Received: 30 January 2024 | Revised: 23 April 2024 | Accepted: 30 April 2024

DOI: 10.1111/mec.17383

ORIGINAL ARTICLE

MOLECULAR ECOLOGY WILEY

Natural selection on feralization genes contributed to the invasive spread of wild pigs throughout the United States

Niek W. G. Barmentlo[1,2] | Patrick G. Meirmans[2] | William H. Stiver[3] | Joseph G. Yarkovich[3] | Blake E. McCann[4] | Antoinette J. Piaggio[5] | Dominic Wright[6] | Timothy J. Smyser[5] | Mirte Bosse[1,7]

[1]Section Ecology & Evolution, Amsterdam Institute for Life and Environment (A-LIFE), Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

[2]Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands

[3]Great Smoky Mountains National Park, Gatlinburg, Tennessee, USA

[4]Theodore Roosevelt National Park, Medora, North Dakota, USA

[5]USDA APHIS WS National Wildlife Research Center, Fort Collins, Colorado, USA

[6]Department of Physics, Chemistry and Biology, Linköping University, Linköping, Sweden

[7]Wageningen University & Research – Animal Breeding and Genomics, Wageningen, The Netherlands

Correspondence
Timothy J. Smyser, USDA APHIS WS National Wildlife Research Center, Fort Collins, CO, USA.
Email: timothy.j.smyser@usda.gov

Mirte Bosse and Niek W. G. Barmentlo, Section Ecology & Evolution, Amsterdam Institute for Life and Environment (A-LIFE), Vrije Universiteit Amsterdam, Amsterdam, The Netherlands.
Email: m.bosse@vu.nl and n.w.g.barmentlo@vu.nl

Funding information
USDA APHIS National Feral Swine Damage Management Program

Handling Editor: David Coltman

Abstract

Despite a long presence in the contiguous United States (US), the distribution of invasive wild pigs (*Sus scrofa* × *domesticus*) has expanded rapidly since the 1980s, suggesting a more recent evolutionary shift towards greater invasiveness. Contemporary populations of wild pigs represent exoferal hybrid descendants of domestic pigs and European wild boar, with such hybridization expected to enrich genetic diversity and increase the adaptive potential of populations. Our objective was to characterize how genetic enrichment through hybridization increases the invasiveness of populations by identifying signals of selection and the ancestral origins of selected loci. Our study focused on invasive wild pigs within Great Smoky Mountains National Park, which represents a hybrid population descendent from the admixture of established populations of feral pigs and an introduction of European wild boar to North America. Accordingly, we genotyped 881 wild pigs with multiple high-density single-nucleotide polymorphism (SNP) arrays. We found 233 markers under putative selection spread over 79 regions across 16 out of 18 autosomes, which contained genes involved in traits affecting feralization. Among these, genes were found to be related to skull formation and neurogenesis, with two genes, TYRP1 and TYR, also encoding for crucial melanogenesis enzymes. The most common haplotypes associated with regions under selection for the Great Smoky Mountains population were also common among other populations throughout the region, indicating a key role of putatively selective variants in the fitness of invasive populations. Interestingly, many of these haplotypes were absent among European wild boar reference genotypes, indicating feralization through genetic adaptation.

KEYWORDS
admixture, coat colouration, invasive bridgehead effect, invasive pigs, selective sweep

Timothy J. Smyser and Mirte Bosse contributed equally to this work.

1 | INTRODUCTION

Founder effects are expected to dominate genetic processes among non-native species introduced into new ecosystems, whereby released propagules capture a subset of the diversity found within native populations (Estoup et al., 2016). Upon introduction, individuals must survive, reproduce and potentially adapt to unique selective pressures encountered in novel habitats to give rise to invasive populations (Kolbe et al., 2004). Thus, the loss of genetic diversity incurred during the initial bottleneck would be expected to restrain both demographic and evolutionary processes and concomitantly diminish the likelihood of a successful invasion. Empirical studies investigating the genetic processes of biological invasion, however, have demonstrated that successful invasions are seldom characterized by the expected loss of genetic diversity (Estoup et al., 2016). Rather, various attributes of the introduction process, such as large propagule size, sustained gene flow from source populations, or admixture of multiple lineages in the introduced range, likely ameliorate the effects of an initial bottleneck (Comeault et al., 2020; Estoup et al., 2016).

Introduction from disparate source populations, in particular, may promote successful invasion given that admixture can efficiently offset expected losses of genetic diversity (Dlugosch et al., 2015; Dlugosch & Parker, 2008). One particular case of genetic release through admixture, called exoferalization, represents the hybridization of domestic and wild lineages among invasive populations and the assembly of gene combinations that have been shaped by natural and artificial selection (Gering et al., 2019). This particular genetic enrichment mechanism can lead to more rapid adaptation to the local environment by two means. First, the introgression of wild genetic material into feral animals from domestic origins can release descendant hybrid populations from the negative fitness effects conveyed by the domestication syndrome (Wright et al., 2020). The domestication syndrome is a phenomenon in which specific physical and physiological traits have repeatedly been modified during the domestication of several species. For example, in pigs, such traits include (but are not limited to) decreased overall intelligence, loss of coat pigmentation, decreased function of the olfactory system and deformation of the skull (Fulgione et al., 2017; Maselli et al., 2014). Generally speaking, these traits are thought to have a negative fitness effect on free-living animals and are selected against in the natural environment. Second, certain traits that arose through intensive breeding can be beneficial, such as increased litter size or larger body size (in the presence of sufficient food resources) (Fulgione et al., 2016). In that sense, artificial selection would increase the phenotypic range of a free-living population, on which natural selection could subsequently operate.

Through exoferalization, the resulting admixed populations may serve as a precursor to what has been characterized as an invasive bridgehead effect. This effect describes a process in which a primary invasion gives rise to adept invaders that pose a heightened risk for secondary invasions across a novel landscape, often hypothesized to be mediated through evolved invasiveness (Lombaert et al., 2010). This phenomenon is often explained through rapid local adaptation and increased genetic health. Conversely, secondary invasion success could also be attributed to increased abundance within invasive populations or similar patterns of human movement that contributed to the initial introduction (Bertelsmeier & Keller, 2018). Regardless, admixture from multiple source populations has been hypothesized to, at least, release introduced populations from inbreeding depression while potentially enabling unique gene assemblies and novel epistatic interactions that could increase fitness and, by extension, invasiveness (Kolbe et al., 2004; Lavergne & Molofsky, 2007). The increasing availability of high-resolution genomic tools enables the testing of hypotheses related to the response of enriched diversity to selective pressures while beginning to elucidate evolutionary mechanisms that contribute to heighted invasiveness (North et al., 2021).

Wild pigs (*Sus scrofa*) are recognized as among the most destructive invasive species in the world, with populations established on all continents except Antarctica (Lewis et al., 2017; Lowe et al., 2000). Although the origins of introduction may vary among the global regions invaded by wild pigs, both domestic pigs and wild boar have contributed to invasive wild pig populations distributed throughout much of the contiguous United States (Mayer & Brisbin, 1991; Smyser et al., 2020, 2024). Though most of the spread of wild pigs in the US occurred over the last 40 years, free-living populations of domestic pigs were initially established in the contiguous US in the 1500s, introduced with Spanish exploration (Mayer & Brisbin, 1991). Once established, populations were continuously augmented through the mid-1900s as a consequence of the incidental escape of pigs seasonally released into forested habitats to fatten on fallen mast crops (White, 2011). In the late 1800s and early 1900s, a time in which feral populations of domestic pigs were well-established, wild boar were first imported to the US from native populations in Europe to stock captive hunting preserves. Subsequent escapes from such preserves created opportunities for wild boar and feral domestic pigs to interbreed. The 1912 importation of wild boar to Hooper Bald, Graham County, North Carolina, adjacent to the present-day boundaries of Great Smoky Mountains National Park, was perhaps the most consequential of these introductions (Bratton & Power, 1975; Buderman et al., 2023; Mayer & Brisbin, 1991). When the Hooper Bald hunting preserve failed as a commercial enterprise in 1922, wild boar escaped and began to interbreed with the established feral pig populations, with hybrid populations establishing within the park by the late-1940s (Bratton & Power, 1975; Mayer & Brisbin, 1991; Stegeman, 1938). Shortly thereafter, wild pigs from the region were noted for wild boar phenotypic characteristics, a morphotype deemed to be more desirable for hunting than that of typical feral pigs (Mayer & Brisbin, 1991). Accordingly, from the late 1920s through the 1970s, wild pigs from the region were used as a source for the deliberate creation of new wild pig populations (to create additional hunting opportunities for the public) or released to augment established feral pig populations (as a means of increasing the phenotypic appeal of recipient populations) (Mayer & Brisbin, 1991).

Ancestry analyses of wild pigs throughout invaded regions within the contiguous US demonstrate that contemporary populations overwhelmingly represent hybrids of domestic pigs and wild boar (Smyser et al., 2020, 2024). The exoferal origin of invasive wild pigs diverges from the historical record that documents a long and sustained period

in which escaped domestic pigs contributed to wild populations in contrast to very limited wild boar releases. The discrepancy between the historical record and genetic patterns suggests greater fitness and associated heightened invasiveness of wild boar × domestic pig hybrids, exemplified by the Great Smoky Mountain population. Accordingly, we sought to leverage the unique molecular resources available for the *S. scrofa* domestic-wild species complex to elucidate genomic processes shaping invasive wild pig populations, specifically, and how the interaction between genetic diversity and selective pressures contributes to invasiveness more broadly. In addition, we aimed to unravel whether the hybrid genomic background of US wild pigs facilitated feralization and local adaptation.

The goal of this study was to characterize the genomic response of an exoferal population of invasive wild pigs to natural and anthropogenic selective pressures and evaluate whether hybrid origins contribute to heightened invasiveness related to genetic enrichment. Utilizing high-resolution Single-Nucleotide Polymorphism (SNP) genotypes generated for wild pigs collected in Great Smoky Mountains National Park and surrounding populations, our objectives are to identify loci and associated genes that demonstrate a response to selection, evaluate the association of identified genomic regions to domestic pig versus wild boar origins, and determine whether signatures of selection observed among wild pigs conform to the predictions of the invasive bridgehead hypothesis. Great Smoky Mountains National Park represents an ideal study system to address these objectives as: (1) the associated wild pig population descends from among the earliest releases of wild boar in the US, (2) the long-term management as a national park (established in 1934) likely limits human-facilitated immigration that could disrupt evolutionary processes and (3) the park encompasses >2000 km^2 of remote, forested habitat, in which the population is subjected to a wide array of natural selective pressures.

2 | MATERIALS AND METHODS

2.1 | Wild pig sampling, and genetic and ecological context

Genetic material was collected from wild pigs from the Great Smoky Mountain National Park population (GRSM, $n=537$) and spatially disjunct but proximate populations in Tennessee, South Carolina, North Carolina and West Virginia (TN, SC, NC, WV; $n=344$) (Figure 1, Table S1). These populations represent hybrid swarms from the original Hooper bald introduction and possess modestly greater ancestry from European wild boar (mean=0.57) than Western heritage breeds (mean=0.29; breeds developed in Europe or in North America from European stock) (Dataset S7 (Smyser et al., 2020)). Genetic samples (hair or a pinna biopsy) were collected by personnel from the National Park Service, Wildlife Services, or other cooperators from animals euthanized as part of invasive species control or disease surveillance efforts. For each animal, the location of sampling was recorded.

