


## ARTICLE

## Methods, Tools, and Technologies

# Probabilistic genetic identification of wild boar hybridization to support control of invasive wild pigs (*Sus scrofa*)

Timothy J. Smyser<sup>1</sup> | Peter Pfaffelhuber<sup>2</sup> | Rachael M. Giglio<sup>1</sup> |  
 Matthew G. DeSaix<sup>1,3</sup> | Amy J. Davis<sup>1</sup>  | Courtney F. Bowden<sup>1</sup> |  
 Michael A. Tabak<sup>4</sup> | Arianna Manunza<sup>5</sup> | Valentin Adrian Bâlțeanu<sup>6</sup> |  
 Marcel Amills<sup>7</sup> | Laura Iacolina<sup>8,9</sup> | Pamela Walker<sup>10</sup> | Carl Lessard<sup>10</sup> |  
 Antoinette J. Piaggio<sup>1</sup>

<sup>1</sup>United States Department of Agriculture, Animal and Plant Health Inspection Service, Wildlife Services, National Wildlife Research Center, Fort Collins, Colorado, USA

<sup>2</sup>Department of Mathematical Stochastics, University of Freiburg, Freiburg im Breisgau, Baden-Württemberg, Germany

<sup>3</sup>Department of Biology, Colorado State University, Fort Collins, Colorado, USA

<sup>4</sup>Quantitative Science Consulting, LLC, Melancthon, Ontario, Canada

<sup>5</sup>National Research Council, Institute of Agricultural Biology and Biotechnology, Milan, Italy

<sup>6</sup>Laboratory of Genomics, Animal Breeding, Biodiversity and Molecular Pathology, University of Agricultural Sciences and Veterinary Medicine Cluj-Napoca, Cluj-Napoca, Romania

<sup>7</sup>Centre for Research in Agricultural Genomics (CRAG), CSIC-IRTA-UAB-UB, Universitat Autònoma de Barcelona, Bellaterra, Catalonia, Spain

<sup>8</sup>Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia

<sup>9</sup>Department of Veterinary Medicine, University of Sassari, Sassari, Italy

<sup>10</sup>Animal Genetic Resources of Canada, Agriculture and Agri-Food Canada, Saskatoon, Saskatchewan, Canada

**Correspondence**

Timothy J. Smyser

Email: [timothy.j.smyser@usda.gov](mailto:timothy.j.smyser@usda.gov)

**Funding information**

USDA-APHIS-WS-National Feral Swine Damage Management Program; USDA-APHIS-WS-National Wildlife Disease Program; Ministry of Research, Innovation and Digitization of Romania, Grant/Award Number: TE 113/2022

**Handling Editor:** Matthew A. Mumma

**Abstract**

The rapid expansion of wild pigs (*Sus scrofa*) throughout the United States has been fueled by unlawful introductions, with invasive populations causing extensive crop losses, damaging native ecosystems, and serving as a reservoir for disease. Multiple states have passed laws prohibiting the possession or transport of wild pigs. However, genetic and phenotypic similarities between domestic pigs and invasive wild pigs—which overwhelmingly represent domestic pig and wild boar hybrids—pose a challenge for the enforcement of such regulations. We sought to exploit wild boar ancestry as a common attribute among the vast majority of invasive wild pigs as a means of genetically differentiating wild pigs from breeds of domestic pig found within the United States. We organized reference high-density single nucleotide polymorphism genotypes (1039 samples from 33 domestic breeds and 382 samples from

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2024 The Authors. *Ecosphere* published by Wiley Periodicals LLC on behalf of The Ecological Society of America. This article has been contributed to by U.S. Government employees and their work is in the public domain in the USA.

