of Pig in Northern Cold Region, Harbin 150030, China

⁷These authors contributed equally

*Correspondence: xiuqinyang@neau.edu.cn (X.Y.); zhichias@163.com (L.Z.); yinyulong@isa.ac.cn (Y.Y.); liudi1963@163.com (D.L.); rbh19901116@163.com (B.R.)

Received: December 15, 2025; Accepted: May 22, 2026; Published Online: May 26, 2026; <https://doi.org/10.59717/j.xinn-life.2026.100222>

© 2026 The Author(s). This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

Citation: Rong B., Niu N., Zong W., et al. (2026). *Sus scrofa domesticus* reveals the genetic evolution of adaptive traits during early divergence. *The Innovation Life* 4:100222.

Domestication has generated extensive phenotypic diversity within *Sus scrofa*, producing marked divergence from wild boars in morphology, behavior, and metabolic physiology across numerous distinct breeds. Yet domestic pigs and wild boars remain interfertile, as effective reproductive barriers have not emerged. This condition, phenotypic divergence without reproductive isolation, makes domestic pigs an informative model for studying genetic processes during the earliest phases of speciation, capturing the tension between selection-driven innovation and gene flow—maintained continuity. Here, we categorize domestic pig breeds by historical gene flow intensity with wild boars (high vs. low) and examine their contrasting genetic architectures, assessing how environmental adaptation and artificial selection shape genomic homogeneity and heterogeneity. Using bile acid (BA) metabolism as an illustrative example, we show that domestic pigs experienced a functional shift in which cytochrome P450 family 8 subfamily B member 1 (CYP8B1) silencing and 4 subfamily A member 21 (CYP4A21) neofunctionalization replaced cholic acid (CA) with hyocholic acid (HCA) as the primary BA. This transition suggests that phenotypic diversification can precede reproductive isolation and that evolutionary changes in regulatory regions may occur before alterations in coding sequences.

INTRODUCTION

The genetic mechanisms underlying phenotypic divergence before reproductive isolation is established remain incompletely characterized.^{1–3} Most traditional model systems, including *Drosophila* and mice, involve lineages that are already fully or nearly reproductively isolated, limiting insight into how selection, gene flow, and genomic restructuring interact during initial divergence.^{4,5} Domestic pigs offer a contrasting system: over roughly 10,000 years of domestication, they have diverged from wild boars in morphology, metabolism, and behavior, yet bidirectional gene flow persists between the two groups.^{6–10} This coexistence of differentiation and genetic exchange enables investigation of how targeted selection shapes genetic architecture while ongoing gene flow constrains divergence — conditions likely resembling the earliest stages of natural speciation.

Compared with cattle, sheep, and chickens, where domestication entailed stronger reproductive separation, *Sus scrofa domesticus* retains considerable genetic continuity with wild boars despite clear phenotypic divergence.^{11,12} Furthermore, unlike dogs, selected intensely for discrete morphological traits, pigs predominantly exhibit polygenic quantitative variation, such as differences in vertebral number, more closely paralleling gradual trait evolution in natural speciation.^{8,12–14} The global distribution of pig breeds, spanning Chinese indigenous populations, European local varieties, and American Criollo pigs, together with their substantial genetic diversity, provides a powerful framework for dissecting how geographic isolation, environmental adaptation, and human-mediated selection collectively shape genomic architecture.^{8,15,16} However, the domestication history of pigs remains elusive due to phenotypic plasticity, persistent gene flow and the lack of reproductive isolation from wild boars.^{16–22} As a result, although traditional

morphological and isotopic approaches have been widely used, comparative genomics and high-throughput sequencing are now indispensable for resolving evolutionary lineages and adaptive processes in *Sus scrofa*.^{17,20,23–27}

Early models of pig domestication assumed strict reproductive isolation between wild and domestic populations.²⁰ Archaeological evidence and whole-genome sequencing have since revealed a far more complex history of multiple domestication centers and extensive subspecific hybridization.^{18,21,28} Gene flow between wild boars and domestic pigs was not restricted to initial domestication but has continued throughout their shared evolutionary history, shaping the genetic composition of both.^{6,9,11,29,30} The resulting reservoir of ancestral polymorphisms enabled rapid responses to artificial selection, facilitating ecotype diversification within a relatively short timeframe,¹⁰ *Sus scrofa* therefore offers an exceptional opportunity to examine how standing variation generates phenotypic novelty during early divergence.

Beyond morphology, domestication has substantially reshaped the metabolic physiology of *Sus scrofa domesticus* relative to wild boars.^{8,31–35} Despite renewed interest from pan-genomic studies, the molecular basis of these metabolic differences remains poorly resolved. BA metabolism offers a prominent example: adult domestic pigs produce HCA (3 α ,6 α ,7 α -trihydroxy-5 β -cholan-24-oic acid) as their primary BA, whereas wild boars additionally retain CA (3 α ,5 β ,7 α -trihydroxycholan-24-oic acid) production.^{34,36–41} Growing evidence linking HCA to glucolipid metabolism suggests this shift may represent an adaptive response to the stable, energy-rich diets characteristic of domestication, likely reflecting functional replacement between two cytochrome P450 enzymes involving both neofunctionalization and noncoding regulatory change.^{36,39,42–47}

In this review, we use domestic pigs as a model to examine how selection intensity and gene flow shape trait genetic architecture prior to reproductive isolation. Focusing on BA metabolism in *Sus scrofa*, we highlight a pattern in which regulatory modifications precede coding sequence changes, illustrating how functional gene replacement drives phenotypic innovation during early lineage divergence.

GENETIC REMODELING OF *SUS SCROFA* UNDER DOMESTICATION: INSIGHTS INTO THE EARLY STAGES OF DIVERGENCE

Although pig domestication spans only ~10,000 years, a brief interval relative to the ~20-million-year evolutionary history of Suidae, strong human-mediated selection has nevertheless produced rapid and pronounced phenotypic divergence from wild boars.^{9,10,22} Speciation generally begins when a single ancestral population splits into independently evolving lineages, a gradual process shaped by natural selection, genetic drift, recombination, and mutation.^{16,18,48} Mutations generate the initial variation required for divergence; over successive generations, selection increases the frequency of advantageous alleles, producing lineage-specific haplotypes (Figure 1) that can spread through gene flow.¹⁸ Whether these haplotypes persist depends on local conditions, beneficial variants may be maintained or amplified, others lost, translating genetic variation into phenotypic diversity across heterogeneous environments.²⁴



Figure 1. Global distribution and genetic diversity of domestic pigs > 500 local breeds: Human-mediated selection during *Sus scrofa* domestication has been driven by socio-economic objectives, resulting in the enrichment of specific haplotypes and marked genetic divergence from wild boars. Over time, this process has given rise to distinct breeds with specialized characteristics. **>80% market shares:** The commercial pigs (including Duroc, Large White, Landrace, Hampshire, and Pietrain) dominate the global pork industry, especially in China, which is the largest swine production and consumption market in the world. Notably, this dominance is not unique to China; similar patterns are observed in Southeast Asian countries as well as in South Korea and Japan. These breeds have been selectively bred for traits such as rapid growth, high feed efficiency, and consistent meat quality, making them highly favored by industrial-scale pig farms. Their widespread adoption reflects the global trend towards standardized and efficient pork production, driven by the need to meet rising demand for pork in a cost-effective and sustainable manner. **89 local breeds:** China possesses the most diverse genetic resources of domestic pigs in the world, among which the Taihu pigs are particularly notable. These pigs are renowned globally for their exceptional reproductive performance and serve as one of the primary parental breeds for improving the reproductive traits of the five major commercial pig breeds. **Fat-type:** Prior to the Industrial Revolution, early breeding objectives focused on the widespread demand for fat in Eurasian societies at the time. Moreover, under traditional rearing conditions, fat-type pig breeds typically exhibited stronger survival capabilities and adaptability. As a result, numerous fattening genotypes were subject to directional selection during the domestication process. **Criollo pigs** are not only an important genetic resource but also play a significant role in the economy and social development of local communities. They can utilize local feed resources and convert household waste into animal protein, providing a sustainable economic source for rural communities. Moreover, the meat of Criollo pigs is of high quality, with a high lean meat content and excellent flavor, making it of significant commercial value. **0 local breed:** Current research indicates that there are no indigenous pig species in Oceania. The existing breeds in the region are primarily the result of human colonization, involving the introduction and breeding of pigs that incorporate Eurasian genetic lineages.

Eurasian wild boars, the ancestral source population, harbor substantial genetic diversity shaped by Pleistocene climatic oscillations and prolonged geographic isolation.^{18,21,28} Genome-wide analyses reveal two major clades that diverged approximately 1.0–1.6 million years ago: a western lineage (Iberian and Central European populations) and an eastern lineage (Chinese and Japanese populations) (Figure 2). Despite this deep split, extensive post-divergence gene flow and regional substructure persist across both clades, which differ in nucleotide diversity ($\pi = 0.0032\text{--}0.0045$), linkage disequilibrium ($LD_{100} = 30\text{--}80\text{ kb}$), and genetic differentiation ($F_{ST} = 0.15\text{--}0.32$ between clades).^{13,18,21} Distinct adaptive signatures also occur in key loci. For example, Western Eurasian wild boars carry unique haplotypes at nuclear receptor subfamily 6 group A member 1 (*NR6A1*) and leptin receptor (*LEPR*) loci, whereas Eastern populations exhibit fixation of specific cytochrome P450 family 4 subfamily F member 2 (*CYP4F2*) variants linked to cold adaptation.^{15,22,25,31,49} This ancestral genetic variation provided the substrate from which domestic pig diversity later emerged and offers a benchmark for assessing historical introgression from wild boars into modern breeds.^{11,15,26,50}

Gene flow-mediated genetic architecture: Domestic pig diversity shaped by ancestral divergence

Population genomic analyses indicate that most local pig breeds neither derive from a single ancestral source nor evolved in isolation.^{8–11,16,51} Their genomes typically reflect contributions from both wild boars and other domestic populations,^{8,50–53} arising from gene flow among regional breeds, recurrent hybridization with wild boars in loosely managed systems, and deliberate crossbreeding for production traits.^{26,29,50–53} Introgression from

commercial lines is also used to maintain effective population size (N_e) and limit inbreeding in endangered local breeds.^{51–53} Geographic separation and human-controlled breeding nevertheless remain major forces structuring differentiation among populations.^{51,53}

Drawing on published genomic and phenotypic data from multiple *Sus scrofa domesticus* breeds, domestic lineages can be broadly categorized by historical gene flow intensity from wild boars, using: (1) fixation frequency of domestication alleles and wild haplotype introgression, assessed via f -statistics (F_{ST}) and N_e ; (2) linkage disequilibrium (LD) decay patterns; and (3) the magnitude of within-population phenotypic variation.

Breeds with high historical gene flow from wild boar (HHGF) retained substantial genetic continuity in early differentiation

HHGF populations experience recurrent wild boar hybridization under relatively weak artificial selection, with the defining genomic consequence being incomplete fixation of domestication alleles, as ongoing introgression continually replenishes ancestral variation.^{15,50,53} Representative breeds include European free-range populations, Turopolje, Iberian, Alentejana, and East Balkan pigs, and Chinese indigenous breeds such as *Meishan* and *Min*.^{15,50,53,54}

Ancestral allele retention at domestication loci illustrates this continuity. Vertebral number is an economically important trait directly affecting carcass length and the yield of high-value cuts.¹⁵ Wild boars typically carry a fixed count of 19 vertebrae, while domestic pigs commonly have 21–23, variation strongly associated with *NR6A1*.⁴⁹ The ancestral *NR6A1* c.748C allele is fixed in wild boars, while the derived c.748T allele, which increases vertebral number, predominates in many European breeds and is fixed in commercial