In previous work, we elucidated the ancestral sources that contributed to wild pig populations throughout the contiguous US (Smyser et al., 2020). Therefore, for comparison purposes, pre-existing genotypic data from European wild boar and Western heritage breeds were used as proxies for the ancestral populations of invasive swine in the US (Alexandri et al., 2017; Burgos-Paz et al., 2013; Goedbloed et al., 2013; Iacolina et al., 2016; Roberts & Lamberson, 2015; Smyser et al., 2020; Yang et al., 2017). The Western Heritage reference cluster includes breeds predominantly raised using traditional (i.e., extensive as opposed to intensive) husbandry practices and traditional breeding methods as opposed to intensive, genomics-based mate selection as it is typically implemented with modern commercial breeds. The wild boar reference cluster includes animals sampled across most of Europe,

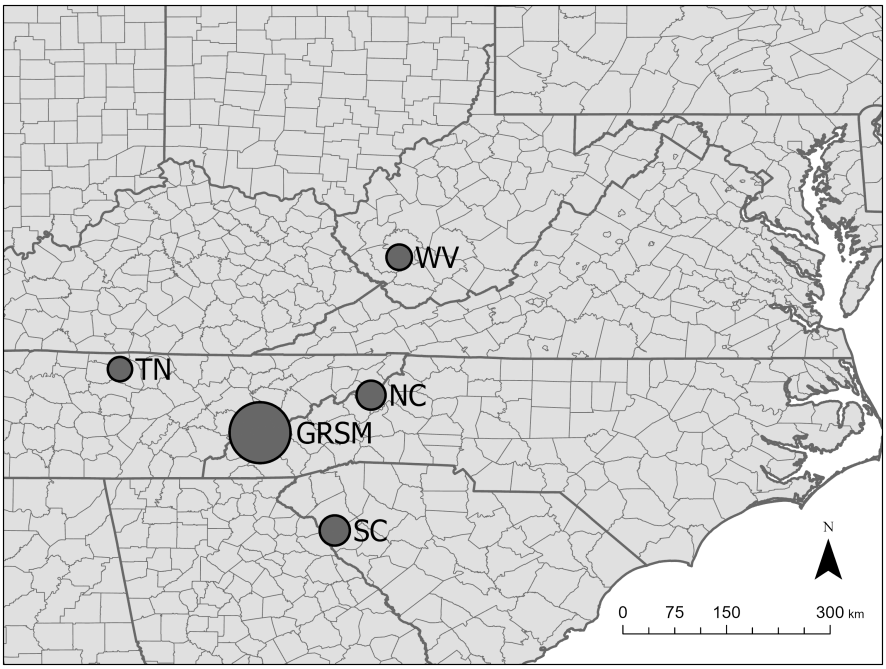

FIGURE 1 Sampling of Invasive Swine populations. This figure represents the sampling locations within the Great Smoky Mountain National Park (GRSM) and the outlying states. The circles within each state represent the general sampling areas, with the size of the circle representing the relative sample size. NC, North Carolina; SC, South Carolina; TN, Tennessee; WV, West Virginia.

spanning from Spain to Eastern Europe. These reference clusters were used to compare the relative genetic diversity and level of inbreeding of GRSM. Additionally, we aimed to assess the ancestral origin of the selective signals found by comparing GRSM with these reference clusters. To accurately interpret the origin of selective signals derived from either the Western Heritage or European wild boar reference clusters, an extra reference cluster not associated with the wild pigs was needed. To this end, we used a collection of Asian *S. scrofa* genotypes that included both local domestic breeds and endemic wild boar (Burgos-Paz et al., 2013; Smyser et al., 2020; Yang et al., 2017).

Adult GRSM wild pigs generally show a black coat (regarded as a domestic pig trait), whereas juveniles display a wild-type striped pattern (characteristic of wild boar; Mayer & Brisbin, 1991). The erect ears and straight tails of GRSM wild pigs are consistent with wild boar phenotypic characteristics, while the reproductive output of wild pigs in general tends to be more similar to domestic output (Chinn et al., 2021, 2022). Based on anecdotal evidence from park managers, black bears and bobcats exert some predation pressure; however, human predation (culling) is the predominant pressure. As for the habitat, Great Smoky Mountain National Park is characterized by closed canopy forests such as hardwood forests (dominant overstory species: *Quercus rubra*) and birchwood forests (most dominant overstory species: Quercus rubra; Jenkins, 2007). In that sense, the primary diet of GRSM wild pigs consists of plant material (about 90%), with mast years largely influencing winter survival and reproductive success (Salinas et al., 2015; Scott, 1973).

2.2 | Genotyping and filtering

All tissue samples analysed for this study were genotyped using the GeneSeek Genomic Profiler (GGP) for Porcine SNP Array (Illumina BeadChip microarrays [San Diego, California] licensed exclusive to GeneSeek, a Neogen Corporation [Lincoln, Nebraska]; Ramos et al., 2009), yielding 62,128 biallelic loci distributed across all 18 autosomes. Genotypes were filtered using PLINK v1.90b6.21 64-bit (Chang et al., 2015), with filtering steps adapted to the specific requirements of the associated analyses. As a general rule, all filtering was performed after subsetting chromosomes or merging datasets. For all analyses, SNP loci with call rates lower than 95% were removed, and individuals missing >10% of SNP loci were discarded. Generally, these quality control steps resulted in a dataset with 530 genotypes from GRSM. For analyses based on genetic stratification, such as principal components analysis, admixture analysis and genetic diversity, genotypes were also filtered by removing all SNP loci with a minor allele frequency lower than 5%.

To complement analyses with the GGP SNP array, 53 animals, strategically selected to represent the genetic range observed within GRSM, were genotyped with the high-resolution Axiom Porcine Genotyping Array (ThermoFisher, Applied Biosystems, Waltham, MA; yielding 592,053 SNP loci for analysis). The selection of animals was based on family clusters, with at least one animal selected from each observed subcluster.

2.3 | Genetic diversity, inbreeding and population structure

Genetic relatedness to other populations and genetic structure within GRSM were assessed with principal components analysis (PCA), admixture analyses and *F*-statistics. For PCA, we used the --pca function of PLINK with default settings. This calculates the first 20 principal components by sliding along windows of SNP loci (Figure 2, Figures S3 and S4). For genetic clustering analyses, we utilized ADMIXTURE v1.3 to delineate genetic clusters (Alexander et al., 2009). Generally, we chose the number of ancestry clusters based on the lowest value for the cross-validation statistic calculated by ADMIXTURE, combined with our knowledge of the population history. In other words, we compared the CV scores of ADMIXTURE runs over a range of clusters (*k*) from 1 to 8 and selected the iteration with the lowest CV score as the most informative value of *k*. In cases where CV scores were very similar across values of *k*, we chose the ADMIXTURE run with ancestral clusters that were closest to what we expected based on the historical records of the populations (Figures S1 and S2). Finally, overall population genetic differences between GRSM and the surrounding populations were characterized with Wright's F_{ST} analyses (Holsinger & Weir, 2009; Weir & Cockerham, 1984) using the --fst function of PLINK with default settings.

2.4 | Genetic diversity and inbreeding

To assess genetic diversity, heterozygosity and level of inbreeding were calculated. Measures of genetic diversity also served as a validation for selective sweep analyses, as selective sweep analyses are sensitive to the (near) fixation of large regions through drift. To characterize the genetic diversity of invasive wild pigs, we calculated the total heterozygosity of the GRSM population and compared these values to both the reference clusters and the surrounding populations. This was done using the PLINK function --het, which calculates individual heterozygosity as 1 minus the number of homozygous SNPs divided by the total number of SNPs genotyped for that individual.

To evaluate the level of inbreeding, the genotypes were assessed for the presence of long runs of homozygosity (ROH). For all individuals, the number of ROHs, the average length of ROHs and the total length of ROHs were calculated using the function --homozyg of PLINK. We restricted the characterization of ROHs to regions with a minimum SNP density of 80 SNPs/kb, a maximum gap size of 600 kb and required complete homozygosity (excluding regions with >1 heterozygous loci) as described by Meyermans et al. (2020). Additionally, datasets used for ROH calculation were only filtered for SNP call rates (≥95%) and individual call rates (≥90%), and thus not for minor allele frequency or linkage disequilibrium. The number of ROH segments and the length of these fragments were used to calculate the fraction of ROHs (fROH) of the total autosomal genetic material per individual. As inbreeding leads to increased homozygosity, the fROH is a relative measurement of inbreeding. As a relative measurement, fROH values for GRSM were compared with the reference clusters and the surrounding invasive populations. To

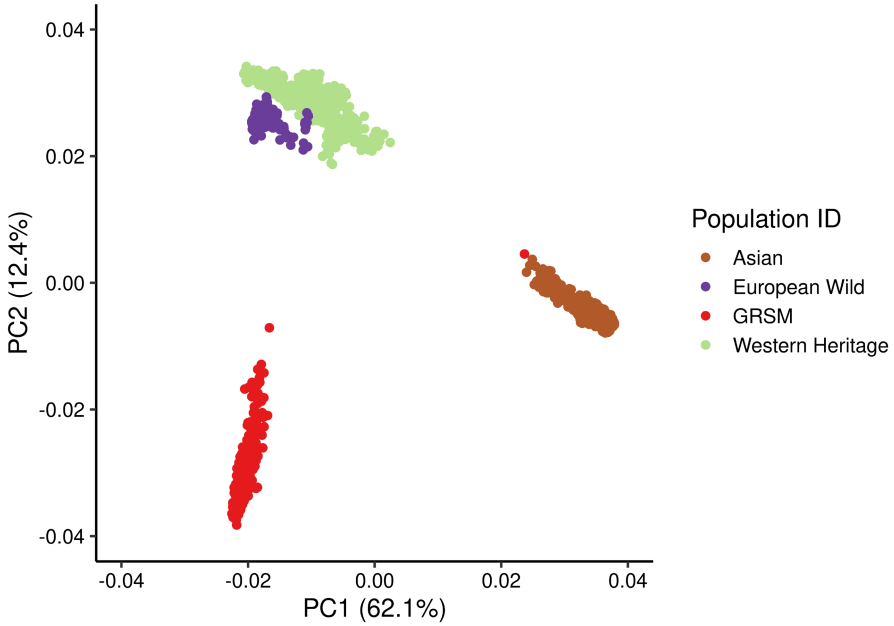

FIGURE 2 Principal Component Analysis comparing the GRSM population with the reference clusters. Represented is the genetic clustering of the subpopulations within the GRSM. The algorithm calculated the top 20 PC axes, of which the top two are represented here. Note that the GRSM represents its own genetic cluster, while still showing the signature of a hybrid swarm.

further infer family structure and detect inbreeding, we also performed Identity-By-Descent analyses using the --genome function of PLINK on specific populations. These analyses were used to calculate the relative probability that an individual was more related to another individual than the baseline would predict. Detecting inbreeding is relevant as drift effects tend to disrupt selective sweep analyses; therefore, discarding highly related individuals is beneficial. Although some pairwise dyads in the dataset were distantly related, no pair of individuals exceeded the relatedness threshold (Proportion Identity-By-Descent = 0.75) specified for selective sweep analysis (Figure S6). To support this analysis, we also performed a relatedness analysis using the --relatedness2 function from VCFTOOLS v0.1.16 (Danecek et al., 2011). The output was checked for an average relatedness score as well as individual pairs showing a relatedness score equal to or higher than 0.3. The average relatedness score was −0.054, while 26 pairwise comparisons out of 281,430 possible comparisons showed a relatedness score of at least 0.3. Relatedness analyses were predominantly performed to assess the general genetic structure of the sampled groups. In that sense, we allowed some family structure as this would not bias our analyses, and thus we were able to use a relatively high Identity-By-Descent threshold.