16 wild boar populations) into five genetically cohesive reference groups: mixed-commercial breeds, Durocs, heritage breeds, primitive breeds, and wild boar. Building upon well-established genetic clustering approaches, we structured the test statistic to describe the difference in the likelihood of a given genotype's ancestry vectors (sensu genetic clustering analysis) if derived strictly from the four described domestic pig reference groups versus allowing for admixture from the wild boar group. By fitting statistical distributions to test statistics of reference domestic pigs, we characterized the distribution of the null hypothesis that a given genotype descends strictly from domestic pig reference groups. We tested the approach with simulated genotypes and empirical data from an additional 29 breeds of domestic pig represented by 435 unique genotypes; all associated test statistics for simulated and empirical domestic pig challenge sets fell within the distribution of reference domestic pigs. We then evaluated 6566 invasive wild pigs sampled across the contiguous United States, of which 63% exceeded the maximum threshold for domestic pigs and could be statistically classified as possessing wild boar ancestry. This approach provides a scientific foundation to enforce regulations prohibiting the possession of this destructive invasive species. Further, this computationally efficient and generalizable approach could be readily adapted to quantify gene flow among ecological systems of conservation or management concern.

#### KEYWORDS

ancestry, exoferalization, feral swine, genetic clustering, genetic identification, hybridization, invasive species, single nucleotide polymorphism, *Sus scrofa*, wild pig

## INTRODUCTION

Humans, both intentionally and inadvertently, have transported organisms across biogeographic barriers and introduced alien species into nonnative habitats (Bertelsmeier & Keller, 2018; Hulme, 2009; Lowe et al., 2000; Mack et al., 2000). With the rate of such introduction accelerating from the Age of Exploration (1450–1650) through the Industrial Revolution (1760–1840) and into the present Era of Globalization (1870–present), invasive species have become widely recognized as among the most critical threats for the persistence of endemic species and ecosystems (Bellard et al., 2016; Hulme, 2009; Mack et al., 2000; Wilcove et al., 1998). Wilcove et al.'s (1998) examination of the causative agents of decline for species protected under the United States Endangered Species Act identified invasive species as a threat to 49% of listed species—second only to habitat destruction and fragmentation. Similarly, Bellard et al. (2016) identified invasive species as the second most common threat among IUCN (International Union for Conservation of Nature) Red

List species that are characterized as extinct or extinct in the wild, with vertebrates and island endemic species being most severely impacted. Various biological attributes of both the invasive species and the ecosystems into which they are introduced influence the full consequences of such introductions; however, the species with the greatest effect on invaded ecosystems are generally those that alter ecosystem processes (Lowe et al., 2000; Mack et al., 2000; Parker et al., 1999).

Invasive wild pigs have been identified as among the most ecologically destructive invasive species in the world, introducing novel rooting and wallowing behaviors that alter ecosystem processes and restructure ecological communities within invaded habitats (Barrios-Garcia & Simberloff, 2013; Boughton & Boughton, 2014; Cushman et al., 2004; Lowe et al., 2000). Further, wild pigs are highly fecund and very adaptable, which allows this species to readily invade both human-dominated and natural ecosystems (Ballari & Barrios-Garcia, 2014; Comer & Mayer, 2009; Gaillard et al., 1998; Koen et al., 2018; Lewis et al., 2017; Petrelli et al., 2022). Although wild pig diets are largely comprised of plant matter, studies

have demonstrated that predation pressure exerted by wild pigs upon prey populations within invaded habitats was comparatively greater than predation pressure exerted by wild boar within their native range (Ballari & Barrios-Garcia, 2014). The overrepresentation of a limited number of prey items in diet analyses has demonstrated that wild pigs switch to seasonally available prey items with an intensity of predation pressure that may diminish or extirpate local prey populations (Strickland et al., 2020). Beyond ecological damage, invasion by wild pigs can also be economically costly, with the greatest losses incurred within the agricultural sector (Anderson et al., 2016; Carlisle et al., 2021; Fern et al., 2021; McKee et al., 2020; Poudyal et al., 2017). Extensive damage by invasive wild pigs has been reported for many crop types, with the potential to greatly reduce yields of corn, peanuts, and pecan within the invaded range (Anderson et al., 2016; McKee et al., 2020). Similarly, with livestock production, wild pigs have been associated with direct predation on livestock, reduced livestock weight conversion due to wild pig-induced stress or consumption of feed, and damage to pasture lands and related infrastructure (Anderson et al., 2019; Carlisle et al., 2021; Miller et al., 2017). Given that wild pigs serve as a reservoir for numerous diseases that have been eliminated from domestic herds, field observations of wild pigs accessing either livestock feed or areas where livestock are kept demonstrate a sustained risk of disease spillover (Anderson et al., 2019; Brown et al., 2020, 2021; Carlisle et al., 2021). Additionally, wild pigs serve as a unique risk for the introduction of foreign animal diseases (e.g., African swine fever, classic swine fever, or foot-and-mouth disease), with the potential for a novel pathogen to establish and spread undetected among wild populations and subsequently spillover into domestic herds with great economic costs (Brown et al., 2020, 2021).