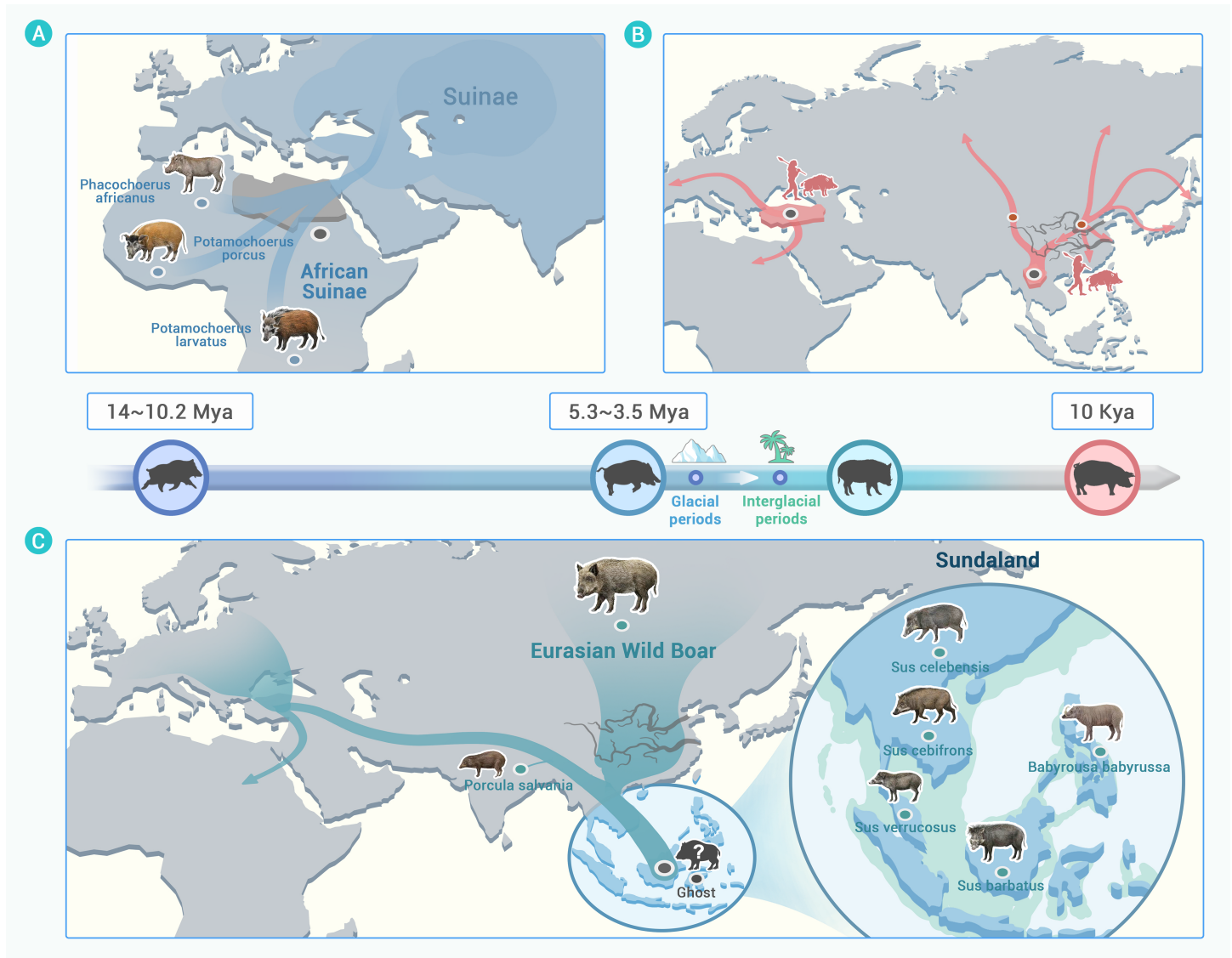


Figure 2. Schematic illustration of the geographic evolution of *Sus scrofa* within the family Suidae (A) Around 14 Mya, the other species within Suidae that were widely distributed across Eurasia and North Africa were gradually replaced by the emerging Suinae species in Eurasia, which ultimately diversified into the ancestors of numerous Sub-Saharan suids in Africa. Eurasian Suinae subsequently diversified into multiple genera. (B) Subsequently, *Sus* species emerged and migrated into Island Southeast Asia. The repeated alternation of ice ages and interglacial periods-mediated fluctuating sea levels facilitated the dispersal and diversification of *Sus* species across the Sunda Shelf. (C) The regions of Anatolia and China have played a foundational role in the formation of various lineages of Eurasian domestic pigs.

lines.^{15,22,49,55} HHGF populations retain substantial ancestral allele frequencies due to wild boar introgression: European HHGF breeds including Turpolje, Alentejana, Apulo-Calabrese, Majorcan Black, Black Slavonian, Iberian, and Nero Siciliano show *NR6A1* C allele frequencies of 5–43%, while some unsystematically managed Chinese indigenous breeds reach 41–75%.^{15,22,56}

Despite regional differences in skeletal traits, increased fat deposition was a shared breeding objective across both Europe and Asia.^{56–61} Lard was historically essential for cooking, food preservation, and energy in pre-industrial societies, making adiposity a central selection target.^{59–61} *LEPR* haplotype surveys across Eurasian wild boars and domestic pigs reveal two distinct phylogenetic groups reflecting deeper patterns of continental divergence.^{15,25,56,59} European populations harbor 11 haplotypes, nine unique to wild boars, with the *LEPR* H1 haplotype predominating (frequency = 51/108 across 54 Eurasian populations). Asian lineages, including Chinese Meishan, Vietnamese pigs, and Siberian wild boars, form separate clusters with species-specific haplotypes (H14, H21, H23, H24) absent from European lineages, consistent with independent domestication events, deep ancestral divergence, and limited post-domestication *LEPR* gene flow between continents.^{15,25,56,57,59}

Selective sweep analyses further reveal distinct selection pressures. In European Iberian pigs, the *LEPR* H1 haplotype is completely fixed ($\pi = 0$) and

deviates significantly from neutrality (HKA test: $\chi^2 = 5.657$, $P = 0.0174$). By contrast, Asian domestic breeds such as Meishan and Vietnamese pigs retain substantially higher haplotype diversity ($Hd = 0.811$).^{15,25,56,57,59} Although the c.1987T allele (*LEPR* H1), which is associated with increased fat deposition, is nearly fixed in Iberian pigs and the closely related Alentejana breed, both known for pronounced adiposity, this allele remains uncommon in many other HHGF populations.^{15,25,56} Because lipid deposition is a polygenic trait, the genetic loci underlying fat accumulation differ among fat-type European breeds.^{56,58} HHGF populations therefore exhibit selected phenotypes but retain substantial genome-wide diversity due to ongoing introgression, which reduces LD through recombination.^{11,62–66}

Reduced LD is thus another defining genomic feature of HHGF breeds, reflecting sustained connectivity with wild populations.^{12,15,29,67} In contrast to intensively selected commercial breeds maintained in closed breeding systems, such as Duroc, Landrace, and Large-white, which show extended haplotype homozygosity, HHGF populations exhibit weaker LD and higher recombination rates due to continuous genetic exchange.^{11,12,29,67,68} Gene flow between wild and domestic pigs has been a persistent component of domestication rather than an event limited to its earliest stages.^{9,11,69,48} Even under modern industrial farming systems, domestic pig-derived single nucleotide polymorphisms (SNPs) continue to introgress into Eurasian wild boar popu-

lations.^{29,48,53,69} Hybridization with multiple ancestral sources reshuffles haplotypes and introduces divergent alleles, which can obscure genetic differentiation among domestic populations.¹⁶

The evolutionary consequences of gene flow depend on its interaction with selection. Advantageous alleles are retained and may spread, deleterious variants are removed by purifying selection, and neutral variants persist as background diversity.^{15,48,70,71} This selective filtering operates primarily at individual loci rather than entire haplotypes, maintaining genomic mosaicism while limiting LD expansion.^{62,63} Continuous gene flow thus preserves high standing genetic variation that may facilitate future adaptive responses, provided N_e remains large enough to prevent stochastic loss of beneficial alleles.^{62,63,72}

This genetic plasticity can nonetheless be threatened by population decline or restricted gene exchange.⁷² Pedigree analysis of the Okinawan Agu pig illustrates the risk: over 18 years (2003–2020), the population lost approximately 5% of its genetic diversity through unequal founder contributions (2%) and genetic drift (3%).⁷² Following the transition to a closed breeding system in 2012, the inbreeding coefficient rose sharply, peaking at 10.5% in 2015, while N_e declined to just 14.6, well below the FAO-recommended threshold of 50 for long-term genetic viability.⁷²

Although artificial selection independently fixed loci affecting behavior and morphology in both European and Asian domestic pigs, extensive gene flow has broadly weakened genome-wide LD, preventing the expansion of large homozygous regions, a pattern particularly evident in frequently admixed local breeds.¹¹ Frantz *et al.* described this structure as "islands of domestication": loci under strong selection embedded within a genomic background of continuous gene exchange.¹¹ Notably, gene flow does not uniformly reduce LD; directional introgression can introduce large linked haplotypes, transiently elevating LD in recipient populations.^{73,74} The Korean pig system illustrates how similar LD patterns can arise from distinct evolutionary histories.⁷⁴ Jeju Black pigs, geographically isolated on the Korean peninsula, display lower LD than mainland Korean pigs intensively hybridized with commercial breeds, a difference more likely reflecting the historically large N_e of Jeju pigs than recent admixture.^{30,74} Low LD alone is therefore insufficient evidence of recent gene flow; LD decay patterns and haplotype sharing with source populations must be evaluated jointly.^{11,26,75} Inferring historical gene flow from single summary statistics such as F_{ST} or N_e risks conflating distinct demographic processes.^{74–77}

Unlike the Korean case, where isolation preserves low LD, Chinese indigenous pigs show remarkable genomic resilience to commercial line introgression, maintaining weak LD despite historical admixture.^{67,73,74} This pattern is broadly shared across Chinese breed groups, indicating a common demographic foundation rather than a breed-specific anomaly.^{48,78} Several factors likely underlie this stability: historically large N_e from pooled village herds and circulating breeding boars; weaker, more decentralized artificial selection producing soft sweeps or no sweeps at many loci; and asymmetric introgression from European commercial pigs.^{9,61,67} Mitochondrial and Y-chromosome evidence shows that Asian haplotypes contributed substantially to European breeding stocks, while reciprocal European nuclear introgression into Chinese populations has remained limited.^{9,67} Chinese breeds consequently display smaller LD blocks and faster LD decay than European breeds, with introduced high-LD segments gradually broken down through recombination within their large- N_e , weakly selected genomic background.^{67,75}

The Mekong River Basin and north-central China are widely recognized as the principal centers of East Asian pig domestication, contributing foundational lineages that later spread across East Asia and into Oceania (Figure 2). Mitochondrial phylogenetic analyses indicate that Mekong wild boars supplied the ancestral haplotypes underlying most East Asian domestic pig lineages.⁷⁰ Asian wild boars, particularly those from the Mekong Basin and the middle–lower Yangtze region, show higher nucleotide diversity ($\pi = 0.0032–0.0045$) and haplotype diversity ($H_d = 0.946$) than European wild boars ($\pi = 0.0028–0.0034$; $H_d = 0.845$).^{70,79} This high ancestral diversity is preserved in East Asian domestic pigs (especially Chinese pigs), which display substantially greater nucleotide diversity ($\pi = 3.31 \times 10^{-3}–4.03 \times 10^{-3}$) than European domestic breeds ($\pi = 1.66 \times 10^{-3}$ to 2.30×10^{-3}) (Table 1 and Supplementary Information).^{80,81} Moreover, East Asian pigs exhibit much

shorter LD decay distances (0.286–0.807 Kb) compared with European pigs (2.367–4.439 Kb), further reflecting their greater standing genetic variation.^{80,81} Chinese indigenous breeds have played a central role in maintaining and disseminating this ancestral diversity.^{70,82–84} Breeds such as Meishan, Min, and Erhualian retain extensive haplotype sharing with regional wild boars, with F_{ST} values of 0.08–0.12 between Chinese domestic pigs and Asian wild boars, indicating moderate differentiation alongside ongoing genetic connectivity.^{8,9,70,82–85} Selective sweep analyses (iHS, XP-EHH) further identify genes involved in cold tolerance, heat adaptation, and metabolic resilience in these HHGF populations, traits likely inherited from wild boar ancestors.^{8,9,70,82–85} Accordingly, Chinese HHGF breeds maintain higher nucleotide diversity ($\pi = 2.43 \times 10^{-3}–2.73 \times 10^{-3}$) and greater structural variant (SV) diversity than European pigs ($\pi = 1.68 \times 10^{-3}–1.76 \times 10^{-3}$), with faster LD decay (half-life: 332–615 nt vs. 882 nt in European commercial pigs).^{81,84} Elevated SV diversity in these gene-flow-permissive populations is consistent with a soft-sweep model, reflecting a large reservoir of standing structural variation available to selection without requiring *de novo* mutation.

Taken together, HHGF breeds distributed across Eurasian local pig populations share a characteristic genomic architecture, "selective fixation within a background of gene flux," in which islands of domestication are embedded within a genome highly permeable to gene flow. Early-stage divergence in this system does not proceed as a simple linear progression toward genome-wide fixation, but as a dynamic balance: persistent gene flow maintains genetic continuity while selection simultaneously stabilizes advantageous traits driving phenotypic differentiation. Through this interplay, *Sus scrofa domesticus* continues to diverge from wild boars while remaining genomically connected to them.

Breeds with low historical gene flow from wild boar (LHGF): Accelerated genetic fixation

Modern perceptions of the domestic pig's phenotype largely derive from the global spread of Euramerican lean-type commercial breeds during the twentieth century.^{14,24,61} The foundations of these lines, however, were established much earlier: following the Industrial Revolution, intensive selective breeding and crossbreeding rapidly reshaped pig morphology, progressively integrating desirable traits from multiple European and Asian breeds into large-scale commercial varieties optimized for agricultural production (Figure 1).

From an evolutionary perspective, modern commercial breeds, Duroc, Large-white, Landrace, Hampshire, and Pietrain, represent the most derived endpoint of the domestication continuum.^{24,26,28,86,87} Their rapid growth, muscularity, and feed efficiency reflect sustained directional selection over many generations within tightly controlled breeding systems that severely restrict gene flow with wild boars or other populations.^{15,24,61,87–89} Consequently, LHGF breeds display accelerated fixation of domestication alleles and clear genomic signatures of selection, including strong selective sweeps, extensive haplotype homozygosity, and reduced genomic flexibility.^{8,14,32,62}

A well-characterized example occurs at the melanocortin 4 receptor gene on chromosome 1, where mutations influence daily weight gain, lean meat percentage, and backfat thickness.⁹⁰ Comparative allele frequency analyses across wild boars, European local breeds, and commercial lines reveal a clear domestication gradient.¹⁵ The Asn298 (A) allele, associated with enhanced growth, is far more frequent in commercial breeds than in traditional populations, while wild boars carry only the ancestral G allele.^{15,90–92} Long-term selection has further expanded homozygous regions in commercial lines while eliminating alternative variants. Whole-genome analyses show that Large-white and Landrace possess substantially lower nucleotide diversity than HHGF breeds such as Rongchang or Tibetan pigs, with two- to three-fold increases in runs of homozygosity length and genomic inbreeding coefficients (FROH).⁸⁸ Selective sweep regions overlap with several known quantitative trait loci, such as *NR6A1*.^{14,49}

Heterozygosity (Expected/observed heterozygosity, H_e/H_o) is widely used to assess within-population genetic diversity.^{14–16} Although commercial breeds originally incorporated multiple ancestral lineages, prolonged intensive selection has reduced their diversity relative to local populations.^{15,52} Their large global population sizes slow diversity loss from genetic drift but

Table 1. Representative advantageous traits and candidate genes in Chinese local breeds.