2.5 | Selective sweep analyses based on extended haplotype homozygosity

Given that recent (strong) selection leads to an increase in the frequency of beneficial alleles and eventual fixation, SNPs that are physically linked to such selectively advantageous alleles are often also fixed. As a result, long haplotypes will form in the genomic regions surrounding beneficial alleles and will contain many neutral variable loci and one or a few selective loci. Using this principle, we screened the GRSM genotypes for long and frequent haplotypes, indicative of selective sweeps. To generate haplotypes, the data was phased with SHAPEIT v2.17 (Delaneau et al., 2012), using default settings. After phasing, the data were filtered for MAF (≥5%), as low-frequency alleles can lead to false positives in a sweep analysis. The haplotypes were then scanned for the presence of extended homozygosity. Using the package *rehh* in R (Gautier & Vitalis, 2012), the Extended Haplotype Homozygosity (EHH) was calculated, which is defined as the probability that two chromosomes carrying the core haplotype are identical by descent over a certain region (Tang et al., 2007). The EHH values were calculated using the approach of Sabeti et al. (2007). Using these EHHs, the integral of decay per SNP was calculated, which itself was used to calculate the integrated haplotype score (iHS) for individual SNP loci following Voight et al. (2006). SNP loci with an iHS corresponding to a p-value ≤ .01 were interpreted as being putatively under selection and identified as focal SNP loci. This selective sweep protocol was also applied to genotypes produced with the high-resolution Axiom Porcine Genotyping Array data to investigate patterns that arose from the main selective sweep analysis on a finer scale.

To compare haplotype structure between GRSM and the reference clusters (European wild boar and Western heritage breeds of domestic pig), we also performed a cross-population EHH analysis (XP-EHH). This analysis compares EHH values for specific alleles between two populations. Using this principle, haplotypes that are more elongated in one population than in the other population can be detected, indicating local selection in the first population. Note that this analysis was only used as a supporting analysis for the iHS-based sweep analysis, so the main selective sweep analysis was only performed on GRSM to reduce noise generated by different genetic backgrounds.

To further confirm the iHS SNP loci as putatively being under selection, the occurrence of ROHs containing focal iHS SNP loci was compared to the occurrence of ROHs containing neutral SNP loci. Similar to the presence of extended haplotypes due to recent selection, recent selection can also lead to the formation of long ROHs through the effects of the linkage between neutral variants and selected variants. Therefore, a ROH can be indicative of recent selection in a particular

region (Bosse et al., 2012). The calculation of ROHs for this purpose was done in the same manner as described above for the fROH.

For the validated iHS focal SNP loci, phased states of surrounding SNP loci were used to determine the genomic regions under selection. Around each focal marker from the iHS analysis, the site-specific EHH (EHHS) was calculated. Regions under selection were determined as the regions in which an EHHS threshold of 0.3 was exceeded (Figure S9). If two regions physically overlapped for more than 40 percent (in bp), the focal marker with the lowest iHS was discarded. The resulting subset of SNP loci was used in all further analyses.

2.6 | Spread of haplotypes under selection through invasive populations

Using phased haplotype states (as generated by SHAPEIT), we investigated the frequencies at which the variants that were under selection in GRSM occurred in the surrounding populations. Specifically, we compared the haplotype frequencies containing SNP loci under putative selection in GRSM with comparable haplotype frequencies among other invasive populations (Tennessee, South Carolina, North Carolina and West Virginia). Using the physical positions of the iHS focal SNP loci, haplotypes were determined based on the combined phased data of GRSM and the four surrounding populations. All haplotypes were approximately 0.5 Mbp, with the focal marker at the centre (Table S4 and Figure S13).

2.7 | Inferring genes under putative selection and related GO terms

Additionally, we sought to identify the genes that were located in the regions under putative selection as determined by the approaches described above. For this, gene transfer data from the Ensembl database was used, based on the Sscrofa11.1 genome assembly (GCA_000003025.6). From this dataset, all genes (including both introns and exons) overlapping the regions under selection were extracted. As the regions under selection were relatively long, they contained a large number of different genes. Therefore, we used Gene Ontology (GO) terms to assess which genes were most likely causing the selective signal. Additionally, genes found by previous research to be related to domestication were identified and evaluated for signatures of selection according to our analysis (Maga et al., 2015). For the GO term analysis, the gene stable IDs of genes under putative selection were submitted to Ensembl BioMart to return gene names. Using these gene names, the Functional Annotation Tool of DAVID v. 6.8 (Huang et al., 2009) was used with default settings to extract Gene Ontology (GO) terms for both KEGG pathways and biological processes (filtered for broad GO terms) based on human gene annotations. As GO terms tend to be fairly wide-ranging, two sets of genes under putative selection were more closely examined. First, we assessed genes associated with melanogenesis given that the coat colouration of wild pigs appears to be different than that of domestic pigs, indicating a basis for morphological adaptation (Chinn et al., 2021; Mayer & Brisbin, 1991). The importance of melanogenesis as a metabolic pathway under apparent selection was also revealed by the gene ontology analysis. Second, genes under selection were cross-referenced with a list of 92 craniofacial genes present in *Sus scrofa* (a subset of the list of Maga et al., 2015). These skull morphology genes are of specific interest since the occipital wall of the skull in domestic pigs is different from that of wild boar (Dinu, 2009).

2.8 | Origin of haplotypes under selection based on identity-by-descent

As wild pigs most likely descend from both European wild boar and Western heritage breeds (Smyser et al., 2020, 2024; Stegeman, 1938), we were interested in identifying the ancestral source of the SNP loci under selection. To this end, we identified which haplotypes containing focal SNP loci were shared between GRSM and the reference clusters (including the Asian cluster as a reference). As recombination events are likely to have occurred both within GRSM and within the reference clusters, it would be improbable to find identical long haplotypes between populations. It is more likely to find shorter/fragmented haplotypes, which could miss the SNP loci in our genotype dataset. To allow for small genetic differences while comparing haplotype structure, an approach based on Identity-By-Descent was used, which allowed us to determine shared haplotypes based on the estimated ancestral state. The combined genotype data of GRSM and reference clusters was phased using the Identity-By-Descent dependent method of BEAGLE v5.2 with default settings (Browning et al., 2018). The phased haplotypes were assessed for the probability that they were based on the same ancestral haplotype by means of the software REFINED IBD (Browning & Browning, 2013). Shared haplotypes were determined within and between the previously determined family clusters. Shared segments were recorded if the segments had a LOD score of at least three, after trimming 0.001 cM off the ends of the haplotypes. The location and length of haplotypes that GRSM shared with the reference clusters were determined (Figure S11). The relative proportion of shared haplotypes per reference cluster was calculated by dividing the number of shared haplotypes with the potential maximum shared haplotypes based on sample size and all potential positions for shared haplotypes.

3 | RESULTS

3.1 | Genetic diversity and inbreeding of wild pigs

In total, we genotyped 881 animals with the GGP assay (62,128 biallelic loci), 537 from the GRSM and the remaining 344 from North Carolina (NC), South Carolina (SC), Tennessee (TN) and West Virginia (WV) (hereafter, outlying populations).

The invasive wild pig population within the Great Smoky Mountains National Park demonstrated similar levels of genetic diversity relative

to reference populations of Western heritage breeds of domestic pig or European wild boar (Figure 3a,b). Observed heterozygosity among GRSM (Het = 0.22, n_{loci} = 28,368) was similar to that observed for Western Heritage breeds (0.26) and the European wild boar (0.22). Likewise, the inbreeding coefficient derived from homozygous segments (fROH) for GRSM of 0.098 (n_{loci} = 28,368) was comparable to Western Heritage (0.10) and slightly higher than European wild boar (0.051).

Levels of genetic diversity and inbreeding were similar across all invasive populations (Figure 3c,d). Average heterozygosity for GRSM was 0.23 (n_{loci} = 57,654), whereas heterozygosity among surrounding invasive populations ranged from 0.17 to 0.29 (WV: 0.17, NC: 0.24, SC: 0.26, TN: 0.29). A similar trend was observed for fROH, with an average fROH value within GRSM of 0.3 (n_{loci} = 57,654), compared to values between 0.16 and 0.44 for the other populations. The PCA, ADMIXTURE and F_{ST} analyses all demonstrated that the GRSM population had a distinct genetic signature compared to the outlying populations (Figures S3 and S6).

3.2 | Integrated haplotype scores and runs of homozygosity indicate selective sweeps

In calculating the integrated haplotype score (iHS) for all available SNP loci (n_{loci} = 57,472), 233 loci were found to be under putative selection (Figure 4). After correcting negative values, the average iHS of putatively selected loci was 2.86 (SD 0.23). These loci were validated as being under recent selection by evaluating their occurrence in ROHs (Figures S7 and S8): SNPs that were under putative selection, as inferred from the iHS analysis, were more often contained in ROHs (on average in 266.5 individuals (SD 131.0) of the total 530 individuals), compared to SNPs not under putative selection (present in a ROH in 154.0 (SD 69.9) individuals). To validate these results, a bootstrapping analysis of 100 repeats sampling 233 random markers was performed. This analysis resulted in an average of 154.7 individuals (SD 4.79) possessing a ROH per assessed marker.

The putatively selected focal SNP loci were found to be closely linked and non-randomly distributed across the genome, allowing us to identify specific chromosomal regions under selection (e.g., red segments in Figure 4b). This was done by calculating the site-specific Extended Haplotype Homozygosity (EHHS) for all loci surrounding the focal SNP and then determining the upper and lower boundaries for the regions under selection based on an EHHS threshold of 0.3. After discarding duplicate regions under selection, this analysis resulted in 79 unique regions suspected to be under selection, which were used in the following analyses (Dataset S1).

When performing the iHS and EHHS analyses on the high-resolution genotypic data (n_{loci} = 373,532), we found 430 SNP loci under selection distributed over 126 individual regions (Dataset S2).