Although invasive wild pigs represent a global challenge, this species is of particular concern in the United States where abundance and distribution have expanded exponentially since the 1980s (McClure et al., 2015; Nolte & Anderson, 2015; Snow et al., 2017). In the absence of a native wild suid, any free-ranging swine encountered in the United States would be considered an invasive wild pig regardless of domestic origin (derived from livestock production systems or the pet trade [i.e., Vietnamese potbellied pigs]), wild origin (Eurasian wild boar), or hybrids of these lineages. However, previous genetic analysis has demonstrated that the vast majority of invasive wild pigs found within the United States represent domestic pig-wild boar hybrids (Smyser et al., 2020). Free-ranging pig populations were first established in North America in the 1500s as a

consequence of deliberate introductions by the Spanish as a means of provisioning exploratory forces (Mayer & Brisbin, 1991; Zadik, 2005). From the initial introduction through the mid-1900s, populations of feral pigs that would have descended strictly from domestic stock were incidentally augmented through the escape of domestic pigs reared using traditional husbandry practices (pannage), in which animals were seasonally released into forested ecosystems to fatten on fallen mast (Mayer & Brisbin, 1991; White, 2011). However, with growing interest in recreational hunting during the late 1800s and early 1900s, wild boar were imported to the United States from native populations in Europe to stock private hunting preserves (Bratton, 1975; Mayer & Brisbin, 1991; Stegeman, 1938). Ultimately, wild boar escaped from preserves and began to interbreed with established populations of feral pigs, thus producing initial hybrid populations (Mayer & Brisbin, 1991).

Hybrid populations, with a phenotype intermediate to that of wild boar and domestic pigs, possessed characteristics that were deemed to be desirable for hunting (Mayer & Brisbin, 1991). Accordingly, from the 1920s through the 1970s, hybrid populations were frequently used as a source by state wildlife management agencies, commercial hunting preserves, and private individuals for the purposes of establishing new huntable populations (i.e., introduction) or used to augment established populations as a means of increasing the phenotypic appeal of local animals (Mayer & Brisbin, 1991). This complex history of hybridization and translocation, with some evidence of increased fitness of hybrid animals (Chinn et al., 2021; Fulgione et al., 2016; Mayer & Brisbin, 1991; Waithman et al., 1999), has given rise to contemporary populations of invasive wild pigs in which the vast majority of animals (96.6%) observed within the contiguous United States are of mixed ancestry with associations to both Western heritage breeds of domestic pig and European wild boar (Smyser et al., 2020).

As the era of widespread translocations for the creation of hunting opportunities transitioned to recognition of the ecological destruction and economic cost caused by wild pigs, various federal and state agencies have sought to mitigate the damage and curtail the spread of this invasive species through the enactment of new regulations and the initiation of control efforts. Despite the shifting paradigm in wild pig management, the rapid expansion of invasive wild pigs observed since the 1980s has been driven by anthropogenic translocation, with animals collected from established populations and released into uninvaded habitats (Hernández et al., 2018; Smyser et al., 2020; Tabak et al., 2017). Such unlawful translocation represents a tragedy of the commons (Hardin, 1968) in which individuals releasing wild pigs