Advanced traits	Typical breed	Candidate genes	Genetic mutation types	Validation level	References
Cold-resistance	Min	PRDM1		GWAS/QTL study, Functional validation in vitro	S.1
Diarrhea-resistance	Pudong White	MUC4	SNP	"Candidate gene association study, Phenotypic correlation analysis"	S.2
		MUC13			
PRRSV-resistance		USP18		Functional validation in vitro, Population genetics analysis	S.3
PCV2-resistance	Dapulian	Myostatin		Functional validation in vitro, Haplotype association analysis	S.4
Meat quality		MRC1	Indel	Functional validation in vitro, Indel polymorphism analysis	S.5
Small size	Bama	AFF4		GWAS, Population genetics analysis, Functional validation in vitro	S.6
High-altitude adaptability	Tibetan	EPAS1	SNP	"Population genetics analysis Selective sweep analysis, Functional validation in vitro"	S.7
		CYP4F2			
		LIMCH1			
		PPP2R5A			
Meat quality	Erhualian	KLHL31		"GWAS/QTL study, Phenotypic correlation analysis, Population genetics analysis"	S.8, 9
		MYH3			
		LDHA			
		WDR16			
		UCP3	Complex mutation	GWAS/QTL study, PCR-RFLP genotyping, Phenotypic association analysis	S.10
	Jinhua	CACNA1S	SNP	"PCR-RFLP genotyping, Haplotype analysis Phenotypic association analysis, Population genetics analysis, PCR-SSCP genotyping, SNP detection, Sequence homology analysis"	S11, 12
		H-FABP			
		ACSL4			
Fat accumulation		FRZB		SNP detection, PCR genotyping, Functional validation in vitro	S.13
	Laiwu	APOB	Complex mutation	"GWAS, meta-analysis, Haplotype-based association analysis, indel/SNP genotyping, F-drop test, Population genetics analysis"	S.14
	Erhualian	CLSTN2		"SLAF-seq genotyping, Population genetics analysis, Differentially selected region identification, functional enrichment analysis, GWAS, SNP genotyping"	S15, 16
		ESR1			
Reproductive performance	Meishan	OLFML2A	SNP	GWAS, Meta-analysis, Bayesian fine mapping, TWAS, PheWAS, Functional validation in vitro	S.17
	Wannan Black	FSHR		PCR-SSCP genotyping, SNP detection, Haplotype analysis, Phenotypic association analysis	S.18
Skin thickness	Chenghua	FBN1		Selective sweep analysis, XP-CLR analysis, SNP detection, Functional validation in vitro	S.19

*Note: PRRSV: Porcine reproductive and respiratory syndrome virus; PCV2: Porcine circovirus type 2; *PRDM1*: PR/SET Domain 1; *MUC4*: Mucin 4; *MUC13*: Mucin 13; *USP18*: Ubiquitin-specific protease 18; *MSTN*: Myostatin; *MRC1*: Mannose receptor C-Type 1; *AFF4*: AF4/FMR2 family member 4; *EPAS1*: Endothelial PAS domain-containing protein 1; *CYP4F2*: Cytochrome P450 4F2; *LIMCH1*: LIM and calponin-homology domain 1; *PPP2R5A*: Protein phosphatase 2 regulatory subunit B'Alpha; *KLHL31*: Kelch-like protein 31; *MYH3*: Myosin heavy chain 3; *LDHA*: Lactate dehydrogenase A; *WDR16*: WD repeat domain 16; *UCP3*: Uncoupling Protein 3; *CACNA1S*: Calcium voltage-gated channel subunit alpha1 S; *H-FABP*: Heart-type fatty acid-binding protein; *ACSL4*: Acyl-CoA synthetase long-chain family member 4; *FRZB*: Frizzled-related protein B; *APOB*: Apolipoprotein B; *CLSTN2*: Calsyntenin-2; *ESR1*: Estrogen receptor 1; *OLFML2A*: Olfactomedin-Like 2A; *FSHR*: Follicle-stimulating hormone receptor; *FBN1*: Fibrillin-1; SNPs: Single nucleotide polymorphisms; Indel: Insertion-deletion polymorphism; GWAS: Genome-wide association study; QTL: Quantitative trait locus; PCR-RFLP: Polymerase chain reaction-restriction fragment length polymorphism; PCR-SSCP: Polymerase chain reaction-single strand conformation polymorphism; SLAF-seq: Specific-locus amplified fragment sequencing; TWAS: Transcriptome-wide association study; PheWAS: Phenome-wide association study; XP-CLR: Cross population composite likelihood ratio.

the widespread dominance of commercial lines has sharply reduced the niches available to local breeds, driving rapid declines in census population sizes.^{15,50,53,89} Many local breeds have consequently undergone severe bottlenecks accompanied by inbreeding depression and elevated FROH.⁵¹ Under these conditions, heterozygosity alone cannot distinguish directional selection in commercial systems from random drift in declining local populations, accurate inference of historical gene flow therefore requires integrative evaluation of FST, Ne, and FROH, across multiple generations.

Intense artificial selection in LHGF breeds has compressed neutral genetic variation and eroded adaptive potential, potentially compromising the ability to respond to environmental change and increasing disease susceptibility.^{52,71,89} Evidence for such risks emerges from studies of balancing selection at the mucin 13 (*MUC13*) gene.³² Analysis of a second filial generation (F_2) population derived from White Duroc boars and prolific Chinese Erhualian sows, generated without artificial selection, identified two haplotypes (Hap1 and Hap2) defined by mutations in *MUC13*. Association analyses across more than 300 phenotypes revealed clear functional trade-offs: Hap1 enhanced several growth traits but increased susceptibility to enterotoxigenic *Escherichia coli* F4ac-induced diarrhea.³² This illustrates how exclusive focus on production traits can inadvertently compromise disease resistance and long-term viability.³²

Overall, reduced gene flow in LHGF breeds removes a key constraint on differentiation, allowing directional selection to rapidly fix adaptive alleles and reinforce phenotypic distinctions. The contrasting genomic structures of HHGF and LHGF pigs reflect their differing evolutionary histories: HHGF populations maintain a mosaic of ancestral polymorphisms and adaptive loci shaped by continuous bidirectional exchange with wild boars, while commercial European pigs display stronger selection signatures and reduced diversity. This contrast parallels natural speciation, where reduced gene flow facilitates fixation of ecologically advantageous traits and accelerates genetic divergence.

Divergent selection-driven homogeneity and heterogeneity of genetic architecture in early divergence

Divergent selection imposed by heterogeneous environments is a major force shaping early population divergence.^{93,94} In domestic pigs, geographic variation in climate, diet, and pathogen pressure has generated distinct adaptive genetic architectures paralleling ecological divergence in natural systems.^{32-34,48,90,95} Acting on standing variation introduced through gene flow, selection preserves advantageous alleles and reshapes trait architecture through regulatory modification, polygenic pathway optimization, and consolidation of adaptive phenotypes.^{81,91-94}

Environmental adaptation: Genomic variation and phenotypic remodeling

Tibetan pigs above 3,000 m elevation show strong positive selection in the regulatory region upstream of *CYP4F2*, a candidate locus associated with hypoxia tolerance and circulatory regulation, with pronounced differentiation from low-altitude Chinese breeds.³¹ This adaptive variant displays pronounced differentiation from low-altitude Chinese breeds.³¹ Likewise, a "Map-to-Pan" assembly framework combined with transposable element (TE) insertion polymorphism analysis has revealed numerous adaptive TE insertions in Tibetan wild boars.³⁷ Most belong to long interspersed nuclear element-1, endogenous retrovirus-like (ERVL), and hobo-Ac-Tam3 (hAT) families and exhibit pronounced population differentiation. For example, an ERVL insertion in protein kinase C epsilon, which regulates hypoxia-induced calcium signaling and pulmonary vascular contraction, is nearly fixed (frequency > 0.95) in high-altitude Tibetan populations but remains rare (< 0.15) in lowland Chinese domestic breeds and European commercial pigs. Similarly, a hAT insertion in dedicator of cytokinesis 4, involved in vascular lumen morphogenesis, shows a frequency difference exceeding 0.8 between Tibetan and non-Tibetan lineages.⁸⁷

Chinese indigenous pigs also display strong geographic structure in thermal adaptation genes along climatic gradients.³⁴ Southern breeds such as the Rongchang pig, adapted to hot and humid subtropical environments, show pronounced selection on heat shock protein (HSP) genes that protect cells under heat stress. These breeds also carry immune regulatory variants linked to thermal resilience, including promoter deletions in GTPase of immunity

associated protein 6.³⁴ Northern breeds such as the Min pig exhibit selection signals for cold tolerance, particularly in genes regulating adipogenic differentiation and tissue remodeling, including *tumor protein, translationally controlled 1*.³⁴ An SV further illustrates this pattern: a 257-bp porcine repetitive element insertion within an intron of *HSPH1*, ~12.6 kb downstream of its transcription start site on chromosome 11, segregates at variable frequencies across Eurasian pigs and is absent from the Sscrofa11.1 reference genome.⁸⁷ It occurs in nearly half of European commercial pig individuals (47%) but is nearly absent from Chinese indigenous breeds and Tibetan wild boars.⁸⁷ Critically, the insertion is associated with significant *HSPH1* down-regulation — given *HSPH1*'s role in thermotolerance and cellular stress responses, this exemplifies how TE-mediated SVs contribute to adaptive divergence under distinct environmental pressures and represent a valuable resource for resolving early divergence patterns.^{24,84,87}

Recent telomere-to-telomere assemblies from long-read sequencing have enabled more refined selective sweep analyses, revealing previously unrecognized genetic determinants of body size divergence.⁹⁵ Comparative analyses between two Chinese small-bodied breeds (Wuzhishan and Bamaxiang) and two large commercial breeds (Large-white and Duroc) identified 133 candidate genes associated with body stature. Among these, *polypeptide N-acetylgalactosaminyltransferase 13* (*GALNT13*) exhibited particularly strong selection signatures.⁹⁵ The gene is highly expressed in skeletal tissues, including thoracic vertebrae and epiphyseal cartilage, but shows significantly reduced expression in small-bodied pigs.⁹⁵ Functional assays demonstrated that *GALNT13* overexpression suppresses chondrocyte proliferation while promoting differentiation, indicating a regulatory role in skeletal development and body size.⁹⁵ These findings complement previously identified body-size regulators, collectively illustrating how divergent regional adaptation and distinct breeding objectives have produced contrasting evolutionary trajectories across local pig populations.^{15,96-98}

Planned breeding-driven directed gene flow: Promoting genomic homogeneity across domestic populations

Human-directed introgression combined with sustained artificial selection can generate stable breeds that integrate advantageous traits from genetically distinct lineages.^{99,100} Evidence from Neolithic pig genomes shows that even during early domestication, when gene flow from wild boars remained common, selection consistently targeted key phenotypic traits.¹⁷ An FST outlier analysis comparing Neolithic European pigs with modern wild boars identified strong selection signals in loci associated with coat color, behavior, and pathogen resistance.¹⁷ Notably, v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog (*KIT*) ligand, which influences coat patterning, experienced selection across multiple Neolithic populations, complementing the role of melanocortin receptor 1 (*MC1R*) in stabilizing black coat color.¹⁷ This pigmentation phenotype represents a hallmark domestication trait that remains conserved in modern HHGF breeds.¹⁰¹ The transition from the camouflaged wild boar coat—maintained by purifying selection at *MC1R*—to the diverse pigmentation of domestic pigs reflects a fundamental shift from predator avoidance to human-driven selection acting on *MC1R* and *KIT*.¹⁰²⁻¹⁰⁴ Epistatic interactions between these loci generate white, piebald, and belted patterns through altered melanocyte distribution and pigment-type switching, providing a paradigmatic example of regulatory rewiring during early lineage divergence.¹⁰³⁻¹⁰⁵ The persistence of these signals despite recurrent wild boar introgression indicates that deliberate human selection effectively counteracted hybridization-driven homogenization. Regional differences in selective intensity, stronger pressure on behavioral genes in closed-breeding Central European herds than in more commensal northwest European systems, further illustrate that domestication proceeded as a mosaic shaped by local husbandry practices.¹⁷

Chinese indigenous pigs generally reach sexual maturity earlier than European domestic breeds.⁶¹ The *aryl hydrocarbon receptor* (*AHR*) gene, which regulates ovarian development through the hypothalamus–pituitary–gonadal axis, exhibits strong signatures of positive selection in several South and East Chinese breeds known for exceptional reproductive performance, including Erhualian, Jinhua, Luchuan, and Neijiang pigs.⁸ In pigs, *AHR* expression fluctuates across the estrous cycle and appears to participate in reproductive regulation. Four nonsynonymous mutations have been identified within exon

11 of the *AHR* gene in domestic pigs.^{26,92} One variant originating from Chinese pig populations has been preferentially selected in modern European Large-white pigs and is associated with increased litter size in several commercial European breeds.²⁶ This Asian-derived haplotype currently exists in strong LD within Large-white populations but has not become fixed in their European ancestral breeds.²⁶ These observations indicate that enhanced reproductive capacity remained a central objective in European breeding programs. The relevant genetic variants were likely introduced during nineteenth-century crossbreeding with highly prolific Chinese breeds, which substantially contributed to the genetic architecture of modern commercial pigs.^{26,87}