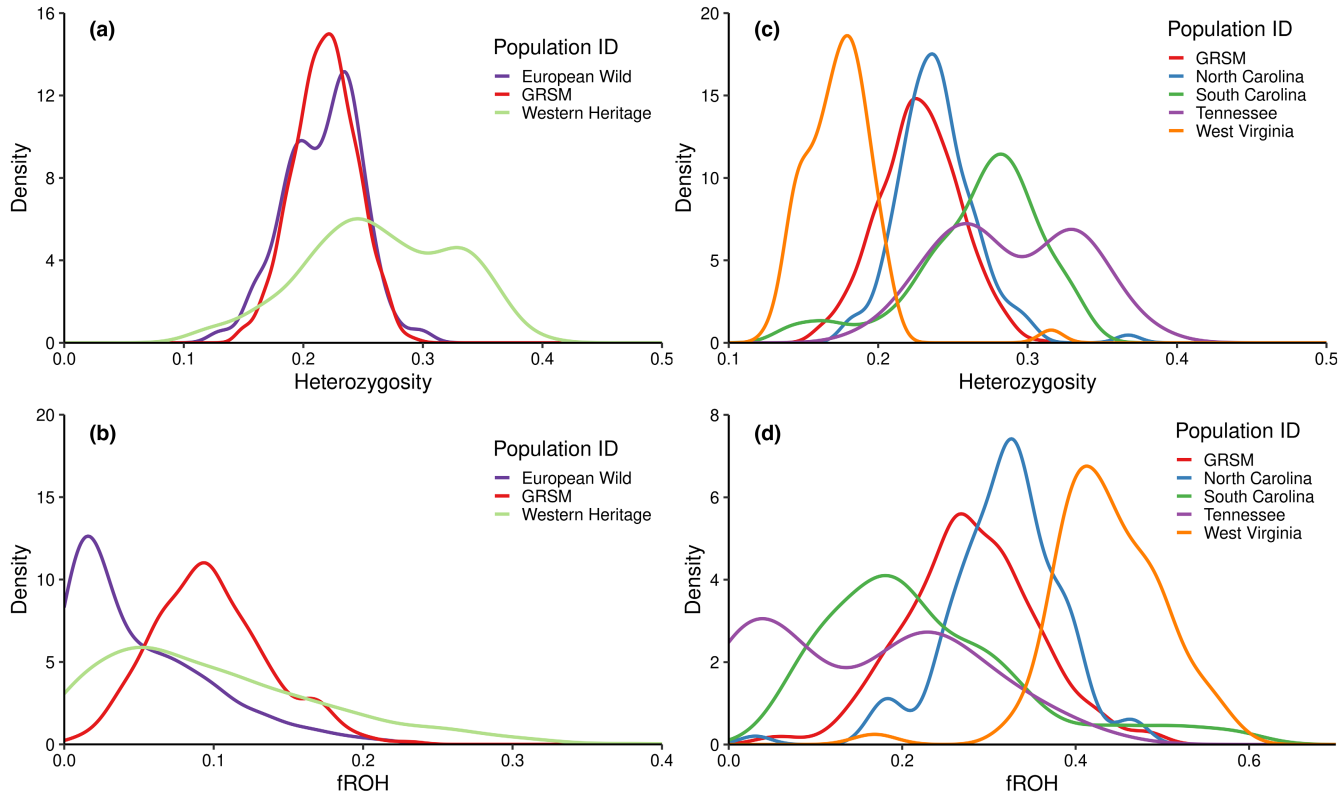

FIGURE 3 Genetic diversity and level of inbreeding of invasive and native pigs. In this figure, the genetic diversity in terms of heterozygosity (a, c) and the level of inbreeding in terms of fraction Runs of Homozygosity (fROH) (b, d) of Western heritage, European wild and invasive populations are represented. The measurements within a sub-figure are relative to each other; therefore, the heterozygosity of GRSM is 0.22 in (a) and 0.28 in (b).

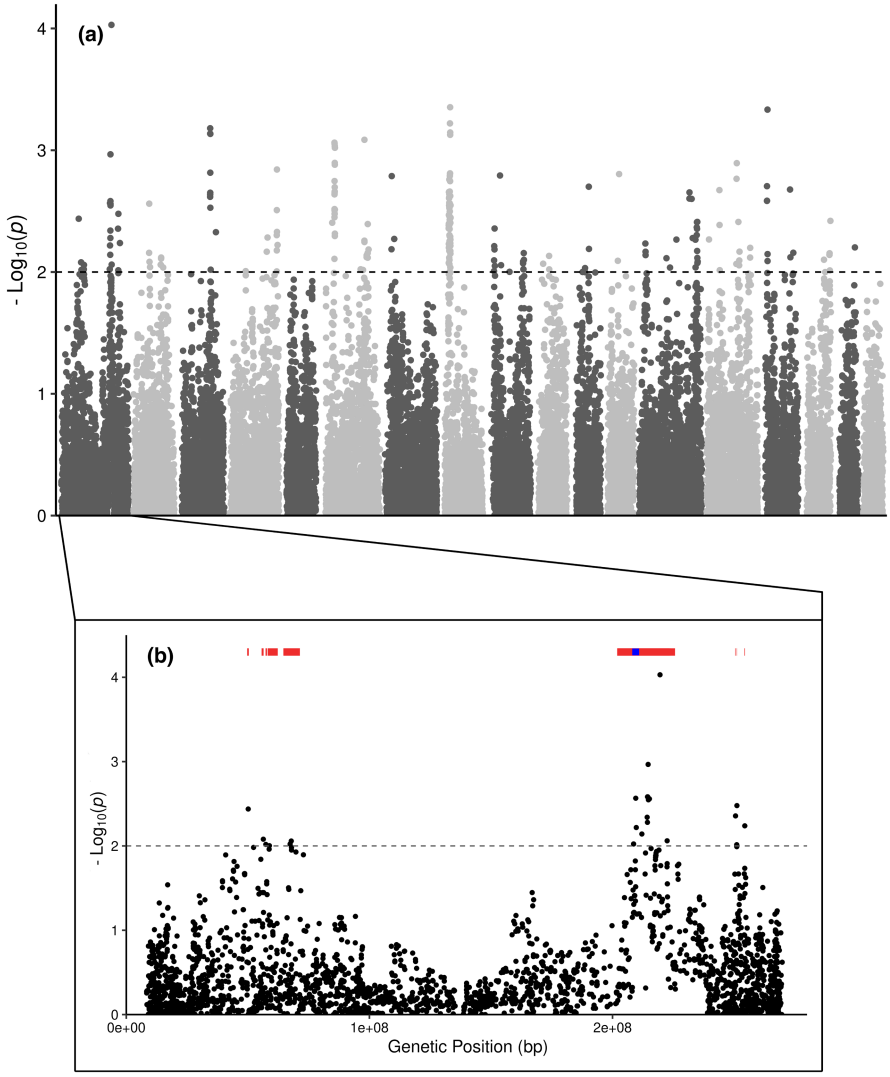

FIGURE 4 Integrated Haplotype Score for extended haplotypes. For each individual genetic marker on each chromosome, the Extended Haplotype Homozygosity (EHH) was calculated. Using the integrals of these values, the integrated Haplotype Score (iHS) could be calculated. This is a measurement for the relative chance that a marker is found in a long haplotype. All markers exceeding a threshold of log(*p*-value) 2 were considered significant (grey segmented line). (a) All chromosomes, while (b) zooms-in on chromosome 1. In (b), the red segments depict the regions under selection derived from these markers. The blue marker represents the location of a genetic marker associated with the melanization gene *TYRP1*.

Because of the higher resolution of this data, these regions were smaller than those inferred from the GGP data, therefore allowing a closer examination of selective signals. In total, 13 out of the 79 regions identified with the GGP data overlapped (in part) with at least one of the 126 regions from the high-resolution genotypic data, while many other regions were nearly overlapping. One important observation was that the major peak on chromosome 1, highlighted in Figure 2b, was narrower using the high-resolution genotypic data while still incorporating the *TYRP1* gene (Figure S5). Similarly, the supporting XP-EHH analyses also revealed the *TYRP1* gene to be in extended haplotypes in the GRSM compared to both the European wild boar and the Western heritage clusters (Figure S10).

3.3 | Shared haplotypes show the European origin of markers under selection

Most of the putatively selected haplotypes identified within the GRSM population for which ancestry could be determined were associated with Western heritage origins, followed by European wild boar (Figure 5 and Figure S12). Our analysis revealed that at 12,250 genetic regions, GRSM shared 153,162 haplotypes with the other populations (Western Heritage: 116,283; European wild boar: 25,611; Asian: 16,336). Thus, of the haplotypes for which the ancestry could be determined, 66% were of Western heritage ancestry, versus 25% and 9% for European wild boar and Asian swine, respectively (percentages are corrected for sample size). Additionally, the haplotypes that GRSM shared with Western heritage breeds were also longer (4.61 Mb, SD 4.28) than the haplotypes shared with European wild boar (3.89 Mb, SD 2.61) or the Asian reference cluster (3.85 Mb, SD 3.12). Combined, these results point to a European origin (both wild and domestic), with the important note that a unique genetic signature is found in GRSM, as most of its haplotypes were unique to this population (710,253 unique haplotypes).

A subset of the data only consisting of haplotypes that contained at least one of the 233 iHS focal SNP loci showed largely similar origins. GRSM shared 7403 haplotypes with Western heritage, 1363 haplotypes with European wild boars and 1068 haplotypes with Asian animals. With relative proportions of 74%, 19% and 7%, respectively, associations of variants under selection also demonstrate a European origin.

A subset of the data only consisting of haplotypes that contained at least one of the 233 iHS focal SNP loci showed largely similar origins.

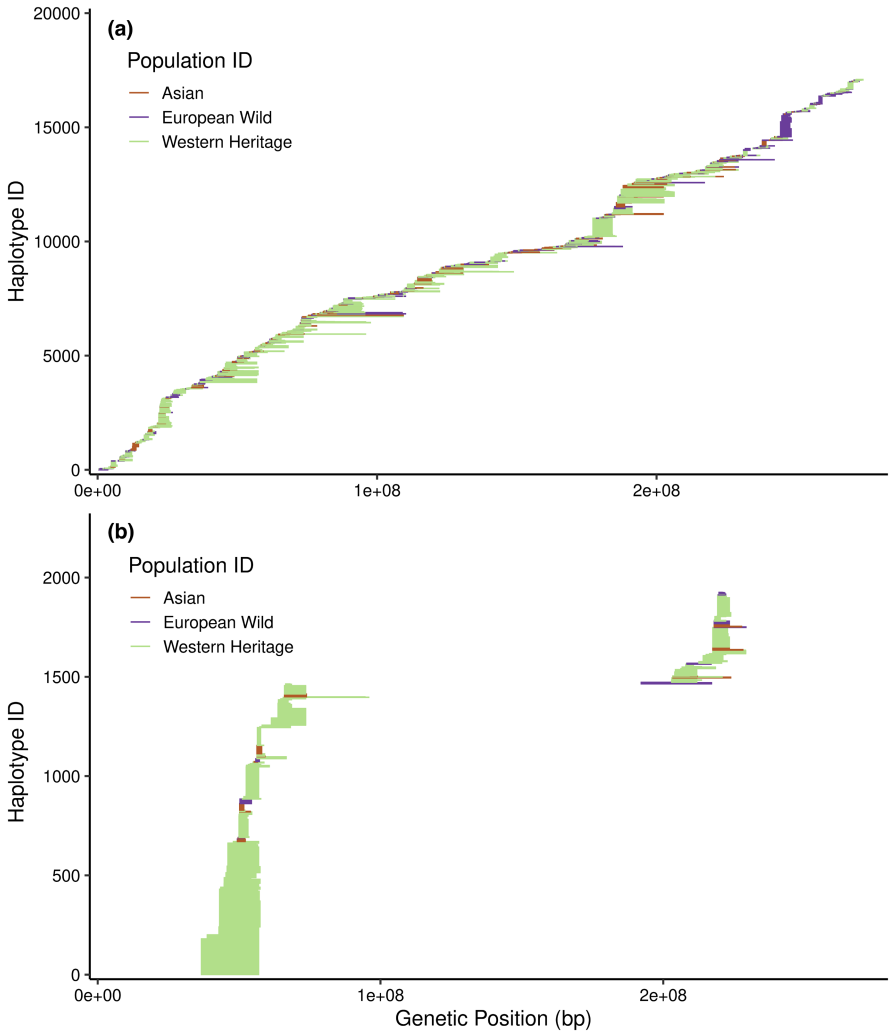

FIGURE 5 Haplotypes on Chromosome 1 that the GRSM shares with Western Heritage, European Wild and Asian clusters. Represented here are all haplotypes (a) and only the haplotypes containing iHS markers (b) from chromosome 1 that GRSM shares with any of the reference clusters. The differentiation of haplotypes is based on an identity-by-descent approach that calculates the likelihood that two haplotypes are derived from the same ancestral haplotype. Only haplotypes with a LOD score of 3 or higher were considered identical.