seek to attain some utility from their presence, putatively related to future hunting opportunities, whereas the ecological and economic costs are borne by many. To deter unauthorized introductions, various jurisdictions have passed laws prohibiting the possession, transport, or release of invasive wild pigs (e.g., Michigan Compiled Laws § 324.41303; Michigan Compiled Laws § 324.41305; Michigan Department of Natural Resources Invasive Species Order 40.4(b); Mississippi Code of Rules 40-2-7.1; Mississippi House Bill 1069; Missouri House Bill 369; Missouri Revised Statute § 270.010). However, the enforcement of such prohibitions is complicated by multiple factors: (1) invasive wild pigs are genetically and phenotypically difficult to differentiate from domestic pigs (Lorenzini et al., 2020); (2) ancestry analyses (sensu Smyser et al., 2020) provide insufficient resolution to differentiate some pig breeds (i.e., those that are most genetically similar to wild boar) from invasive wild pigs; and (3) wild pig populations are genetically heterogeneous, in contrast to the genetic management imposed upon domestic breeds, with populations influenced by unique histories of introduction, gene flow among populations, and natural selection. Accordingly, our goal was to develop a statistical test to genetically differentiate invasive wild pigs from domestic pigs. Given that populations of wild pigs are genetically varied across the invaded range, we sought to leverage the genetic signature of wild boar hybridization—a genetic characteristic shared among the vast majority of wild pigs in the contiguous United States—as a means of differentiating these invasive animals from breeds of domestic pig. Our specific objective was to describe the probability that a high-resolution single nucleotide polymorphism (SNP) genotype for an individual of unknown origin could be attributable to allele frequencies found strictly among domestic breeds versus requiring wild boar hybridization.

## METHODS

### Assembly of genetic reference set

Differentiating invasive wild pigs encountered in the contiguous United States from domestic breeds required the assembly of a specific genetic reference set as opposed to the more general approach used previously to describe ancestral contributions to invasive wild pigs from the global domestic pig–wild boar species complex (Smyser et al., 2020). To meet this specific objective, we restricted the reference set to European populations of wild boar (previous work demonstrated no association of invasive wild pigs with Asian lineages of

wild boar; Smyser et al., 2020), commercial breeds (represent the majority of animals raised for the US pork market), and Western heritage breeds (characterized as traditional, historic, or heirloom breeds of European origin or North American origin descending from European stock with a breed description that predates industrial agriculture; Livestock Conservancy, 2023; Appendix S1: Figure S1). Although heritage breeds may represent only a small fraction of US pork production, such breeds are the most genetically proximate to wild boar and, concomitantly, the most difficult to differentiate (Smyser et al., 2020). Accordingly, we deliberately included both major commercial breeds and heritage breeds in the reference set to minimize the risk of misclassifying a domestic pig of any breed as possessing wild boar ancestry. We assembled the reference set from previously published high-resolution SNP genotypes, restricting analysis to genotypes produced with Illumina BeadArray technology (San Diego, California) across multiple commercially available arrays (Illumina PorcineSNP60, Illumina PorcineSNP60 v2, Genomic Profiler for Porcine HD, licensed exclusively to GeneSeek, a Neogen Corporation, Lansing, Michigan; Ramos et al., 2009). We augmented previously published genotypes (detailed in Smyser et al., 2020) with a subset of novel genotypes produced for this study (Appendix S1: Table S1). We restricted our analyses to loci that were available across all datasets (influenced by loci shared across arrays and the extent to which publicly available datasets were filtered by authors prior to publication) and mapped to autosomes (Sscrofa11.1 genome assembly; Warr et al., 2020). In sum, we included 33 breeds and 16 populations of European wild boar, representing a total of 1421 reference samples (Table 1) genotyped at 28,545 biallelic loci.

With reference genotypes identified, we used a combination of ADMIXTURE version 1.3.0 (Alexander et al., 2009) and principal components analysis (R package adegenet; Jombart, 2008; R version 4.3.0; R Core Team, 2023) to characterize the genetic structure within the reference set and organize reference genotypes into genetically cohesive groups (Appendix S1: Figures S1–S3). Combining the results of these efforts, we identified the following five reference groups: (1) mixed-commercial breeds (Landrace and Yorkshire [also commonly referred to as Large Whites] often reared with intensive husbandry practices), (2) Durocs (a common commercial breed genetically distinct from other commercial breeds), (3) heritage breeds (18 traditional breeds of pigs, generally raised with extensive husbandry practices), (4) primitive breeds (a unique suite of 12 ancient breeds, distinct from heritage breeds based on their genetic similarity to wild boar), and (5) European wild boar. With the delineation of these five reference