Gene flow between geographically distant populations can also occur indirectly through intermediate breeds.⁷⁵ A notable example involves the transmission of commercial pig haplotypes into Tibetan pigs, mediated by Lijiang pigs acting as genetic intermediaries.⁷⁵ More broadly, genomic homogeneity arising from planned breeding is not unique to modern improvement programs. Shared cultural traditions and long-standing agricultural practices can also produce consistent selective sweeps across geographically distributed breeds. One example is the T-box transcription factor 19 (*TBX19*) locus, where a common haplotype is shared among nearly all Chinese domestic pig breeds but differs markedly from those found in European domestic pigs and Eurasian wild boars.⁸ Within domestic pigs, the *TBX19* region exhibits extremely low genetic diversity compared with the SNP-rich genomes of Chinese wild boars, indicating a strong selective sweep during domestication.⁸ The *TBX19* gene plays a critical role in pituitary cell differentiation and regulates expression of the *proopiomelanocortin* gene, which produces adrenocorticotrophic hormone and modulates metabolic processes through the hypothalamic–pituitary–adrenal axis.⁸ In addition, *TBX19* indirectly affects behavioral traits by influencing corticosteroid and androgen levels produced by the adrenal cortex.⁸ Interestingly, at this locus Chinese domestic pigs show greater sequence similarity to *Sus celebensis* than to Eurasian wild boars or European domestic pigs, leaving the origin of this unusual haplotype unresolved. Current hypotheses include ancient introgression from an unsampled lineage or convergent evolution at this genomic region.⁸

Taken together, environmental selection is a central driver of divergence, but its influence on genetic architecture is strongly modulated by the historical intensity of gene flow. Across global pig populations, this interaction has produced distinct evolutionary outcomes: HHGF breeds retain greater genomic flexibility for continued adaptation, whereas LHGF breeds consolidate lineage-specific traits through targeted allele fixation. Convergent patterns across geographic regions reflect shared evolutionary priorities during early differentiation, and similar pressures related to energy acquisition, reproduction, and environmental adaptation generate parallel genetic responses among breeds experiencing comparable gene flow regimes. LHGF breeds worldwide consequently tend to fix alleles associated with growth efficiency and reproductive performance, while HHGF breeds maintain higher diversity in immune and metabolic pathways, preserving adaptive capacity without prematurely entering reproductive isolation.^{24,106}

The early diversification of the Suidae family was largely unified by the evolutionary expansion of *Sus scrofa*.¹⁰ Human domestication has progressively fragmented this once cohesive lineage into numerous domesticated populations and breed types.¹⁰ Both visible and cryptic differences between domestic pigs and their wild counterparts ultimately trace back to modifications in genetic architecture and gene expression.^{10,12,15,35} Continued development of high-quality reference genomes and trait-associated genetic markers will therefore be essential for dissecting the natural and anthropogenic selection processes that shaped *Sus* evolution and for clarifying the genetic and regulatory mechanisms underlying phenotypic diversity.

GENE FUNCTIONAL REPLACEMENT IN THE EARLY DIVERGENCE OF *SUS SCROFA DOMESTICUS*: BA STRUCTURE AS A MODEL

Domestication imposed novel ecological and nutritional pressures on pig populations, redirecting their evolutionary trajectories and reshaping metabolic regulation relative to wild boars. One striking example involves the BA metabolic system of *Sus scrofa domesticus*.^{36,40} In domestic pigs, the loss of CA synthesis through functional silencing of CYP8B1, together with the emergence of HCA production via neofunctionalization of CYP4A21, repre-

sents a coordinated genetic transition rather than a set of random mutations.³⁸ This metabolic reorganization appears closely tied to the domestication niche, where animals rely on stable and energy-rich diets rather than unpredictable wild foraging. Domestic pigs also show strong resistance to metabolic disorders associated with high-energy feeding during fattening,³⁶ suggesting that the shift in BA composition reflects a largely hidden physiological adaptation. Consequently, changes in the BA profile provide a useful framework for examining the molecular processes underlying genetic remodeling during the early stages of divergence. More broadly, this metabolic transition illustrates how early differentiation reconciles two opposing forces: directional selection, which fixes advantageous traits, and ongoing gene flow, which preserves genetic continuity. Although these dynamics differ along the HHGF–LHGF spectrum, they converge at the level of metabolic pathways through regulatory flexibility. Structural changes in regulatory genes controlling BA composition therefore represent an informative model for investigating the evolution of genetic architecture during early divergence in domestic pigs.

Evolutionary patterns of BA profiles in wild and domestic pigs

Environmental differences often impose directional selection that drives ecological divergence and adaptive evolution.¹⁰⁷ Variations in food availability and nutritional composition are particularly important in shaping metabolic strategies among related species.¹⁰⁷ The ability to evolve novel catalytic enzymes is therefore central to the emergence of adaptive metabolic pathways.³⁸ Because lipids perform fundamental roles in cellular structure and physiology, lipid metabolism is deeply involved in mammalian adaptive evolution.¹⁰⁷ BAs, the final products of cholesterol metabolism, are essential for lipid digestion and the absorption of fat-soluble vitamins, while also regulating systemic and local lipid homeostasis.³⁵ As dietary lipid resources vary across ecological niches, the structure and composition of BA pools show remarkable diversity among species.¹⁰⁸ Both the molecular structure of individual BA monomers and the relative abundance of different BA classes provide valuable information for evolutionary and taxonomic studies. For example, in many teleost fishes, including *Amia* and most teleost lineages, the capacity to synthesize key stereoisomerases involved in CA formation is well established, whereas this process remains incomplete in Cyprinidae.¹⁰⁸ Such variation in BA biosynthetic capacity has been used to reconsider phylogenetic relationships among teleost groups, illustrating the broader evolutionary relevance of BA metabolism.

Across vertebrates, BA pools broadly track dietary ecology, with CA dominating in carnivorous lineages where it regulates cholesterol turnover, whereas chenodeoxycholic acid (CDCA) prevails in herbivores—although ruminants present a notable exception, exhibiting CA-rich profiles possibly driven by microbial fermentation-derived lipid loads.^{108–116} Much of this interspecific diversity stems from lineage-specific hydroxylation of CDCA catalysed by cytochrome P450 enzymes rather than direct modification of CA, underscoring the complex interplay between feeding strategy, digestive physiology, and biosynthetic pathway evolution.^{117–120}

Domestic pigs display a distinctive BA composition. In adult animals, HCA constitutes the principal trihydroxy primary BA synthesized *de novo* in the liver and occurs in proportions comparable to CA in humans.^{39,108} Wild boars, and possibly warthogs, retain the capacity to synthesize CA, and their livers typically produce more CA than HCA.^{40,41} This contrast likely reflects ecological differences in feeding strategies. Wild boars and minimally domesticated local breeds exploit diverse habitats and therefore consume a broader range of foods than intensively managed commercial pigs.^{45,69,121} Their dietary diversity contributes to exceptional ecological adaptability and partly explains the species' success as a globally invasive mammal, now distributed across all continents except Antarctica and many Pacific and American islands.¹²¹ Acorns and other hard mast from *Quercus* species represent key food resources during autumn and winter, whereas herbaceous vegetation dominates the diet in spring and summer.^{122,123} Although plant matter forms the bulk of their diet, wild boars opportunistically consume fungi, earthworms, insect larvae, mollusks, crustaceans, fish, reptiles, birds, and small mammals when conditions permit.⁴⁵ Dietary lipid variation can promote adaptive divergence among consumer populations, as illustrated by comparisons between urban and wild pig populations.¹²⁴ Urban boars tend to be larger and heavier

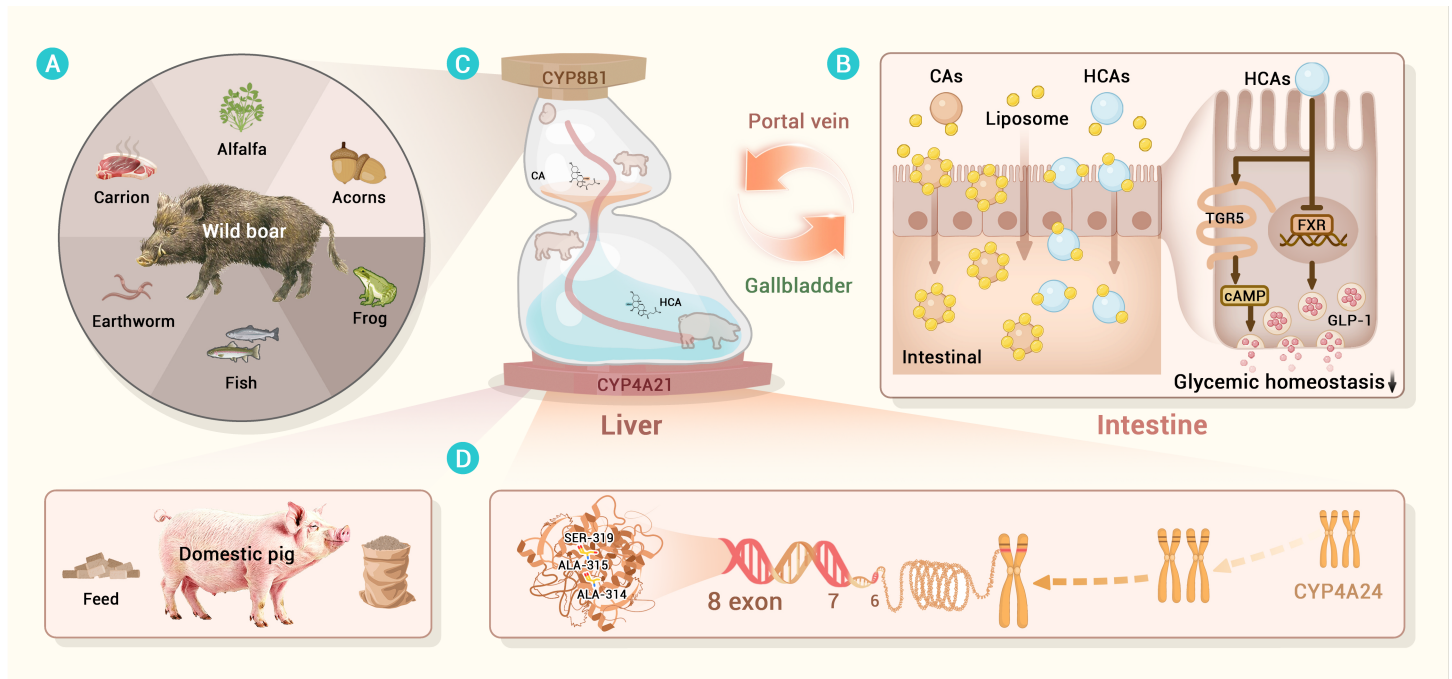


Figure 3. Genetic basis and metabolic modifications of elevated hyocholeic acid species (HCAs) in the total bile acid pool during the evolution from wild boar to domestic pigs (A) Despite a more monotonous diet compared to wild boars, domestic pigs have a stable energy supply and balanced nutrition. (B) Interspecies studies reveal that high levels of HCA stimulate Takeda G protein-coupled receptor 5 (TGR5)-glucagon-like peptide-1 (GLP-1) signaling in enteroendocrine L-cells, promoting rapid postprandial blood glucose clearance, which underpins diabetes resistance in domestic pigs and is conserved in humans and mouse models. (C) Cytochrome P450 family 4 subfamily A member 21 (CYP4A21) expression in the liver of domestic pigs increases during development, progressively replacing cytochrome P450 family 8 subfamily B member 1 (CYP8B1) to dominate the synthesis of trihydroxy bile acids. (D) Proposed evolutionary origin of CYP4A21 gene from cytochrome P450 family 4 subfamily A member 24 (CYP4A24) gene.

and display higher serum triglyceride concentrations but lower creatine levels than their wild counterparts, suggesting greater access to high-fat food sources.¹²⁴ In contrast, wild boars must forage actively to satisfy energy demands, and the retained ability to synthesize CA may enhance lipid absorption efficiency under such conditions.¹¹⁰⁻¹¹²

For highly domesticated pigs, however, food scarcity rarely threatens survival. Instead, the main physiological challenge arises from chronic positive energy balance associated with intensive feeding regimes (Figure 3). Functional studies indicate that high concentrations of HCA stimulate Takeda G protein-coupled receptor 5–glucagon-like peptide-1 signaling in enteroendocrine L-cells, thereby accelerating postprandial glucose clearance.³⁶ This mechanism contributes to a broader regulatory system that enables domestic pigs to maintain glucose homeostasis even after consuming energy-dense diets. Such regulation likely underlies their well-known resistance to diabetes during fattening and is also conserved in human and mouse models.³⁶ At the same time, HCA is less efficient than CA or CDCA at emulsifying and transporting dietary lipids in the small intestine because of its lower hydrophobicity.³⁵ Experimental studies in mice show that increased hydrophilicity of the BA pool can impair intestinal lipid absorption and induce weight loss.¹²⁵ A similar mechanism may contribute to the decline in apparent fat digestibility observed in piglets during weaning, from approximately 95% to 65–80%, in addition to the immature digestive system and limited enzyme secretion characteristic of this developmental stage.¹²⁶ Overall, the increasing proportion of HCA within the BA pool appears to enhance tolerance to prolonged energy-rich feeding during domestic pig production (Figure 3).