GRSM shared 7403 haplotypes with Western heritage, 1363 haplotypes with European wild boar and 1068 haplotypes with Asian animals. With relative proportions of 74%, 19% and 7%, respectively, associations of variants under selection also demonstrate a European origin.

3.4 | iHS focal marker haplotype frequencies show haplotype spread through invasive populations

Most haplotypes containing iHS focal SNP loci that had high frequencies in GRSM were also common among the surrounding invasive populations, even though recombination events and multiple ancestries could easily affect haplotype structure. As an example, Table S4 represents the haplotype frequencies consisting of SNP loci surrounding the melanistic gene *TYRP1* (Ren et al., 2011; Wu et al., 2016). The most frequent haplotype for GRSM (0.91 of all chromosomes) is also frequently present in the outlying populations (TN: 0.60, SC: 0.47, NC: 0.60, WV: 0.30). When considering all 79 regions under selection, we observed that an average of 76% of all chromosomes present in GRSM carried the same haplotype. The most frequent haplotypes of GRSM were found to have average frequencies ranging between 0.16 and 0.43 in the outlying populations (TN: 0.16, SC: 0.20, NC: 0.43, WV: 0.23) (Figure S13).

3.5 | Gene ontology enrichment analysis

The 79 regions that we found to be under putative selection based on the iHS analysis contained a total of 1235 genes (Dataset S3), 645 of which we were able to identify the gene name. On this gene set, a gene ontology (GO) term enrichment analysis resulted in 155 significant GO terms (p-value < .05) (Dataset S5). Most of the enriched terms were fairly uninformative (e.g., roughly 25 are somewhat related to ion transport), but one pattern did emerge – 7 GO terms associated with neurogenesis were enriched (12 terms for $p < .10$). Most notably, GO:0022008: neurogenesis ($p = .002$, 60 genes) and GO:0007399: nervous system development ($p < .001$, 88 genes) appeared enriched.

Additionally, an analysis of enriched KEGG pathway genes suggested melanogenesis ($p = .13$, 6 genes) was more enriched than most other GO terms. This non-significant result stimulated a closer investigation of the causative genes under putative selection (Table 1). Most of these six genes were broad receptors or transcription factors; however, two specific genes were identified, *TYR* and *TYRP1*, that encode

TABLE 1 Description of genes under putative selection.

Gene name	Phenotypic trait	Gene description	Genotype array	Literature
TYRP1	Coat colouration	Encodes an enzyme that converts DHICA into Indolequinone. Associated with brown and blond colouration in Asian pigs	Porcine SNP Array and High-resolution	Ren et al. (2011), Wu et al. (2016)
TYR	Coat colouration	Encodes an enzyme that converts Tyrosine into Dopaquinone. Knockdown causes total albinism in pigs.	Porcine SNP Array	Wu et al. (2012), Zhou et al. (2015)
COL9A1	Craniofacial formation	Encodes for collagen expressed in the eye socket, among other tissues.	Porcine SNP Array	Van Camp et al. (2006)
FREM1	Craniofacial formation	Encodes for a basement membrane protein. Mutations are associated with Craniosynostosis.	Porcine SNP Array	Vissers et al. (2011)
SIX3	Craniofacial formation	Transcription factor that suppresses WNT3. Important in the development of the forebrain and eyes.	Porcine SNP Array	Lagutin et al. (2003)
PIBF1	Craniofacial formation	Transcription factor induced by progesterone. Usually associated with the immune system, but mutations are also associated with craniofacial disorders.	Porcine SNP Array	Maga et al. (2015)
SMARCE1	Craniofacial formation	Encoded protein is part of SWI/SNF, a chromatin remodelling complex.	Porcine SNP Array	Fowles et al. (2003)
TP63	Craniofacial formation	Encodes for a transcription factor needed for proper olfactory development.	Porcine SNP Array	Durante et al. (2020)
Olfaction genes	Olfaction	Clusters of olfactory receptors	High-resolution	NA

Note: Represented are some of the most interesting genes located in regions under selection in GRSM individuals. The column "Phenotypic trait" represents only one of the functions that these genes are involved in.

for enzymes crucial in melanogenesis and are known to cause albinism when impaired (Wu et al., 2012). For *TYRP1*, three SNP loci surrounding this gene were found at a high frequency in GRSM (freq: 0.88), but less frequent among European wild boar (freq: 0.29) and Western heritage (0.12) (Table S2). When investigating the 53 animals with complementary high-resolution genotypes, we observed a specific haplotype surrounding *TYRP1* that was shared among 80% of all sampled chromosome copies and contained five SNP loci spanning 11.95 kb (Table S3). These results combined indicate a unique pattern for GRSM at this particular locus, suggesting selection for melanogenesis.

As the skulls of invasive wild pigs have a unique morphotype (Mayer & Brisbin, 1991), the genes under putative selection were cross-referenced with a list of 92 known craniofacial genes. From this list, six genes appeared to be under selection (Table 1), which may suggest a selective signal on skull formation.

A GO analysis on the 126 regions (899 total genes, Dataset S4) under selection according to the iHS analysis on the high-resolution genotype data revealed another interesting GO term signal; multiple enriched GO terms were associated with olfaction (Dataset S6). Among these enriched GO terms were 'GO:0050907: detection of chemical stimulus involved in sensory perception' ($p < .001$, 32 genes) and 'GO:0007608: sensory perception of smell' ($p < .001$, 33 genes). Closer examination revealed that two selective signals in olfactory receptor clusters on two different chromosomes are mainly causative of these signals. Combined with the neurological signal, one could speculate that this indicates selective pressure on traits affecting foraging behaviour and predator detection, as predation pressure (including population control) and new food types are strong selective pressures.

4 | DISCUSSION

Understanding the ecological, demographic and evolutionary processes that contribute to the establishment and spread of invasive species is a major challenge for invasion biology. Studying the importance of genetic diversity in limiting or enabling the capacity for a species to adapt to local environments has become feasible with the recent rise of genomics (Welles & Dlugosch, 2018). Using these new possibilities, our research demonstrates the effect hybridization, and specifically exoferalization, can have on increasing genetic diversity, adaptive potential and consequently the invasiveness of wild pigs.

4.1 | Admixed characteristics explained with genetic data

Based on the notion that the wild pigs in GRSM are hybrid descendants of wild boar and domestic pigs, we wanted to assess the implications of admixture on their adaptive potential and, by extension, invasiveness. Our results demonstrate that the invasive wild pigs of GRSM have undergone adaptation by means of directional selection, implying higher invasiveness. The selective

signals suggest increased fitness after selection at these localized genetic sites. This, combined with the fact that adaptation to novel environments (including human influences) increases the fitness of invasive populations, provides us with a genetic basis for increased invasiveness. Not only did the analysis indicate selection at genes related to foraging behaviour and cryptic colouration, but genetic stratification analyses provided a basis for the possibility of adaptation. Comparisons of heterozygosity and fROH demonstrated GRSM maintained genetic diversity similar to Western heritage breeds and European wild boar. The process of exoferalization, therefore, elevated genetic diversity beyond levels observed among populations established from a low number of founding individuals of either wild boar or Western heritage breed origins.

When assessing the speculative effects of the selective sweeps, the GO enrichment analyses suggested selective pressure on the neural system, revealing enrichment of genes involved in neurogenesis (GO:0022008, 60 genes) and nervous system development (GO:0007399, 88 genes). Interestingly, feralization has been suggested to involve evolution in brain size and composition, including behavioural adaptation (Henriksen et al., 2018). Notably, neurological capacity has been found to correlate positively with effective foraging behaviour and anti-predator behaviour (Croney et al., 2003), whereas the domestication process selected animals with behavioural traits that were more manageable in production settings (Kruska, 2005). Thus, artificial selection resulted in a severe decrease of the brain–body ratio in several domestic pig breeds compared to wild boar (Kruska, 1970; Maselli et al., 2014). Multiple olfaction GO terms were enriched, including 'Sensory perception of smell' (GO:0007608, 33 genes), hinting at the olfactory receptor family known to play an important role in rapid pig evolution (Liu et al., 2022; Maga et al., 2015). Our results are in line with previous research among Italian feral pigs that describes olfactory traits reverting to predomestication phenotypes (Petrelli et al., 2022). Selection on olfaction would stimulate foraging behaviour in a similar way as increased neurological development, and perhaps these selective signals are the result of the same environmental selective pressure. Additionally, consistent population control efforts within GRSM since 1959 are likely to have led to anti-predator behavioural responses associated with humans, potentially with adaptations in both morphology and physiology (Buderman et al., 2023; Mayer & Brisbin, 1991). The adaptive response to both environmental and anthropogenic selective pressures will lead to increasingly invasive animals due to the direct extension of fitness to the invasiveness of these introduced species.

Another interesting (though admittedly speculative) pattern is that six known craniofacial genes appear to be under putative selection. These genes encode for two structural proteins (Van Camp et al., 2006; Vissers et al., 2011) and four transcriptional regulators (Durante et al., 2020; Fowles et al., 2003; Lagutin et al., 2003; Maga et al., 2015). The observed selective signal is exciting, as the skull shape has been observed to be under directional selection. For example, across the *S. scrofa* domestic-wild species complex, a gradient of skull occipital angle morphologies has been described that ranges from ±65° in domestic pigs to ±95° in wild boar (Mayer & Brisbin, 1991). Interestingly, when comparing long term feral pig populations and short-term feral pig populations, the long-term feral pigs' occipital angle was found to shift towards a greater angle (and thus starting to resemble wild type). Dinu (2009) proposed that the difference in skull shape that has emerged among domestic pigs is due to (1) artificial selection for more manageable pigs and (2) directional selection to relieve stress on jaw muscles associated with domestic diets. The presence of selective signals on skull morphology genes in GRSM could therefore indicate a similar return to wild type skulls.

Two genes encoding for crucial enzymes in the melanogenesis pathway were found in selective sweeps: *TYR* (tyrosinase, Chr 9) and *TYRP1* (tyrosinase related protein 1, Chr 1). Both genes are associated with coat pigmentation, as *TYR* knockdown experiments have been shown to cause total albinism in pigs and mice (Wu et al., 2012; Zhou et al., 2015). In the melanosome, tyrosine is converted to dopaquinone by the activity of tyrosinase, which in turn is converted into either eumelanin, which is mediated by *TYRP1*, or pheomelanin (Ito & Wakamatsu, 2003). The ratio between these types of melanins determines the coat colouration. While mutations in *TYR* have not yet been associated with a specific coat colouration, a mutation in *TYRP1* is thought to cause the unique brown colouration of Chinese pigs (Ren et al., 2011). Similarly, examples are known of mutations in this gene being associated with different coat colouration morphologies in domestic breeds of chicken, goat and rabbit (Becker et al., 2015; Li et al., 2019; Utzeri et al., 2014). These studies corroborate our interpretation of selection associated with this gene as inferred from the unique allele frequencies for three SNP loci encompassing *TYRP1* in GRSM compared to the reference clusters. An interesting aspect is the fact that in multiple animal species, similar melanogenesis-relevant genes cause behavioural changes, specifically fearfulness and aggressiveness (McKinnon & Pierotti, 2010). This, combined with the neurological developmental changes, could have interesting implications for the behaviour of invasive wild pigs.

The putative selective pressure on colouration genes could be explained by the fact that coat colouration appears to influence neonate survival. Survival rates of wild pig neonates with striped coats (a characteristic of wild boar that is generally absent in domestic pigs) appear to be higher than those with domestic pelage patterns (e.g., fully black or black and white spotted) (Chinn et al., 2021). Furthermore, selection of coat colouration could provide increased protection against sun exposure. Wild boar are susceptible to UV damage, partly explaining their need for wallowing (Bracke, 2011; Newell et al., 2021). A wild type coat offers greater UV protection; however, one could question the selective pressure imposed by UV exposure in the closed canopy forests characteristic of the Great Smoky Mountains region.

The fact that genes responsible for neurological development, coat colouration and skull formation appear to be under selection in invasive wild pigs, is interesting from an evolutionary perspective in the context of the role of these traits in the domestication syndrome. The domestication syndrome describes a phenomenon in which certain trait morphs have been selected for in multiple domestic species and are generally considered to result in low fitness in the wild (Wilkins et al., 2014; Wright

et al., 2020). As domesticated traits often decrease effective foraging, a return to the ancestral state could imply increased invasiveness among GRSM. The signatures of selection indicate that genes specifically associated with the domestication syndrome are under selection, solidifying an evolutionary hypothesis in which processes of natural selection favour the ancestral, wild type of those traits. This pattern has also been observed in other domesticated populations that became feral, such as rabbits, chickens and dingos (Henriksen et al., 2018). Interestingly, our analyses contradict the assumption that these haplotypes are wild boar-derived, as the selected haplotypes were genetically most similar to haplotypes observed among heritage breeds. Therefore, it is not merely the introgression of wild boar alleles, but the genome-wide combination of domestic and wild ancestry that forms the basis of the invasiveness of the wild pigs, strikingly similar to patterns observed among Hawaiian feral chickens (Johnsson et al., 2016).

4.2 | From exoferalization to invasive bridgehead

The exoferalization patterns combined with the selective signals observed in this study paint an invasion scenario in which hybridization enriched genetic diversity and fostered local adaptation. When translating this to the bridgehead effect, we would in fact expect GRSM, a known invasive bridgehead, to be genetically diverse with low inbreeding depression (Ascunce et al., 2011; Lombaert et al., 2010). Otherwise, all secondary populations would quickly succumb to the effects of inbreeding and concomitant genetic load. Therefore, the observed genetic diversity and signatures of selection support the hypothesis that the GRSM population, descending from the wild boar introduced into Hooper Bald, would be a strong bridgehead population. We also found that other invasive populations surrounding GRSM had high diversity levels and similar haplotype structures at selective genetic sites, increasing the likelihood of a population achieving invasive success through local adaptation.

Critiques on the application of evolutionary theory in invasion ecological theory (including IBE) mainly stem from the fact that there is no research-based phenotypic or genetic evidence for adaptation in these cases (Bertelsmeier & Keller, 2018). Generally, such evolutionary invasion studies tend to only provide ancestry models showing that one invasive population contributed to the emergence of many new invasive populations (Javal et al., 2019). Here, we identify the genetic basis for adaptation, in combination with the pre-existing evidence that these animals are morphologically distinct from both wild boar and heritage breeds (Mayer & Brisbin, 1991). Additionally, our analyses provide a genetic basis for directional adaptation to the wild, contributing to feralization, as some of the genes under intense selection identified in the present work have been associated with traits involved in domestication.

Generally, when considering an invasive population as an invasive bridgehead, failed introductions are provided as evidence that a given invasive population is the only population that could be established (Keller et al., 2012; Rahbari et al., 2017). For example, in developing the concept of an IBE, Lombaert et al. (2010) demonstrated that multiple intentional introductions of *Harmonia axyridis* for pest control did not result in invasive populations. However, one particularly invasive *H. axyridis* population became established, resulting in secondary introductions. This is clearly not the case for wild pigs, given that multiple introductions are known that have led to self-propagating invasive populations (Mayer & Brisbin,1991). However, it is clear that the population in the Great Smoky Mountains region, descending from the Hooper Bald wild boar introduction, has served as a source for secondary introductions of invasive wild pigs into many areas not geographically linked to the GRSM. Historical records describe situations in which hunting associations actively sourced hybrid swine from the region for the purposes of establishing additional populations for recreational hunting (Lewis, 1966). The fact that the translocation of wild pigs from the Great Smoky Mountains region led to new invasive populations helps strengthen our conclusion that this population has adapted to be more invasive.

As for our hypothesis that exoferalization led to the formation of an invasive bridgehead, we argue that this might be a common phenomenon, though generally not described as such in the literature. Populations formed through exoferalization and invasive bridgehead populations often have one aspect in common: they are genetically admixed (Blumenfeld et al., 2021; Henriksen et al., 2018; Javal et al., 2019; Keller et al., 2012; Kruska, 2005). As such, we advocate for the inclusion of evolutionary hypotheses related to feralization in invasive bridgeheads to include species that have been domesticated. This would allow for greater insights into the evolutionary implications of admixture that have been restrictively classified as feralization.

5 | CONCLUSION

The current research provides support for the hypothesis that the heightened invasiveness of populations descending from the Hooper Bald wild boar introduction was (at least in part) attributable to exoferalization between wild boar and heritage breeds. The GRSM population and surrounding populations were found to be genetically diverse, enabling subsequent adaptation to novel ecosystems. The signatures of the selection observed among invasive wild pig populations are likely to be the result of directional selection against domestic phenotypes, such as loss of coat pigmentation. On the other hand, we also found evidence for selection acting on haplotypes derived from domestic pigs. These results suggest that the unique combination of wild and domestic alleles that occurred within GRSM allowed the population to function as an invasive bridgehead population, with descendant populations demonstrating increased invasiveness. As the spread of wild pigs has continued in the US, intensive containment and control efforts are needed. Research on the genetic make-up and adaptability of these animals increases our understanding of their invasive nature and allows for better regulation of these populations.

AUTHOR CONTRIBUTIONS

N.W.G.B. performed the analyses, with advice from M.B., T.J.S. and P.G.M.; W.H.S., J.G.Y., B.E.M.C., T.J.S. and A.J.P. were responsible

for tissue sample collection; T.J.S and A.J.P. oversaw genotyping; N.W.G.B., M.B., T.J.S. and P.G.M. wrote the paper with editorial assistance from A.J.P., D.W., W.H.S., J.G.Y. and B.E.M.C.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude for sample collection from our colleagues working in the field (National Park Service personnel within GRSM and Wildlife Services personnel within surrounding invasive populations); this work would not be possible without your valued efforts, support and contributions. Similarly, we thank the members of the WS-NWRC Wildlife Genetics Project for their input on the statistical analysis and modelling. The project was financially supported by the USDA APHIS WS National Feral Swine Damage Management Program. The findings and conclusions in this publication are those of the authors and should not be construed to represent an official US Government determination or policy. Mention of companies or commercial products does not imply recommendation or endorsement by the US Government over others not mentioned. Product names are mentioned solely to report factually on the methodologies used in the described analyses.

CONFLICT OF INTEREST STATEMENT

The authors declare no competing interests.

OPEN RESEARCH BADGES

This article has earned an Open Data badge for making publicly available the digitally-shareable data necessary to reproduce the reported results. The data is available at https://doi.org/10.48338/VU01-Q7LFEL.

DATA AVAILABILITY STATEMENT

All code can be found at: https://github.com/Niekbarmentlo/-Feral-swine-selective-sweep-pipeline. All genotyping data is available at https://doi.org/10.48338/VU01-Q7LFEL.

ORCID

Niek W. G. Barmentlo https://orcid.org/0000-0001-5256-3591
Patrick G. Meirmans https://orcid.org/0000-0002-6395-8107
Dominic Wright https://orcid.org/0000-0003-2329-2635
Timothy J. Smyser https://orcid.org/0000-0003-4542-3077
Mirte Bosse https://orcid.org/0000-0003-2433-2483

REFERENCES

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*(9), 1655–1664. https://doi.org/10.1101/gr.094052.109

Alexandri, P., Megens, H. J., Crooijmans, R. P. M. A., Groenen, M. A. M., Goedbloed, D. J., Herrero-Medrano, J. M., Rund, L. A., Schook, L. B., Chatzinikos, E., Triantaphyllidis, C., & Triantafyllidis, A. (2017). Distinguishing migration events of different timing for wild boar in the Balkans. *Journal of Biogeography*, *44*(2), 259–270. https://doi.org/10.1111/jbi.12861

Ascunce, M. S., Chin-Cheng, Y., Jane, O., Calcaterra, L., Wu, W.-J., Cheng-Jen, S., Goudet, J., Ross, K. G., & Shoemaker, D. (2011). Global invasion history of the fire AntSolenopsis invicta. *Science*, *331*(6020), 1066–1068. https://doi.org/10.1371/journal

Becker, D., Otto, M., Ammann, P., Keller, I., Drögemüller, C., & Leeb, T. (2015). The brown coat colour of Coppernecked goats is Asso. with a non-synonymous variant at the TYRP1 locus on chromosome 8. *Animal Genetics*, *46*(1), 50–54. https://doi.org/10.1111/age.12240

Bertelsmeier, C., & Keller, L. (2018). Bridgehead effects and role of adaptive evolution in invasive populations. *Trends in Ecology & Evolution*, *33*(7), 527–534. https://doi.org/10.1016/J.TREE.2018.04.014

Blumenfeld, A. J., Eyer, P. A., Husseneder, C., Mo, J., Johnson, L. N. L., Wang, C., Kenneth Grace, J., Chouvenc, T., Wang, S., & Vargo, E. L. (2021). Bridgehead effect and multiple introductions shape the global invasion history of a termite. *Communications Biology*, *4*(1), 196. https://doi.org/10.1038/s42003-021-01725-x

Bosse, M., Megens, H. J., Madsen, O., Paudel, Y., Frantz, L. A. F., Schook, L. B., Crooijmans, R. P. M. A., & Groenen, M. A. M. (2012). Regions of homozygosity in the porcine genome: Consequence of demography and the recombination landscape. *PLoS Genetics*, *8*(11), e1003100. https://doi.org/10.1371/journal.pgen.1003100

Bracke, M. B. M. (2011). Review of wallowing in pigs: Description of the behaviour and its motivational basis. *Applied Animal Behaviour Science*, *132*(1–2), 1–13. https://doi.org/10.1016/j.applanim.2011.01.002

Bratton, S. P., & Power, S. (1975). The effect of the European Wild Boar, Sus scrofa, on Gray Beech Forest in the Great Smoky Mountains. *Ecology*, *56*(6), 1356–1366.