Although the biosynthetic pathways for CA and CDCA are well characterized in humans and rodents, the mechanisms governing HCA synthesis in pigs remain incompletely understood.³⁵ Some experimental evidence indicates that *de novo* HCA formation may involve a more complex pathway. *In vitro* studies show that porcine CYP4A21 catalyzes taurine-conjugated CDCA far more efficiently than unconjugated CDCA.¹²⁷ Because conjugation reactions are generally assumed to occur after completion of *de novo* BA synthesis, and because the key enzyme responsible, BA-CoA amino acid N-acyl-transferase, is not located in microsomes, the precise sequence of reactions and the cellular compartments involved remain unresolved.¹²⁸ Consequently,

whether the canonical BA biosynthetic model applies fully to domestic pigs is still uncertain. Nonetheless, it is well established that hepatic CYP4A21 catalyzes hydroxylation of BA precursors at the C₆ position in pigs.^{47,117,127}

Given the complex genetic relationship between wild and domestic pig populations, the regulatory network governing HCA synthesis provides a promising framework for investigating the evolutionary history of *Sus scrofa*. Developing molecular identification systems based on genes controlling HCA biosynthesis and expression could help clarify patterns of divergence and gene flow. Such an approach is valuable not only because relevant biochemical markers are easily measurable, but also because it directly links metabolic phenotypes with underlying regulatory evolution. A key unresolved question is whether metabolic adaptations such as the HCA shift represent a universal component of a domestication syndrome or independent responses to local selection pressures. Current evidence suggests an intermediate scenario: dominance of HCA appears consistent across domestic pig populations, yet the regulatory mechanisms producing this phenotype, such as the timing of CYP8B1 silencing or the magnitude of CYP4A21 expression, may vary quantitatively among breeds. This combination of conserved phenotypic outcome and flexible genetic implementation parallels patterns observed in other domestication traits, including coat color, and supports the idea that domestication operates through canalized phenotypic selection acting on genetically plastic developmental systems.

Hypothesis: CYP4A21-mediated functional replacement of CYP8B1 in *Sus scrofa domestica*

After birth, hepatic CYP8B1 expression declines progressively and becomes nearly undetectable in adulthood, whereas CYP4A21 expression in hepatocytes increases along an opposite developmental trajectory.³⁶ Because these enzymes introduce hydroxyl groups at the C₁₂ and C₆ positions of the steroid nucleus, respectively, their reciprocal expression patterns explain the characteristic BA profile of adult domestic pigs. Specifically, their BA pools are enriched in C₆-hydroxylated acids, including HCA, hyodeoxycholic acid, taurine-HCA, taurohyodeoxycholic acid, glycine-HCA, and glycohyodeoxycholic acid, while C₁₂-hydroxylated acids such as CA, deoxycholic acid, taurine-CA, taurodeoxycholic acid, glycine-CA, and glycodeoxycholic acid are largely absent.³⁹ The detection of CA and HCA at relatively high proportions in the total BA pool of wild boar suggests that the underlying mechanism of this

developmental expression switch likely involving epigenetic rather than coding sequence changes in domestic pigs.^{40,129} Notably, evidence for this switch in domestic pigs derives exclusively from commercial breeds.^{38,130} Although comprehensive BA profiling across fetal and adult stages remains unavailable for diverse breeds, targeted metabolome of BA from limited local varieties indicates that this replacement is largely conserved in domestic breeds.¹³¹ Additionally, high-fat diet-raised experimental pigs (derived from artificially selected miniature breeds) readily develop hepatic lipid metabolic disorders upon high-dose CA supplementation, reversely providing reciprocal evidence that the HCA-to-CA substitution in domestic pigs represents an adaptive evolutionary trait.¹³² Therefore, the tightly regulated spatial and temporal expression of these two enzymes, ultimately producing an HCA-dominated BA pool in domestic pigs, offers an unusually clear system for studying the genetic mechanisms underlying early divergence.³⁸

CYP8B1, a liver-specific enzyme in most mammals, functions as a key rate-limiting component of the microsomal classical BA synthesis pathway and largely determines the final product of the pathway.¹³³ When *CYP8B1* participates in the pathway, synthesis proceeds toward CA; when it is absent, CDCA becomes the primary product.³⁵ Although the *CYP4* gene family represents one of the most ancient branches of the cytochrome P450 superfamily, originating more than 1.3 billion years ago, the specific member *CYP4A21* emerged relatively recently compared with *CYP8B1*, which is highly conserved across mammals and even exhibits a relatively relaxed evolutionary trajectory.^{39,47,134,135} In most mammals, the functional *CYP8B1* gene displays a distinctive architecture characterized by a single exon lacking introns, a single-copy genomic arrangement, and a highly conserved open reading frame.^{133,135} Amino acid sequences of CYP8B1 proteins share more than 74% identity across species, underscoring the remarkable evolutionary conservation of this enzyme.^{133,135} Its widespread distribution, structural simplicity, and functional stability suggest that *CYP8B1* arose early in vertebrate evolution and has been maintained under long-term selective constraints.^{135,136} The absence of introns minimizes transcriptional variability and promotes stable transcript production, thereby reducing the likelihood that disruptive mutations interfere with essential metabolic processes.¹³⁷ Such extensive conservation implies that CYP8B1 performs a fundamental physiological role and may have originated before the diversification of chordates, subsequently becoming fixed as a stable component of vertebrate BA metabolism.¹³⁶ For this reason, CYP8B1 provides a valuable marker for investigating the genetic basis of BA biosynthesis and for reconstructing phylogenetic relationships across taxa.

Functional *CYP8B1* genes occur widely among advanced chordates, including mammals and birds, and are also present in numerous actinopterygian fishes.^{135,138} However, because genomic data remain incomplete for several vertebrate groups, the evolutionary status of *CYP8B1* in amphibians and reptiles is still uncertain. Changes in diet, particularly shifts in lipid composition, are considered the principal selective forces underlying degeneration or loss of *CYP8B1* during speciation.^{135,138,139} In several extant species, including *Heterocephalus glaber*, Elephantidae, and Sirenia, inactivation of *CYP8B1* eliminates CA production from the BA pool.^{135,139,140} In some cases, these loss-of-function events appear to predate the emergence of modern species. For example, genomic comparisons of African and Asian elephants reveal five shared frameshift mutations that inactivate *CYP8B1*.¹³⁵ The same mutations also occur in the genomes of the extinct woolly mammoth (*Mammuthus primigenius*) and the American mastodon (*Mammuthus americanus*), indicating that their common ancestor already lacked a functional gene.¹³⁵ Conversely, evidence of positive selection on *CYP8B1* has been reported in advanced cetaceans, where a single-base deletion near the termination codon extends the open reading frame by approximately 30 base pairs relative to other artiodactyls.^{135,139} Within Artiodactyla, evolutionary transitions in feeding ecology, from herbivory in Bovidae to carnivory in Balaenopteridae, coincide with intensified selection on *CYP8B1*, indirectly supporting observations of relaxed selective pressure in many herbivorous lineages. Critical evidence identifying the core role of dietary shifts in the selective pressure on CYP8B1 remains lacking, but the development-dependent downregulation of CYP8B1 in domestic pigs may imply that dietary structure-mediated selection pressure initially acts on regulatory sequences rather than amino acid-coding sequences. This is supported by the *CYP8B1* promoter exhibiting a

significantly higher evolutionary rate (relaxed selection, dN/dS $\approx$ 0.25–0.40) than housekeeping gene promoters (purifying selection, dN/dS = 0.08 $\pm$ 0.01).^{135,141–144} This is consistent with the notion that dietary shifts drive adaptive changes in *CYP8B1* expression by modifying regulatory sequences, while the amino acid-coding region remains highly conserved (sequence homology >70% across species) to maintain its core catalytic function in BA synthesis.^{133–135} Consequently, relaxed evolution of *CYP8B1* during early divergence likely occurs primarily in non-coding regulatory elements rather than through disruption of the coding sequence.

This progressive transcriptional repression of *CYP8B1* may be linked to potential DNA methylation and histone deacetylation events in its promoter region. For instance, CpG dinucleotides within the *CYP8B1* proximal promoter, particularly the binding site for hepatocyte nuclear factor 4 α , a known transcriptional activator of *CYP8B1*, might undergo methylation during piglet development.^{145–147} Such methylation could potentially block the recruitment of transcriptional activators, thereby contributing to chromatin condensation and subsequent gene silencing.^{148–150} Moreover, it is plausible that the enrichment of the repressive histone mark H3K27me3 (presumably mediated by the polycomb repressive complex 2) and reduced levels of the active histone mark H3K9ac may further reinforce the silent state of *CYP8B1* in adult hepatocytes.^{151–154} This speculation aligns with the well-documented role of these histone modifications in developmental gene silencing across mammalian species, although direct experimental validation in pigs is currently lacking.

CYP4A21 exhibits a very different evolutionary and functional profile. Most members of the CYP4A subfamily catalyze ω - or (ω -1)-hydroxylation of long-chain aliphatic compounds, including fatty acids and prostaglandins, and their genes therefore display both constitutive and inducible expression patterns.¹³⁴ Expression of CYP4A enzymes is highest in liver and kidney tissues, where induction can increase baseline transcription levels several-fold.¹⁵⁵ Among these enzymes, however, CYP4A21 shows a markedly distinct substrate specificity.¹²⁷ To date, *Sus scrofa* is the only known species in which a CYP4A enzyme catalyzes the hydroxylation of steroid substrates. Evolution of the CYP4A subfamily produced a specialized structural region capable of recognizing steroid nuclei before the domestication of *Sus scrofa*. Three amino-acid substitutions—Ala-314, Ala-315, and Ser-319—in a region otherwise conserved within the CYP4A family confer this unusual catalytic property.^{117,127} These substitutions enable CYP4A21 to insert an oxygen atom at the C_{6 α} position of taurine-conjugated CDCA, thereby generating taurine-HCA.

In domestic pigs, the fixation of *CYP4A21* and the distinctive regulatory architecture of *CYP8B1* illustrate how selection can drive adaptive change through functional substitution between genes. The emergence of *CYP4A21* activity is hypothesized to compensate for the decline of CA synthesis caused by reduced *CYP8B1* expression, effectively replacing CA with a novel trihydroxy BA—HCA—in the domestic pig BA pool.³⁹ High sequence homology between *CYP4A21* and other *CYP4* genes suggests that it originated from a common ancestral gene within the CYP4 subfamily, which also produced fatty-acid hydroxylases such as *CYP4A24*.⁴⁷ This evolutionary pattern is consistent with Ohno's hypothesis of gene duplication and neofunctionalization, whereby one gene copy retains the ancestral function while the other accumulates mutations that generate a novel biochemical role.¹⁵⁶ Domestication, particularly the shift in nutrient availability and feeding practices, may have promoted positive selection on this gene. Beyond coding-sequence modifications affecting substrate binding sites, mutations in non-coding regions likely contributed to divergence from the ancestral gene and to the emergence of its specialized metabolic function.⁴⁷

Promoter architecture also supports this interpretation. Promoters of cloned CYP4A genes typically lack canonical TATA boxes but contain multiple binding sites for the transcription factor Sp1.⁴⁷ Porcine *CYP4A* promoters also include YY1 binding motifs.⁴⁷ Such features resemble those of housekeeping genes, suggesting that CYP4A genes maintain basal transcriptional activity.¹⁵⁷ During gene differentiation, additional *cis*-regulatory elements appear to have been incorporated, enabling more precise regulatory control.³⁸ Variation within the 5'-flanking regions of these genes may therefore explain the distinctive spatial and temporal expression pattern of *CYP4A21* and provides an important framework for examining its evolutionary trajectory.

Nevertheless, direct evidence demonstrating whether *CYP4A21* originated specifically within Suidae, or whether its fixation in domestic pigs reflects selection associated with domestication-related dietary change, remains lacking.

Taken together, the shift in BA composition in *Sus scrofa* may offer a clear molecular example of the evolutionary principles governing early divergence. First, divergent selection associated with domestication-specific diets promotes targeted genetic remodeling, including regulatory modification of *CYP8B1* and neofunctionalization of *CYP4A21*. Second, the genetic architecture that emerges balances adaptation and continuity: HCA synthesis becomes established while ancestral metabolic pathways are not entirely eliminated. Third, metabolic specialization can arise before complete reproductive isolation. In this sense, the BA transition supports a “phenotype-first” model of divergence, in which ecological adaptation, here, the shift to energy-rich domestic diets, drives molecular restructuring prior to full genomic separation. Ongoing gene flow between wild and domestic pigs makes this system particularly valuable for examining such early evolutionary dynamics.⁴

PERSPECTIVE

The genetic diversity of *Sus scrofa domesticus*, in which marked phenotypic divergence emerged prior to complete reproductive isolation, makes the domestic pig an exceptional system for investigating the ecological and genetic processes that operate during the earliest stages of lineage divergence. Over the past two decades, population genomics has greatly advanced understanding of pig (*Sus scrofa*) domestication and early speciation. Genome-wide analyses have identified thousands of loci under selection, revealed extensive gene flow between wild boars and domestic breeds, and clarified the genetic basis of key phenotypic traits.^{11,14,21,48} Discoveries ranging from individual loci such as *NR6A1* and *LEPR* to broader breed architectures including HHGF and LHGF populations illustrate how artificial selection and natural introgression jointly shape genomic diversity. However, these approaches largely rely on bulk tissue data and therefore cannot resolve cell-type-specific regulatory mechanisms or fine-scale functional variation.

Emerging single-cell omics and CRISPR-Cas9 tools offer a route to dissect how regulatory and coding variants become established during early divergence under persistent gene flow. Integrating scRNA-seq, scATAC-seq, and epigenomic profiling across HHGF and LHGF breeds with functional perturbation in pig organoids could reveal whether cis-regulatory changes at loci such as *CYP4A21* initiate phenotypic innovation while coding mutations stabilize traits as gene flow declines. Complementary F2 mapping, ancient DNA analysis, and nutritional epigenome editing would further clarify how dietary shifts interact with genetic background to drive adaptive metabolic remodeling, positioning *Sus scrofa domesticus* as a powerful experimental system for testing fundamental hypotheses in speciation biology.

Taken together, *Sus scrofa domesticus* provides a unique bridge between domestication research and speciation biology. Its genomic diversity, ongoing gene flow with wild populations, and well-documented phenotypic variation create an experimentally tractable system for testing fundamental hypotheses about early lineage divergence. Integrating functional genomics, epigenetic profiling, ancient DNA analysis, and targeted genome editing will not only illuminate the evolutionary origins of domestic pig diversity but also offer broader insights into how new lineages emerge in nature, where selection and gene flow operate simultaneously. In this context, domestic species function as powerful natural experiments for addressing central questions in evolutionary biology.

REFERENCES

1. Wolf J.B. and Ellegren H. (2017). Making sense of genomic islands of differentiation in light of speciation. *Nat. Rev. Genet.* **18**:87–100. DOI:10.1038/nrg.2016.133
2. Johnson N.A. (2022). Speciation: Genomic sequence data and the biogeography of speciation. *Natl. Sci. Rev.* **9**:nwac294. DOI:10.1093/nsr/nwac294
3. Jarrett B.J.M., Downing P.A. and Svensson E.I. (2025). Meta-analysis reveals that phenotypic plasticity and divergent selection promote reproductive isolation during incipient speciation. *Nat. Ecol. Evol.* **9**:833–844. DOI:10.1038/s41559-025-02687-7
4. Diaz F., Allan C.W., Chen X., et al. (2022). Divergent evolutionary trajectories shape the postmating transcriptional profiles of conspecific and heterospecifically mated cactophilic *Drosophila* females. *Commun. Biol.* **5**:842. DOI:10.1038/s42003-022-03758-2

5. Fotopulosova V., Tanieli G., Fusek K., et al. (2024). A minimal hybrid sterility genome assembled by chromosome swapping between mouse subspecies (*Mus musculus*). *Mol. Biol. Evol.* **41**:msae211. DOI:10.1093/molbev/msae211
6. Ramírez O., Ojeda A., Tomás A., et al. (2009). Integrating Y-chromosome, mitochondrial, and autosomal data to analyze the origin of pig breeds. *Mol. Biol. Evol.* **26**:2061–72. DOI:10.1093/molbev/msp118
7. Mabry M.E., Rowan T.N., Pires J.C., et al. (2021). Feralization: Confronting the complexity of domestication and evolution. *Trends. Genet.* **37**:302–305. DOI:10.1016/j.tig.2021.01.005
8. Zhu Y., Li W., Yang B., et al. (2017). Signatures of selection and interspecies introgression in the genome of Chinese domestic pigs. *Genome. Biol. Evol.* **9**:2592–2603. DOI:10.1093/gbe/evx186
9. Ai H., Zhang M., Yang B., et al. (2021). Human-mediated admixture and selection shape the diversity on the modern swine (*Sus scrofa*) Y chromosomes. *Mol. Biol. Evol.* **38**:5051–5065. DOI:10.1093/molbev/msab230
10. Liu L., Bosse M., Megens H.J., et al. (2019). Genomic analysis on pygmy hog reveals extensive interbreeding during wild boar expansion. *Nat. Commun.* **10**:1992. DOI:10.1038/s41467-019-10017-2
11. Frantz L.A., Schraiber J.G., Madsen O., et al. (2015). Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nat. Genet.* **47**:1141–8. DOI:10.1038/ng.3394
12. Bosse M., Megens H.J., Madsen O., et al. (2014). Untangling the hybrid nature of modern pig genomes: A mosaic derived from biogeographically distinct and highly divergent *Sus scrofa* populations. *Mol. Ecol.* **23**:4089–102. DOI:10.1111/mec.12807
13. Zhang Z., Zhang Q., Xiao Q., et al. (2018). Distribution of runs of homozygosity in Chinese and Western pig breeds evaluated by reduced-representation sequencing data. *Anim. Genet.* **49**:579–591. DOI:10.1111/age.12730
14. Rubin C.J., Megens H.J., Martinez Barrio A., et al. (2012). Strong signatures of selection in the domestic pig genome. *Proc. Natl. Acad. Sci. USA* **109**:19529–36. DOI:10.1073/pnas.1217149109
15. Bovo S., Ribani A., Muñoz M., et al. (2020). Whole-genome sequencing of European autochthonous and commercial pig breeds allows the detection of signatures of selection for adaptation of genetic resources to different breeding and production systems. *Genet. Sel. Evol.* **52**:33. DOI:10.1186/s12711-020-00553-7
16. Revidatti M.A., Gama L.T., Martin Burriel I., et al. (2021). On the origins of American Criollo pigs: A common genetic background with a lasting Iberian signature. *PLoS One* **16**:e0251879. DOI:10.1371/journal.pone.0251879
17. Erven J.A.M., Mattiangeli V., Dreshaj M., et al. (2025). Archaeogenomic insights into commensalism and regional variation in pig management in Neolithic northwest Europe. *Proc. Natl. Acad. Sci. USA* **122**:e2410235122. DOI:10.1073/pnas.2410235122
18. Frantz L.A., Schraiber J.G., Madsen O., et al. (2013). Genome sequencing reveals fine scale diversification and reticulation history during speciation in *Sus*. *Genome. Biol.* **14**:R107. DOI:10.1186/gb-2013-14-9-r107
19. Bosse M., Megens H.J., Madsen O., et al. (2012). Regions of homozygosity in the porcine genome: Consequence of demography and the recombination landscape. *PLoS. Genet.* **8**:e1003100. DOI:10.1371/journal.pgen.1003100
20. Rowley-Conwy P. (2025). European Neolithic farmers interbred their domestic pigs with wild boar. *Proc. Natl. Acad. Sci. USA* **122**:e2504472122. DOI:10.1073/pnas.2504472122
21. Groenen M.A., Archibald A.L., Uenishi H., et al. (2012). Analyses of pig genomes provide insight into porcine demography and evolution. *Nature* **491**:393–8. DOI:10.1038/nature11622
22. Yang G., Ren J., Zhang Z., et al. (2009). Genetic evidence for the introgression of Western NR6A1 haplotype into Chinese Licha breed associated with increased vertebral number. *Anim. Genet.* **40**:247–50. DOI:10.1111/j.1365-2052.2008.01820.x
23. Larson G., Albarella U., Dobney K., et al. (2007). Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proc. Natl. Acad. Sci. USA* **104**:15276–81. DOI:10.1073/pnas.0703411104
24. Yang L., Yin H., Bai L., et al. (2024). Mapping and functional characterization of structural variation in 1060 pig genomes. *Genome. Biol.* **25**:116. DOI:10.1186/s13059-024-03253-3
25. Pérez-Montarelo D., Rodríguez M.C., Fernández A., et al. (2015). Haplotype diversity of porcine LEP and LEPR genes involved in growth and fatness regulation. *J. Appl. Genet.* **56**:525–533. DOI:10.1007/s13353-015-0284-7
26. Bosse M., Megens H.J., Frantz L.A., et al. (2014). Genomic analysis reveals selection for Asian genes in European pigs following human-mediated introgression. *Nat. Commun.* **5**:4392. DOI:10.1038/ncomms5392
27. Stanton D.W.G., Manin A., Evlin A., et al. (2026). Genomic and morphometric evidence for Austronesian-mediated pig translocation in the Pacific. *Science* **391**:eadv4963. DOI:10.1126/science.adv4963
28. Larson G., Dobney K., Albarella U., et al. (2005). Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science* **307**:1618–21. DOI:10.1126/science.1106927
29. Goedbloed D.J., Megens H.J., Van Hooft P., et al. (2013). Genome-wide single nucleotide polymorphism analysis reveals recent genetic introgression from domestic pigs into Northwest European wild boar populations. *Mol. Ecol.* **22**:856–66. DOI:10.1111/j.1365-294X.2012.05670.x