Browning, B. L., & Browning, S. R. (2013). Improving the accuracy and efficiency of identity-by-descent detection in population data. *Genetics*, *194*(2), 459–471. https://doi.org/10.1534/genetics.113.150029

Browning, B. L., Zhou, Y., & Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *American Journal of Human Genetics*, *103*(3), 338–348. https://doi.org/10.1016/j.ajhg.2018.07.015

Buderman, F. E., Helm, P. J., Clark, J. D., Williamson, R. H., Yarkovich, J., & Mullinax, J. M. (2023). A multi-level modeling approach to guide management of female feral hogs in Great Smoky Mountains National Park. *Biological Invasions*, *25*, 3065–3082. https://doi.org/10.1007/s10530-023-03086-4

Burgos-Paz, W., Souza, C. A., Megens, H. J., Ramayo-Caldas, Y., Melo, M., Lemús-Flores, C., Caal, E., Soto, H. W., Martínez, R., Álvarez, L. A., Aguirre, L., Iñiguez, V., Revidatti, M. A., Martínez-López, O. R., Llambi, S., Esteve-Codina, A., Rodríguez, M. C., Crooijmans, R. P. M. A., Paiva, S. R., … Pérez-Enciso, M. (2013). Porcine colonization of the Americas: A 60k SNP story. *Heredity*, *110*(4), 321–330. https://doi.org/10.1038/hdy.2012.109

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: Rising to the challenge of larger and richer datasets. *GigaScience*, *4*(1), 7. https://doi.org/10.1186/s13742-015-0047-8

Chinn, S. M., Kilgo, J. C., Vukovich, M. A., & Beasley, J. C. (2021). Influence of intrinsic and extrinsic attributes on neonate survival in an invasive large mammal. *Scientific Reports*, *11*(1), 11033. https://doi.org/10.1038/s41598-021-90495-x

Chinn, S. M., Schlichting, P. E., Smyser, T. J., Bowden, C. F., & Beasley, J. C. (2022). Factors influencing pregnancy, litter size, and reproductive parameters of invasive wild pigs. *Journal of Wildlife Management*, *86*(8). https://doi.org/10.1002/jwmg.22304

Comeault, A. A., Wang, J., Tittes, S., Isbell, K., Ingley, S., Hurlbert, A. H., & Matute, D. R. (2020). Genetic diversity and thermal performance in invasive and native populations of African fig flies. *Molecular Biology and Evolution*, *37*(7), 1893–1906. https://doi.org/10.5061/dryad.866t1g1n3

Croney, C. C., Adams, K. M., Washington, C. G., & Stricklin, W. R. (2003). A note on visual, olfactory and spatial cue use in foraging behavior of pigs: Indirectly assessing cognitive abilities. *Applied Animal Behaviour Science*, *83*(4), 303–308. https://doi.org/10.1016/S0168-1591(03)00128-X

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Delaneau, O., Marchini, J., & Zagury, J. F. (2012). A linear complexity phasing method for thousands of genomes. *Nature Methods*, *9*(2), 179–181. https://doi.org/10.1038/nmeth.1785

Dinu, A. (2009). The action of the masticatory muscles and cranial changes in pigs as results of domestication. *Documenta Praehistorica*, *36*(1), 207–218. https://doi.org/10.4312/dp.36.13

Dlugosch, K. M., Cang, F. A., Barker, B. S., Andonian, K., Swope, S. M., & Rieseberg, L. H. (2015). Evolution of invasiveness through increased resource use in a vacant niche. *Nature Plants*, *1*(6), 1-5. https://doi.org/10.1038/NPLANTS.2015.66

Dlugosch, K. M., & Parker, I. M. (2008). Founding events in species invasions: Genetic variation, adaptive evolution, and the role of multiple introductions. *Molecular Ecology*, *17*(1), 431–449. https://doi.org/10.1111/j.1365-294X.2007.03538.x

Durante, M. A., Kurtenbach, S., Sargi, Z. B., Harbour, J. W., Choi, R., Kurtenbach, S., Goss, G. M., Matsunami, H., & Goldstein, B. J. (2020). Single-cell analysis of olfactory neurogenesis and differentiation in adult humans. *Nature Neuroscience*, *23*(3), 323–326. https://doi.org/10.1038/s41593-020-0587-9

Estoup, A., Ravigné, V., Hufbauer, R., Vitalis, R., Gautier, M., & Facon, B. (2016). Is there a genetic paradox of biological invasion? *Annual Review of Ecology, Evolution, and Systematics*, *47*, 51–72. https://doi.org/10.1146/annurev-ecolsys-121415-032116

Fowles, L. F., Bennetts, J. S., Berkman, J. L., Williams, E., Koopman, P., Teasdale, R. D., & Wicking, C. (2003). Genomic screen for genes involved in mammalian craniofacial development. *Genesis*, *35*(2), 73–87. https://doi.org/10.1002/gene.10165

Fulgione, D., Rippa, D., Buglione, M., Trapanese, M., Petrelli, S., & Maselli, V. (2016). Unexpected but welcome. Artificially selected traits may increase fitness in wild boar. *Evolutionary Applications*, *9*(6), 769–776. https://doi.org/10.1111/eva.12383

Fulgione, D., Trapanese, M., Buglione, M., Rippa, D., Polese, G., Maresca, V., & Maselli, V. (2017). Pre-birth sense of smell in the wild boar: The ontogeny of the olfactory mucosa. *Zoology*, *123*, 11–15. https://doi.org/10.1016/j.zool.2017.05.003

Gautier, M., & Vitalis, R. (2012). rehh: An R package to detect footprints of selection in genome-wide SNP data from haplotype structure. *Bioinformatics*, *28*(8), 1176–1177. https://doi.org/10.1093/bioinformatics/bts115

Gering, E., Incorvaia, D., Henriksen, R., Conner, J., Getty, T., & Wright, D. (2019). Getting back to nature: Feralization in animals and plants. *Trends in Ecology and Evolution*, *34*(12), 1137–1151. https://doi.org/10.1016/j.tree.2019.07.018

Goedbloed, D. J., Megens, H. J., Van Hooft, P., Herrero-Medrano, J. M., Lutz, W., Alexandri, P., Crooijmans, R. P. M. A., Groenen, M., Van Wieren, S. E., Ydenberg, R. C., & Prins, H. H. T. (2013). Genome-wide single nucleotide polymorphism analysis reveals recent genetic introgression from domestic pigs into Northwest European wild boar populations. *Molecular Ecology*, *22*(3), 856–866. https://doi.org/10.1111/j.1365-294X.2012.05670.x

Henriksen, R., Gering, E., & Wright, D. (2018). Feralisation-the understudied counterpoint to domestication. In *Origin and evolution of biodiversity* (pp. 183–195). Springer International Publishing. https://doi.org/10.1007/978-3-319-95954-2_11

Holsinger, K. E., & Weir, B. S. (2009). Genetics in geographically structured populations: Defining, estimating and interpreting FST. *Nature Reviews Genetics*, *10*(9), 639–650. https://doi.org/10.1038/nrg2611

Huang, D. W., Sherman, B. T., & Lempicki, R. A. (2009). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, *37*(1), 1–13. https://doi.org/10.1093/nar/gkn923

Iacolina, L., Scandura, M., Goedbloed, D. J., Alexandri, P., Crooijmans, R. P. M. A., Larson, G., Archibald, A., Apollonio, M., Schook, L. B., Groenen, M. A. M., & Megens, H. J. (2016). Genomic diversity and differentiation of a managed Island wild boar population. *Heredity*, *116*(1), 60–67. https://doi.org/10.1038/hdy.2015.70

Ito, S., & Wakamatsu, K. (2003). Quantitative analysis of eumelanin and Pheomelanin in humans, mice, and other animals: A comparative review. *Pigment Cell Research*, *16*(5), 523–531.

Javal, M., Lombaert, E., Tsykun, T., Courtin, C., Kerdelhué, C., Prospero, S., Roques, A., & Roux, G. (2019). Deciphering the worldwide invasion of the Asian long-horned beetle: A recurrent invasion process from the native area together with a bridgehead effect. *Molecular Ecology*, *28*(5), 951–967. https://doi.org/10.1111/mec.15030

Jenkins, M. A. (2007). Vegetation communities of Great Smoky Mountains National Park. *Southeastern Naturalist*, *6*(sp2), 35–56. https://doi.org/10.1656/1528-7092(2007)6[35:vcogsm]2.0.co;2

Johnsson, M., Gering, E., Willis, P., Lopez, S., Van Dorp, L., Hellenthal, G., Henriksen, R., Friberg, U., & Wright, D. (2016). Feralisation targets different genomic loci to domestication in the chicken. *Nature Communications*, *7*, 12950. https://doi.org/10.1038/ncomms12950

Keller, S. R., Gilbert, K. J., Fields, P. D., & Taylor, D. R. (2012). Bayesian inference of a complex invasion history revealed by nuclear and chloroplast genetic diversity in the colonizing plant, *Silene latifolia*. *Molecular Ecology*, *21*(19), 4721–4734. https://doi.org/10.1111/j.1365-294X.2012.05751.x

Kolbe, J. J., Glor, R. E., Schettino, L. R., Lara, A. C., Larson, A., & Losos, J. B. (2004). Genetic variation increases during biological invasion by a Cuban lizard. *Nature*, *431*(7005), 177–181. https://doi.org/10.1038/nature02807

Kruska, D. (1970). Vergleichend cytoarchitektonisehe Untersuchungen an Getfirnen yon Wild-und Hausschweinen. *Anatomy and Embryology*, *131*, 291–324.

Kruska, D. C. T. (2005). On the evolutionary significance of encephalization in some Eutherian mammals: Effects of adaptive radiation, domestication, and feralization. *Brain, Behavior and Evolution*, *65*(2), 73–108. https://doi.org/10.1159/000082979

Lagutin, O. V., Zhu, C. C., Kobayashi, D., Topczewski, J., Shimamura, K., Puelles, L., Russell, H. R. C., McKinnon, P. J., Solnica-Krezel, L., & Oliver, G. (2003). Six3 repression of Wnt signaling in the anterior neuroectoderm is essential for vertebrate forebrain development. *Genes and Development*, *17*(3), 368–379. https://doi.org/10.1101/gad.1059403

Lavergne, S., & Molofsky, J. (2007). Increased genetic variation and evolutionary potential drive the success of an invasive grass. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(10), 3883–3888.

Lewis, J. C. (1966). Observations of pen-reared European hogs released for stocking. *The Journal of Wildlife Management*, *30*(4), 832–835.

Lewis, J. S., Farnsworth, M. L., Burdett, C. L., Theobald, D. M., Gray, M., & Miller, R. S. (2017). Biotic and abiotic factors predicting the global distribution and population density of an invasive large mammal. *Scientific Reports*, *7*, 44152. https://doi.org/10.1038/srep44152

Li, J., Bed'hom, B., Marthey, S., Valade, M., Dureux, A., Moroldo, M., Péchoux, C., Coville, J. L., Gourichon, D., Vieaud, A., Dorshorst, B., Andersson, L., & Tixier-Boichard, M. (2019). A missense mutation in TYRP1 causes the chocolate plumage color in chicken and alters melanosome structure. *Pigment Cell and Melanoma Research*, *32*(3), 381–390. https://doi.org/10.1111/pcmr.12753

Liu, L., Megens, H. J., Crooijmans, R. P. M. A., Bosse, M., Huang, Q., Van Sonsbeek, L., Groenen, M. A. M., & Madsen, O. (2022). The Visayan warty pig (*Sus cebifrons*) genome provides insight into chromosome evolution and sensory adaptation in pigs. *Molecular Biology and Evolution*, *39*(6), msac110. https://doi.org/10.1093/molbev/msac110

Lombaert, E., Guillemaud, T., Cornuet, J. M., Malausa, T., Facon, B., & Estoup, A. (2010). Bridgehead effect in the worldwide invasion of the biocontrol harlequin ladybird. *PLoS One*, *5*(3), e9743. https://doi.org/10.1371/journal.pone.0009743

Lowe, S., Browne, M., Boudjelas, S., & De Poorter, M. (2000). *100 Of the world's worst invasive alien species: A selection from the global invasive species database*. Invasive Species Specialist Group. www.issg.org/booklet.pdf

Murat Maga, A., Navarro, N., Cunningham, M. L., & Cox, T. C. (2015). Quantitative trait loci affecting the 3D skull shape and size in mouse and prioritization of candidate genes in-silico. *Frontiers in Physiology*, *6*, 92. https://doi.org/10.3389/fphys.2015.00092

Maselli, V., Polese, G., Larson, G., Raia, P., Forte, N., Rippa, D., Ligrone, R., Vicidomini, R., & Fulgione, D. (2014). A dysfunctional sense of smell: The irreversibility of olfactory evolution in free-living pigs. *Evolutionary Biology*, *41*(2), 229–239. https://doi.org/10.1007/s11692-013-9262-3

Mayer, J. J., & Brisbin, I. L. (1991). *Wild pigs in the United States: Their history, comparative morphology, and current status*. University of Georgia Press.