30. Amaral A.J., Megens H.J., Crooijmans R.P., et al. (2008). Linkage disequilibrium decay and haplotype block structure in the pig. *Genetics* **179**:569–79. DOI:10.1534/genetics.107.084277
31. Ma Y.F., Han X.M., Huang C.P., et al. (2019). Population genomics analysis revealed origin and high-altitude adaptation of Tibetan pigs. *Sci. Rep.* **9**:11463. DOI:10.1038/s41598-019-47711-6
32. Tong X., Chen D., Hu J., et al. (2023). Accurate haplotype construction and detection of selection signatures enabled by high quality pig genome sequences. *Nat. Commun.* **14**:5126. DOI:10.1038/s41467-023-40434-3
33. Ding J., Cui X., Wang X., et al. (2024). Multi-omics analysis of gut microbiota and metabolites reveals contrasting profiles in domestic pigs and wild boars across urban environments. *Front. Microbiol.* **15**:1450306. DOI:10.3389/fmicb.2024.1450306
34. Zong W., Chen L., Zhang D., et al. (2025). Two telomere-to-telomere pig genome assemblies and pan-genome analyses provide insights into genomic structural landscape and genetic adaptations. *Imeta* **4**:e70013. DOI:10.1002/imt.2.70013
35. Rong B., Zhang Q., Zhang X., et al. (2024). Hyocholic acid: A novel therapeutic strategy for metabolic syndrome. *Innov. Life* **2**:100093. DOI:10.59717/j.xinn-life.2024.100093
36. Zheng X., Chen T., Jiang R., et al. (2021). Hyocholic acid species improve glucose homeostasis through a distinct TGR5 and FXR signaling mechanism. *Cell. Metab.* **33**:791–803. DOI:10.1016/j.cmet.2020.11.017
37. Fan M., Yang Z., Jin L., et al. (2025). Differential roles of individual bile acid in physiology and disease. *Pharmacol. Res.* **218**:107845. DOI:10.1016/j.phrs.2025.107845
38. Lundell K. and Wikvall K. (2003). Gene structure of pig sterol 12 α -hydroxylase (CYP8B1) and expression in fetal liver: Comparison with expression of taurochenodeoxycholic acid 6 α -hydroxylase (CYP4A21). *Biochim. Biophys. Acta* **1634**:86–96. DOI:10.1016/j.bbaplp.2003.09.002
39. Lundell K. and Wikvall K. (2008). Species-specific and age-dependent bile acid composition: Aspects on CYP8B and CYP4A subfamilies in bile acid biosynthesis. *Curr. Drug. Metab.* **9**:323–31. DOI:10.2174/138920008784220574
40. Morita S., Yamasaki H. and Ogura M. (1970). Bile acids of wild boar, *Sus scrofa leucomystax* (Temminck et Schlegel). *Yonago. Acta. Med.* **14**:82–6. DOI:10.3109/00365517009046222
41. Boström H. (1986). Characterization of 6 α -hydroxylation of taurochenodeoxycholic acid in pig liver. *J. Lipid. Res.* **27**:807–12. DOI:10.1016/j.ygyno.2004.08.046
42. Patience J.F., Rossoni-Serão M.C. and Gutiérrez N.A. (2015). A review of feed efficiency in swine: Biology and application. *J. Anim. Sci. Biotechnol.* **6**:33. DOI:10.1186/s40104-015-0031-2
43. Robeson M.S. II, Khanipov K., Golovko G., et al. (2017). Assessing the utility of metabarcoding for diet analyses of the omnivorous wild pig (*Sus scrofa*). *Ecol. Evol.* **8**:185–196. DOI:10.1002/ece3.3638
44. Heng Y.C., Chua J.H.X., Silvaraju S., et al. (2026). Metagenomic insights into the global wild boar faecal microbiome reveal novel taxa and carbohydrate degraders distinguishing wild and domesticated *Sus*. *Microbiome* **14**:123. DOI:10.1186/s40168-026-02392-y
45. Canright V.R., Piaggio A.J. and Beasley J.C. (2024). Invasive wild pig (*Sus scrofa*) diets on barrier islands in the southeastern United States. *Pest. Manag. Sci.* **80**:3901–3911. DOI:10.1002/ps.8093
46. Zhang B., Sun Z., Yu Z., et al. (2022). Transcriptome and targeted metabolome analysis provide insights into bile acids' new roles and mechanisms on fat deposition and meat quality in lamb. *Food. Res. Int.* **162**:111941. DOI:10.1016/j.foodres.2022.111941
47. Lundell K. (2004). The porcine taurochenodeoxycholic acid 6 α -hydroxylase (CYP4A21) gene: Evolution by gene duplication and gene conversion. *Biochem. J.* **378**:1053–8. DOI:10.1042/BJ20031657
48. Ai H., Fang X., Yang B., et al. (2015). Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nat. Genet.* **47**:217–25. DOI:10.1038/ng.3199
49. Mikawa S., Morozumi T., Shimanuki S.I., et al. (2007). Fine mapping of a swine quantitative trait locus for number of vertebrae and analysis of an orphan nuclear receptor, germ cell nuclear factor (NR6A1). *Genome. Res.* **17**:586–593. DOI:10.1101/gr.6085507
50. Herrero-Medrano J.M., Megens H.J., Groenen M.A., et al. (2013). Conservation genomic analysis of domestic and wild pig populations from the Iberian Peninsula. *BMC. Genet.* **14**:106. DOI:10.1186/1471-2156-14-106
51. Zorc M., Škorput D., Gvozdanović K., et al. (2022). Genetic diversity and population structure of six autochthonous pig breeds from Croatia, Serbia, and Slovenia. *Genet. Sel. Evol.* **54**:30. DOI:10.1186/s12711-022-00718-6
52. Atrian-Afiani F., Berger B., Draxl C., et al. (2023). Selective sweeps in the Austrian Turpolfje and other commercial pig populations. *Animals (Basel)* **13**:3749. DOI:10.3390/ani13243749
53. Ishikawa K., Doneva R., Raichev E.G., et al. (2021). Population genetic structure and diversity of the East Balkan Swine (*Sus scrofa*) in Bulgaria, revealed by mitochondrial DNA and microsatellite analyses. *Anim. Sci. J.* **92**:e13630. DOI:10.1111/asj.13630
54. Liu T., Ji D., Li X., et al. (2024). Population genetics reveals new introgression in the nucleus herd of min pigs. *Genes Genomics* **46**:389–398. DOI:10.1007/s13258-024-01490-1
55. Burgos C., Latorre P., Altarriba J., et al. (2015). Allelic frequencies of NR6A1 and VRTN, two genes that affect vertebrae number in diverse pig breeds: A study of the effects of the VRTN insertion on phenotypic traits of a Duroc $\times$ Landrace-Large White cross. *Meat. Sci.* **100**:150–155. DOI:10.1016/j.meatsci.2014.09.143
56. Muñoz M., Bozzi R., García F., et al. (2018). Diversity across major and candidate genes in European local pig breeds. *PLoS One* **13**:e0207475. DOI:10.1371/journal.pone.0207475
57. Molinero E., Pena R.N., Estany J., et al. (2025). Unravelling novel and closely linked association signals for fat-related traits in pigs using prioritised variants from whole-genome sequence data. *Animal* **19**:101496. DOI:10.1016/j.animal.2025.101496
58. Poklukur K., Mestre C., Škrlep M., et al. (2023). A meta-analysis of genetic and phenotypic diversity of European local pig breeds reveals genomic regions associated with breed differentiation for production traits. *Genet. Sel. Evol.* **55**:88. DOI:10.1186/s12711-023-00858-3
59. Huang Y., Zhou L., Zhang J., et al. (2020). A large-scale comparison of meat quality and intramuscular fatty acid composition among three Chinese indigenous pig breeds. *Meat. Sci.* **168**:108182. DOI:10.1016/j.meatsci.2020.108182
60. Jørgensen D. (2024). The Medieval Pig. *Woodbridge: The Boydell Press*. DOI:10.1017/9781805432456
61. Wang L., Li K., Yang G., et al. (2011). Animal genetic resources in China: Pigs (ed. China National Commission of Animal Genetic Resources). *China Agricultural Press*. <http://course.syt126.com/?m=course&p=detail&cid=8&kid=341>
62. Innan H. and Kim Y. (2004). Pattern of polymorphism after strong artificial selection in a domestication event. *Proc. Natl. Acad. Sci. USA* **101**:10667–72. DOI:10.1073/pnas.0401720101
63. Samuk K., Owens G.L., Delmore K.E., et al. (2017). Gene flow and selection interact to promote adaptive divergence in regions of low recombination. *Mol. Ecol.* **26**:4378–4390. DOI:10.1111/mec.14226
64. Burbrink F.T., DeBaun D., Foley N.M., et al. (2025). Recombination-aware phylogenomics. *Trends. Ecol. Evol.* **40**:900–912. DOI:10.1016/j.tree.2025.06.011
65. Han Y., Zhao Z., Stanton D.W.G., et al. (2025). Ancient genomics reveals the origin, dispersal, and human management of East Asian domestic pigs. *Mol. Biol. Evol.* **42**:msaf214. DOI:10.1093/molbev/msaf214
66. Barmentlo N.W.G., Meirmans P.G., Stiver W.H., et al. (2024). Natural selection on feralization genes contributed to the invasive spread of wild pigs throughout the United States. *Mol. Ecol.* **33**:e17383. DOI:10.1111/mec.17383
67. Zhang C. and Plastow G. (2011). Genomic diversity in pig (*Sus scrofa*) and its comparison with human and other livestock. *Curr. Genomics* **12**:138–46. DOI:10.2174/138920211795564386
68. Chen Z., Zhou M., Sun Y., et al. (2024). Exploration of genome-wide recombination rate variation patterns at different scales in pigs. *Animals (Basel)* **14**:1345. DOI:10.3390/ani14091345
69. Ribani A., Taurisano V., Karatsidi D., et al. (2023). Signatures of admixture and genetic uniqueness in the autochthonous Greek black pig breed deduced from gene polymorphisms affecting domestication-derived traits. *Animals (Basel)* **13**:1763. DOI:10.3390/ani13111763
70. Wu G.S., Yao Y.G., Qu K.X., et al. (2007). Population phylogenomic analysis of mitochondrial DNA in wild boars and domestic pigs revealed multiple domestication events in East Asia. *Genome. Biol.* **8**:R245. DOI:10.1186/gb-2007-8-11-r245
71. Bosse M., Megens H.J., Madsen O., et al. (2015). Using genome-wide measures of coancestry to maintain diversity and fitness in endangered and domestic pig populations. *Genome. Res.* **25**:970–81. DOI:10.1101/gr.187039.114
72. Touma S., Fusho K. and Arakawa A. et al. (2025). Assessment of genetic diversity and inbreeding in the Okinawa indigenous Agui pig through pedigree analysis. *Anim. Biosci.* **38**:1105–1115. DOI:10.5713/ab.24.0646
73. Qiao R., Li X., Madsen O., et al. (2023). Potential selection for lipid kinase activity and spermatogenesis in Henan native pig breeds and growth shaping by introgression of European genes. *Genet. Sel. Evol.* **55**:64. DOI:10.1186/s12711-023-00841-y
74. Kim J., Cho S., Caetano-Anolles K., et al. (2015). Genome-wide detection and characterization of positive selection in Korean Native Black Pig from Jeju Island. *BMC. Genet.* **16**:3. DOI:10.1186/s12863-014-0160-1
75. Yang R., Jin S., Fang S., et al. (2024). Genetic introgression from commercial European pigs to the indigenous Chinese Lijiang breed and associated changes in phenotypes. *Genet. Sel. Evol.* **56**:24. DOI:10.1186/s12711-024-00893-8
76. Patterson N., Price A.L. and Reich D. (2006). Population structure and eigenanalysis. *PLoS. Genet.* **2**:e190. DOI:10.1371/journal.pgen.0020190
77. Lawson D.J., van Dorp L. and Falush D. (2018). A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nat. Commun.* **9**:3258. DOI:10.1038/s41467-018-05257-7
78. Feng X., Diao S., Liu Y., et al. (2023). Exploring the mechanism of artificial selection signature in Chinese indigenous pigs by leveraging multiple bioinformatics database tools. *BMC Genomics* **24**:743. DOI:10.1186/s12864-023-09848-7
79. Hongo H., Ishiguro N., Watanobe T., et al. (2002). Variation in mitochondrial DNA of Vietnamese pigs: Relationships with Asian domestic pigs and Ryukyu wild boars. *Zool. Sci.* **19**:1329–35. DOI:10.2108/zsj.19.1329

80. Cao M., Yuan T., Li D., et al. (2025). Whole-genome resequencing reveals positive selection and introgression signatures and genetic loci associated with early puberty traits in Chinese indigenous pigs. *Genet. Sel. Evol.* **57**:29. DOI:10.1186/s12711-025-00975-1
81. Li D., Wang Y., Yuan T., et al. (2024). Pangenome and genome variation analyses of pigs unveil genomic facets for their adaptation and agronomic characteristics. *Imeta* **3**:e257. DOI:10.1002/imt2.257
82. Ishihara S., Arakawa A., Ba N.V., et al. (2023). Population structure of Vietnamese pigs using mitochondrial DNA. *Anim. Sci. J.* **94**:e13875. DOI:10.1111/asj.13875
83. A V., Kumar A., Mahala S., et al. (2024). Revelation of genetic diversity and genomic footprints of adaptation in Indian pig breeds. *Gene* **893**:147950. DOI:10.1016/j.gene.2023.147950
84. Du H., Zhou L., Liu Z., et al. (2024). The 1000 Chinese indigenous pig genomes project provides insights into the genomic architecture of pigs. *Nat. Commun.* **15**:10137. DOI:10.1038/s41467-024-54471-z
85. Li Z., Wei S., Li H., et al. (2017). Genome-wide genetic structure and differentially selected regions among Landrace, Erhualian, and Meishan pigs using specific-locus amplified fragment sequencing. *Sci. Rep.* **7**:10063. DOI:10.1038/s41598-017-09969-6
86. Larson G., Liu R., Zhao X., et al. (2010). Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. *Proc. Natl. Acad. Sci. USA* **107**:7686–91. DOI:10.1073/pnas.0912264107
87. Li Z., Liu X., Wang C., et al. (2023). The pig pangenome provides insights into the roles of coding structural variations in genetic diversity and adaptation. *Genome. Res.* **33**:1833–1847. DOI:10.1101/gr.277638.122
88. Yuan H., Li C., Zhao S., et al. (2025). Unraveling the genetic diversity and adaptive traits of laboratory pig breeds within the perspective of whole - genome resequencing. *BMC Genomics* **26**:604. DOI:10.1186/s12864-025-11790-9
89. Wen H., Blackburn H.D., Mulim H.A., et al. (2025). Characterization of genomic diversity and population structure of worldwide Duroc subpopulations and other pig breeds. *Genet. Sel. Evol.* **58**:1. DOI:10.1186/s12711-025-01017-6
90. Kim K.S., Larsen N., Short T., et al. (2000). A missense variant of the porcine melanocortin-4 receptor (MC4R) gene is associated with fatness, growth, and feed intake traits. *Mamm. Genome* **11**:131–5. DOI:10.1007/s003350010025
91. Jokubka R., Maak S., Kerziene S., et al. (2006). Association of a melanocortin 4 receptor (MC4R) polymorphism with performance traits in Lithuanian White pigs. *J. Anim. Breed. Genet.* **123**:17–22. DOI:10.1111/j.1439-0388.2006.00559.x
92. Fontanesi L., Schiavo G., Scotti E., et al. (2015). A retrospective analysis of allele frequency changes of major genes during 20 years of selection in the Italian Large White pig breed. *J. Anim. Breed. Genet.* **132**:239–46. DOI:10.1111/jbg.12127
93. Liu Y., Xu Y., Li G., et al. (2025). Local climate adaptation in Chinese indigenous pig genomes. *Animals* **15**:2412. DOI:10.3390/ani15162412
94. Liu C., Ao N., Liang Y., et al. (2024). Analysis of cis-regulatory changes underlying phenotype divergence shaped by domestication in pigs. *Front. Genet.* **15**:1421859. DOI:10.3389/fgene.2024.1421859
95. Luo Y.B., Huang N., Zha C.W., et al. (2026). Telomere-to-telomere genome assembly of a male pig provides insight into population structure and selection for body stature. *Nat. Genet.* **58**:195–205. DOI:10.1038/s41588-025-02433-6
96. Wang Y., Feng R., Hu X., et al. (2025). Latitude-driven adaptive evolution of body length in diverse pig breeds across China. *BMC. Biol.* **23**:150. DOI:10.1186/s12915-025-02265-0
97. Wang Z., Song B., Yao J., et al. (2024). Whole-genome analysis reveals distinct adaptation signatures to diverse environments in Chinese domestic pigs. *J. Anim. Sci. Biotechnol.* **15**:97. DOI:10.1186/s40104-024-01053-0
98. Gao Y., Feng X., Diao S., et al. (2024). Deciphering genetic characteristics of South China and North China indigenous pigs through selection signatures. *BMC Genomics* **25**:1191. DOI:10.1186/s12864-024-11119-y
99. Du H., Liu Z., Lu SY., et al. (2024). Genomic evidence for human-mediated introgressive hybridization and selection in the developed breed. *BMC Genomics* **25**:331. DOI:10.1186/s12864-024-10259-5
100. Liu C., Huang R., Su G., et al. (2023). Introgression of pigs in Taihu Lake region possibly contributed to the improvement of fertility in Danish Large White pigs. *BMC Genomics* **24**:733. DOI:10.1186/s12864-023-09860-x
101. Fontanesi L. (2022). Invited review: Genetics and genomics of pigmentation variability in pigs: A review. *Livest. Sci.* **265**:105079. DOI:10.1016/j.livsci.2022.105079
102. Giovannini S., Strillacci M.G., Bagnato A., et al. (2023). Genetic and phenotypic characteristics of belted pig breeds: A review. *Animals (Basel)* **13**:3072. DOI:10.3390/ani13193072
103. Fang M., Larson G., Ribeiro H.S., et al. (2009). Contrasting mode of evolution at a coat color locus in wild and domestic pigs. *PLoS. Genet.* **5**:e1000341. DOI:10.1371/journal.pgen.1000341
104. Giuffra E., Evans G., Törnsten A., et al. (1999). The Belt mutation in pigs is an allele at the Dominant white (I/KIT) locus. *Mamm. Genome* **10**:1132–6. DOI:10.1007/s003359901178
105. Hirooka H., de Koning D.J., van Arendonk J.A., et al. (2002). Genome scan reveals new coat color loci in exotic pig cross. *J. Hered.* **93**:1–8. DOI:10.1093/jhered/93.1.1
106. Muñoz M., Bozzi R., García-Casco J., et al. (2019). Genomic diversity, linkage disequilibrium and selection signatures in European local pig breeds assessed with a high density SNP chip. *Sci. Rep.* **9**:13546. DOI:10.1038/s41598-019-49830-6
107. Twining C.W., Bernhardt J.R., Derry A.M., et al. (2021). The evolutionary ecology of fatty-acid variation: Implications for consumer adaptation and diversification. *Ecol. Lett.* **24**:1709–1731. DOI:10.1111/ele.13771
108. Haslewood G.A. (1967). Bile salt evolution. *J. Lipid. Res.* **8**:535–50. DOI:10.1016/S0022-2275(20)38873-8
109. Sasaki T. (1966). Stero-bile acids and bile alcohols: LXXXII. Comparative studies on the bile salts of fishes by thin layer chromatography. *J. Biochem.* **60**:56–62. DOI:10.1093/oxfordjournals.jbchem.a128399
110. Hofmann A.F. and Roda A. (1984). Physicochemical properties of bile acids and their relationship to biological properties: An overview of the problem. *J. Lipid. Res.* **25**:1477–89. DOI:10.1016/S0022-2275(20)34421-7
111. Watt S.M. and Simmonds W.J. (1984). Effects of four taurine-conjugated bile acids on mucosal uptake and lymphatic absorption of cholesterol in the rat. *J. Lipid. Res.* **25**:448–55. DOI:10.1016/S0022-2275(20)37796-8
112. Heuman D.M., Hylemon P.B. and Vlahcevic Z.R. (1989). Regulation of bile acid synthesis. III. Correlation between biliary bile salt hydrophobicity index and the activities of enzymes regulating cholesterol and bile acid synthesis in the rat. *J. Lipid. Res.* **30**:1161–71. DOI:10.1016/S0022-2275(20)38276-6
113. Sheriha G.M., Waller G.R., Chan T., et al. (1968). Composition of bile acids in ruminants. *Lipids* **3**:72–8. DOI:10.1007/BF02530972
114. Martínez-Álvarez M., Mattock J., Auffret M., et al. (2022). Microbiome-driven breeding strategy potentially improves beef fatty acid profile benefiting human health and reduces methane emissions. *Microbiome* **10**:166. DOI:10.1186/s40168-022-01352-6
115. Lin L., Lai Z., Yang H., et al. (2023). Genome-centric investigation of bile acid metabolizing microbiota of dairy cows and associated diet-induced functional implications. *ISME J.* **17**:172–184. DOI:10.1038/s41396-022-01333-5
116. Zheng K., Guo L., Cao Y., et al. (2024). High-concentrate diet decreases lamb fatty acid contents by regulating bile acid composition. *Food. Chem. X.* **24**:101871. DOI:10.1016/j.fochx.2024.101871
117. Araya Z., Hellman U. and Hansson R. (1995). Characterisation of taurochenodeoxycholic acid 6 alpha-hydroxylase from pig liver microsomes. *Eur. J. Biochem.* **231**:855–61. DOI:10.1111/j.1432-1033.1995.0855d.x
118. Takahashi S., Fukami T., Masuo Y., et al. (2016). Cyp2c70 is responsible for the species difference in bile acid metabolism between mice and humans. *J. Lipid. Res.* **57**:2130–2137. DOI:10.1194/jlr.M071183
119. Tavoloni N. (1986). Bile acid structure and bile formation in the guinea pig. *Biochim. Biophys. Acta.* **879**:186–201. DOI:10.1016/0005-2760(86)90102-5
120. Li J. and Dawson P.A. (2019). Animal models to study bile acid metabolism. *Biochim. Biophys. Acta. Mol. Basis. Dis.* **1865**:895–911. DOI:10.1016/j.bbdis.2018.05.011
121. van Hees H.M.J., Ballari S.A., Dieste-Pérez L., et al. (2023). Diet and stomach characteristics of feral piglets (Sus scrofa): Implications for farmed piglets. *J. Anim. Physiol. Anim. Nutr. (Berl.)* **107**:529–540. DOI:10.1111/jpn.13726
122. Canright V.R., Piaggio A.J., Chinn S.M., et al. (2023). DNA metabarcoding reveals consumption of diverse community of amphibians by invasive wild pigs (Sus scrofa) in the southeastern United States. *Sci. Rep.* **13**:20889. DOI:10.1038/s41598-023-48139-9
123. Russo G., Danieli P.P., Primi R., et al. (2017). Stable isotopes in tissues discriminate the diet of free-living wild boar from different areas of central Italy. *PLoS One* **12**:e0183333. DOI:10.1371/journal.pone.0183333
124. Castillo-Contreras R., Mentaberre G., Fernandez Aguilar X., et al. (2021). Wild boar in the city: Phenotypic responses to urbanisation. *Sci. Total. Environ.* **773**:145593. DOI:10.1016/j.scitotenv.2021.145593
125. Sayin S.I., Wahlström A., Felin J., et al. (2013). Gut microbiota regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally occurring FXR antagonist. *Cell. Metab.* **17**:225–35. DOI:10.1016/j.cmet.2013.01.003
126. Lauridsen C. (2020). Effects of dietary fatty acids on gut health and function of pigs pre- and post-weaning. *J. Anim. Sci.* **98**:skaa086. DOI:10.1093/jas/skaa086
127. Lundell K., Hansson R. and Wikvall K. (2001). Cloning and expression of a pig liver taurochenodeoxycholic acid 6alpha-hydroxylase (CYP4A21): A novel member of the CYP4A subfamily. *J. Biol. Chem.* **276**:9606–12. DOI:10.1074/jbc.M006584200
128. Pellicoro A., van den Heuvel F.A., Geuken M., et al. (2007). Human and rat bile acid-CoA: Amino acid N-acyltransferase are liver-specific peroxisomal enzymes: Implications for intracellular bile salt transport. *Hepatology* **45**:340–8. DOI:10.1002/hep.21528
129. Smith Z., Ryerson D. and Kemper J.K. (2013). Epigenomic regulation of bile acid metabolism: Emerging role of transcriptional cofactors. *Mol. Cell. Endocrinol.* **368**:59–70. DOI:10.1016/j.mce.2012.04.008
130. Yin C., Liu X., Fang W., et al. (2025). Hyocholic acid species as the key modulator for cecal epithelial homeostasis in low-birth-weight piglets. *Nutrients* **17**:3415. DOI:10.3390/nu17213415
131. Qi D., Zheng T., Yang M., et al. (2024). Bile acid composition and transcriptome analysis of the liver and small intestine in different species. *Metabolites* **14**:451. DOI:10.3390/metabo14080451
132. Yamada S., Kawaguchi H., Yamada T., et al. (2017). Cholic acid enhances visceral adiposity, atherosclerosis and nonalcoholic fatty liver disease in microminipigs. *J. Atheroscler. Thromb.* **24**:1150–1166. DOI:10.5551/jat.39909