McKinnon, J. S., & Pierotti, M. E. R. (2010). Colour polymorphism and correlated characters: Genetic mechanisms and evolution. *Molecular Ecology*, *19*(23), 5101–5125. https://doi.org/10.1111/j.1365-294X.2010.04846.x

Meyermans, R., Gorssen, W., Buys, N., & Janssens, S. (2020). How to study runs of homozygosity using plink? A guide for analyzing medium density snp data in livestock and pet species. *BMC Genomics*, *21*(1), 94. https://doi.org/10.1186/s12864-020-6463-x

Newell, C., Walker, H., & Caro, T. (2021). Pig pigmentation: testing Gloger's rule. *Journal of Mammalogy*, *102*(6), 1525–1535. https://doi.org/10.1093/jmammal/gyab090

North, H. L., McGaughran, A., & Jiggins, C. D. (2021). Insights into invasive species from whole-genome resequencing. *Molecular Ecology*, *30*(23), 6289–6308. https://doi.org/10.1111/mec.15999

Petrelli, S., Buglione, M., Maselli, V., Troiano, C., Larson, G., Frantz, L., Manin, A., Ricca, E., Baccigalupi, L., Wright, D., Pietri, C., & Fulgione, D. (2022). Population genomic, olfactory, dietary, and gut microbiota analyses demonstrate the unique evolutionary trajectory of feral pigs. *Molecular Ecology*, *31*(1), 220–237. https://doi.org/10.1111/mec.16238

Rahbari, M., Rahlfs, S., Jortzik, E., Bogeski, I., & Becker, K. (2017). H_2O_2 dynamics in the malaria parasite *Plasmodium falciparum*. *PLoS One*, *12*(4), e0174837. https://doi.org/10.1371/journal.pone.0174837

Ramos, A. M., Crooijmans, R. P. M. A., Affara, N. A., Amaral, A. J., Archibald, A. L., Beever, J. E., Bendixen, C., Churcher, C., Clark, R., Dehais, P., Hansen, M. S., Hedegaard, J., Hu, Z. L., Kerstens, H. H., Law, A. S., Megens, H. J., Milan, D., Nonneman, D. J., Rohrer, G. A., … Groenen, M. A. M. (2009). Design of a high density SNP genotyping assay in the pig using SNPs identified and characterized by next generation sequencing technology. *PLoS ONE*, *4*(8). https://doi.org/10.1371/journal.pone.0006524

Ren, J., Mao, H., Zhang, Z., Xiao, S., Ding, N., & Huang, L. (2011). A 6-bp deletion in the TYRP1 gene causes the brown colouration phenotype in Chinese indigenous pigs. *Heredity*, *106*(5), 862–868. https://doi.org/10.1038/hdy.2010.129

Roberts, K. S., & Lamberson, W. R. (2015). Relationships among and variation within rare breeds of swine. *American Society of Animal Science*, *93*, 3810–3813.

Sabeti, P. C., Varilly, P., Fry, B., Lohmueller, J., Hostetter, E., Cotsapas, C., Xie, X., Byrne, E. H., McCarroll, S. A., Gaudet, R., Schaffner, S. F., Lander, E. S., Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Gibbs, R. A., Belmont, J. W., … Stewart, J. (2007). Genome-wide detection and characterization of positive selection in human populations. *Nature*, *449*(7164), 913–918. https://doi.org/10.1038/nature06250

Salinas, R. A., Stiver, W. H., Corn, J. L., Lenhart, S., Collins, C., Madden, M., Vercauteren, K. C., Schmit, B. B., Kasari, E., Odoi, A., Hickling, G., & Mccallum, H. (2015). An individual-based model for feral hogs in great smoky mountains national park. *Natural Resource Modeling*, *28*(1), 18–36. https://doi.org/10.1111/nrm.12055

Scott, C. D. (1973). *Seasonal food habits of European Wild Hogs (Sus Scrofa) in the Great Smoky Mountains National Park*. [Master's thesis, University of Tennessee]. https://trace.tennessee.edu/utk_gradthes

Smyser, T. J., Pfaffelhuber, P., Giglio, R. M., DeSaix, M. G., Davis, A. J., Bowden, C. F., Tabak, M. A., Manunza, A., Bâlteanu, V. A., Amills, M., Iacolina, L., Walker, P., Lessard, C., & Piaggio, A. J. (2024). Probabilistic genetic identification of wild boar hybridization to support control of invasive wild pigs (*Sus scrofa*). *Ecosphere*, *15*(2), e4774. https://doi.org/10.1002/ecs2.4774

Smyser, T. J., Tabak, M. A., Slootmaker, C., Robeson, M. S., Miller, R. S., Bosse, M., Megens, H. J., Groenen, M. A. M., Paiva, S. R., de Faria, D. A., Blackburn, H. D., Schmit, B. S., & Piaggio, A. J. (2020). Mixed ancestry from wild and domestic lineages contributes to the rapid expansion of invasive feral swine. *Molecular Ecology*, *29*(6), 1103–1119. https://doi.org/10.1111/mec.15392

Stegeman, L. C. (1938). American Society of Mammalogists The European Wild Boar in the Cherokee National Forest. *Source. Journal of Mammalogy*, *19*(3), 279–290.

Tang, K., Thornton, K. R., & Stoneking, M. (2007). A new approach for using genome scans to detect recent positive selection in the human genome. *PLoS Biology*, *5*(7), 1587–1602. https://doi.org/10.1371/journal.pbio.0050171

Utzeri, V. J., Ribani, A., & Fontanesi, L. (2014). A premature stop codon in the TYRP1 gene is associated with brown coat colour in the European rabbit (*Oryctolagus cuniculus*). *Animal Genetics*, *45*(4), 600–603. https://doi.org/10.1111/age.12171

Van Camp, G., Snoeckx, R. L., Hilgert, N., Van Den Ende, J., Fukuoka, H., Wagatsuma, M., Suzuki, H., Smets, R. M. E., Vanhoenacker, F., Declau, F., Van De Heyning, P., & Usami, S.-I. (2006). A new autosomal recessive form of stickler syndrome is caused by a mutation in the COL9A1 gene. *The American Journal of Human Genetics*, *79*, 449–457.

Vissers, L. E. L. M., Cox, T. C., Maga, A. M., Short, K. M., Wiradjaja, F., Janssen, I. M., Jehee, F., Bertola, D., Liu, J., Yagnik, G., Sekiguchi, K., Kiyozumi, D., van Bokhoven, H., Marcelis, C., Cunningham, M. L., Anderson, P. J., Boyadjiev, S. A., Passos-Bueno, M. R., Veltman, J. A., … Roscioli, T. (2011). Heterozygous mutations of FREM1 are associated with an increased risk of isolated metopic craniosynostosis in humans and mice. *PLoS Genetics*, *7*(9), e1002278. https://doi.org/10.1371/journal.pgen.1002278

Voight, B. F., Kudaravalli, S., Wen, X., & Pritchard, J. K. (2006). A map of recent positive selection in the human genome. *PLoS Biology*, *4*(3), 446–458. https://doi.org/10.1371/journal.pbio.0040072

Weir, B. S., & Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population. *Structure*, *38*(6), 1358–1370.

Welles, S. R., & Dlugosch, K. M. (2018). *Population genomics of colonization and invasion* (pp. 655–683). Springer.

White, S. (2011). From globalized pig breeds to capitalist pigs: A study in animal cultures and evolutionary history. *Environmental history*, *16*(1), 94–120. https://doi.org/10.1093/envhis/emq143

Wilkins, A. S., Wrangham, R. W., & Tecumseh Fitch, W. (2014). The "domestication syndrome" in mammals: A unified explanation based on neural crest cell behavior and genetics. *Genetics*, *197*(3), 795–808. https://doi.org/10.1534/genetics.114.165423

Wright, D., Henriksen, R., & Johnsson, M. (2020). Defining the domestication syndrome: Comment on Lord et al. 2020. *Trends in Ecology*

& Evolution, *35*(2), 1059–1060. https://doi.org/10.1016/j.tree.2020.08.009

Wu, W., Li, X., Zhou, R., & Li, L. (2012). Bioinformatics analysis of TYR within and among species. *African Journal of Microbiology Research*, *6*(5), 1069–1074. https://doi.org/10.5897/AJMR11.1483

Wu, X., Zhang, Y., Shen, L., Du, J., Luo, J., Liu, C., Pu, Q., Yang, R., Li, X., Bai, L., Tang, G., Zhang, S., & Zhu, L. (2016). A 6-bp deletion in exon 8 and two mutations in introns of TYRP1 are associated with blond coat color in Liangshan pigs. *Gene*, *578*(1), 132–136. https://doi.org/10.1016/j.gene.2015.12.011

Yang, B., Cui, L., Perez-Enciso, M., Traspov, A., Crooijmans, R. P. M. A., Zinovieva, N., Schook, L. B., Archibald, A., Gatphayak, K., Knorr, C., Triantafyllidis, A., Alexandri, P., Semiadi, G., Hanotte, O., Dias, D., Dovč, P., Uimari, P., Iacolina, L., Scandura, M., … Megens, H. J. (2017). Genome-wide SNP data unveils the globalization of domesticated pigs. *Genetics Selection Evolution*, *49*(1), 71. https://doi.org/10.1186/s12711-017-0345-y

Zhou, X., Xin, J., Fan, N., Zou, Q., Huang, J., Ouyang, Z., Zhao, Y., Zhao, B., Yi, X., Guo, L., Esteban, M. A., Zeng, Y., Yang, H., & Lai, L. (2015). Generation of CRISPR/Cas9-mediated gene-targeted pigs via somatic cell nuclear transfer. *Cellular and Molecular Life Sciences*, *72*(6), 1175–1184. https://doi.org/10.1007/s00018-014-1744-7

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Barmentlo, N. W. G., Meirmans, P. G., Stiver, W. H., Yarkovich, J. G., McCann, B. E., Piaggio, A. J., Wright, D., Smyser, T. J., & Bosse, M. (2024). Natural selection on feralization genes contributed to the invasive spread of wild pigs throughout the United States. *Molecular Ecology*, *33*, e17383. https://doi.org/10.1111/mec.17383