133. Gáfvels M., Olin M., Chowdhary B.P., et al. (1999). Structure and chromosomal assignment of the sterol 12alpha-hydroxylase gene (CYP8B1) in human and mouse: Eukaryotic cytochrome P-450 gene devoid of introns. *Genomics* **56**:184–196. DOI:10.1006/geno.1998.5606
134. Simpson A.E. (1997). The cytochrome P450 4 (CYP4) family. *Gen. Pharmacol.* **28**:351–9. DOI:10.1016/s0306-3623(96)00246-7
135. Shinde S.S., Teekas L., Sharma S., et al. (2019). Signatures of relaxed selection in the CYP8B1 gene of birds and mammals. *J. Mol. Evol.* **87**:209–220. DOI:10.1007/s00239-019-09903-6
136. Dehal P., Predki P., Olsen A.S., et al. (2001). Human chromosome 19 and related regions in mouse: Conservative a4d lineage-specific evolution. *Science* **293**:104–11. DOI:10.1126/science.1060310
137. Tennyson C.N., Klamut H.J. and Worton R.G. (1995). The human dystrophin gene requires 16 hours to be transcribed and is cotranscriptionally spliced. *Nat. Genet.* **9**:184–90. DOI:10.1038/ng0295-184
138. Yan B., Sun Y., Fu K., et al. (2023). Effects of glyphosate exposure on gut-liver axis: Metabolomic and mechanistic analysis in grass carp (*Ctenopharyngodon idellus*). *Sci. Total. Environ.* **902**:166062. DOI:10.1016/j.scitotenv.2023.166062
139. Endo Y., Kamei K.I. and Inoue-Murayama M. (2018). Genetic signatures of lipid metabolism evolution in Cetacea since the divergence from terrestrial ancestor. *J. Evol. Biol.* **31**:1655–1665. DOI:10.1111/jeb.13361
140. Sharma V. and Hiller M. (2018). Loss of enzymes in the bile acid synthesis pathway explains differences in bile composition among mammals. *Genome. Biol. Evol.* **10**:3211–3217. DOI:10.1093/gbe/evy243
141. Zhang M. and Chiang J.Y. (2001). Transcriptional regulation of the human sterol 12alpha-hydroxylase gene (CYP8B1): Roles of hepatocyte nuclear factor 4alpha in mediating bile acid repression. *J. Biol. Chem.* **276**:41690–9. DOI:10.1074/jbc.M105117200
142. Arthur R.K. and Ruvinsky I. (2011). Evidence that purifying selection acts on promoter sequences. *Genetics* **189**:1121–6. DOI:10.1534/genetics.111.133637
143. She X., Rohl C.A., Castle J.C., et al. (2009). Definition, conservation and epigenetics of housekeeping and tissue-enriched genes. *BMC Genomics* **10**:269. DOI:10.1186/1471-2164-10-269
144. Wei K., Ma L. and Zhang T. (2019). Characterization of gene promoters in pig: Conservative elements, regulatory motifs and evolutionary trend. *PeerJ* **7**:e7204. DOI:10.7717/peerj.7204
145. Zhang M. and Chiang J.Y. (2001). Transcriptional regulation of the human sterol 12alpha-hydroxylase gene (CYP8B1): Roles of hepatocyte nuclear factor 4alpha in mediating bile acid repression. *J. Biol. Chem.* **276**:41690–9. DOI:10.1074/jbc.M105117200
146. Maj M.A., Burrin D.G. and Manjarin R. (2023). Decreased FXR agonism in the bile acid pool is associated with impaired FXR signaling in a pig model of pediatric NAFLD. *Biomedicines* **11**:3303. DOI:10.3390/biomedicines11123303
147. Schiöth H.B., Boström A., Murphy S.K., et al. (2016). A targeted analysis reveals relevant shifts in the methylation and transcription of genes responsible for bile acid homeostasis and drug metabolism in non-alcoholic fatty liver disease. *BMC Genomics* **17**:462. DOI:10.1186/s12864-016-2814-z
148. Sohn H.A., Go H., An T.H., et al. (2026). DNMT1 Facilitates the progression of MASLD by impeding transcription mediated by HNF4a and PPARa. *Clin. Mol. Hepatol.* DOI: 10.3350/cmh.2025.1099.
149. Leighton G. and Williams D.C. Jr. (2020). The methyl-CpG-binding domain 2 and 3 proteins and formation of the nucleosome remodeling and deacetylase complex. *J. Mol. Biol.* **432**:1624–1639. DOI:10.1016/j.jmb.2019.10.007
150. Kaluscha S., Domcke S., Wirbelauer C., et al. (2022). Evidence that direct inhibition of transcription factor binding is the prevailing mode of gene and repeat repression by DNA methylation. *Nat. Genet.* **54**:1895–1906. DOI:10.1038/s41588-022-01241-6
151. Kim D.H., Kwon S., Byun S., et al. (2016). Critical role of RanBP2-mediated SUMOylation of Small Heterodimer Partner in maintaining bile acid homeostasis. *Nat. Commun.* **7**:12179. DOI:10.1038/ncomms12179
152. Mao W., Ge X., Chen Q., et al. (2025). Epigenetic mechanisms in the transcriptional regulation of circadian rhythm in mammals. *Biology (Basel)*. **14**:42. DOI:10.3390/biology14010042
153. Zhang Y., Lin Y., Cai H., et al. (2025). EZH2 confers lenvatinib resistance in hepatocellular carcinoma by suppressing ACSL1-Mediated ferroptosis. *BMC Cancer* **25**:1638. DOI:10.1186/s12885-025-15086-9
154. Choi J., Ryoo Z.Y., Cho D.H., et al. (2021). Trans-tail regulation-mediated suppression of cryptic transcription. *Exp. Mol. Med.* **53**:1683–1688. DOI:10.1038/s12276-021-00711-x
155. Sharma R., Lake B.G., Foster J., et al. (1988). Microsomal cytochrome P-452 induction and peroxisome proliferation by hypolidaemic agents in rat liver, k mechanistic interrelationship. *Biochem. Pharmac.* **37**:1193–1201. DOI:10.1016/0006-2952(88)90770-8
156. Long M. (2001). Evolution of novel genes. *Curr. Opin. Genet. Dev.* **11**:673–80. DOI:10.1016/S0959-437X(00)00252-5
157. Lee J.S., Galvin K.M. and Shi Y. (1993). Evidence for physical interaction between the zinc-finger transcription factors YY1 and Sp1. *Proc. Natl. Acad. Sci. USA* **90**:6145–9. DOI:10.1073/pnas.90.13.6145

FUNDING AND ACKNOWLEDGMENTS

This work was supported by the grants from the National Key R&D Program of China (2021YFD1200303), National Nature Science Foundation of China's Heilongjiang province (LH2022C027), Scholars Program of Northeast Agricultural University (54916212), Natural Science Foundation of China (32472870). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

AUTHOR CONTRIBUTIONS

Every author contributed to this work. B.R. designed this review. B.R., N.N., W.Z., N.Z., Y.P., R.Z., and X.L. were responsible for collecting the relative data. B.R. N.N., W.Z., X.M., and S.Z. managed statistical data analysis. B.R. and Z.S. were responsible for the manuscript writing. X.Q., L.Z., D.L., B.R., and Y.Y. financed this project. All authors contributed to the manuscript and approved the final version.

DECLARATION OF INTERESTS

The authors declare no competing interests.