**TABLE 1** Organization of reference *Sus scrofa* high-density single nucleotide polymorphism genotypes, representing 33 breeds of domestic pig and 16 populations of wild boar, into five reference groups to inform the statistical differentiation of wild boar and wild boar hybrids from domestic pigs.

| Reference group    | Breed/population      | Sample size    |
|--------------------|-----------------------|----------------|
| Mixed-commercial   |                       | <i>310</i>     |
|                    | Landrace              | 121            |
|                    | Yorkshire/Large White | 189            |
| Duroc              |                       | <i>159</i>     |
|                    | Duroc                 | 159            |
| Heritage           |                       | <i>381</i>     |
|                    | Berkshire             | 80             |
|                    | Choctaw               | 10             |
|                    | Cuino                 | 7              |
|                    | Gloucester Old Spot   | 20             |
|                    | Guinea Hog            | 39             |
|                    | Hairless              | 6              |
|                    | Hampshire             | 68             |
|                    | Large Black           | 23             |
|                    | Leicoma               | 9              |
|                    | Monteiro              | 9              |
|                    | Moura                 | 8              |
|                    | Mulefoot              | 24             |
|                    | Ossabaw               | 17             |
|                    | Piau                  | 8              |
|                    | Poland China          | 4              |
|                    | Red White Belted      | 5              |
|                    | Spotted               | 14             |
|                    | Tamworth              | 30             |
|                    | Primitive             |                |
| Black Slavonian    |                       | 14             |
| Calabrese          |                       | 14             |
| Casertana          |                       | 8              |
| Cinta Senese       |                       | 14             |
| Iberian            |                       | 38             |
| Manchado de Jabugo |                       | 5              |
| Mangalitsa         |                       | 47             |
| Mora Romagnola     |                       | 9              |
| Nera Siciliana     |                       | 13             |
| Sicilian           |                       | 2              |
| Turopolje          |                       | 16             |
| Yucatan            | 9                     |                |
| Wild Boar          |                       | Subtotal = 382 |
|                    | Bulgaria              | 5              |
|                    | Croatia               | 15             |

(Continues)

**TABLE 1** (Continued)

| Reference group | Breed/population | Sample size |
|-----------------|------------------|-------------|
|                 | Finland          | 3           |
|                 | France           | 24          |
|                 | Greece           | 57          |
|                 | Italy            | 15          |
|                 | Luxembourg       | 4           |
|                 | Netherlands      | 56          |
|                 | Poland           | 4           |
|                 | Portugal         | 11          |
|                 | Romania          | 16          |
|                 | Russia           | 18          |
|                 | Sardinia         | 92          |
|                 | Slovenia         | 18          |
|                 | Spain            | 37          |
|                 | Tunisia          | 7           |

Note: Values in italics are totals of each reference group.

groups, we recorded the observed allele frequencies ( $p$ ) for all five reference groups, which we held constant in the parameterization of the statistical test described below.

## Statistical methods

Given that a history of wild boar hybridization, shared among the vast majority of invasive wild pigs within the contiguous United States, was the common characteristic we sought to leverage in the development of a genetic test to differentiate invasive wild pigs from domestic pigs, we configured a test statistic to reflect the change in the likelihood of a given genotype if that genotype was drawn strictly from the allele frequencies represented among a combination of all domestic pig reference groups ( $L$  reference groups = mixed-commercial, Duroc, heritage, and primitive) versus allowing for wild boar ancestry in full (genetically pure wild boar) or in part (wild boar  $\times$  domestic pig hybrid;  $K$  reference groups representing wild boar and four domestic pig reference groups). Specifically, the test statistic, hereafter referred to as delta likelihood, represents the difference in the likelihood of the ancestry vector  $\hat{q}^K$  estimated for a given genotype if drawn from any combination of the allele frequencies ( $p$ ) represented among the  $K=5$  reference groups minus the likelihood of the ancestry vector  $\hat{q}^L$  estimated strictly from allele frequencies among the  $L=4$  domestic pig reference groups. Accordingly, if  $p_{km}$  is the frequency of the reference allele at locus  $m$  in the

$k$ th reference group and  $X_m = \{0, 1, 2\}$  is the number of reference alleles carried by a given genotype, the delta likelihood statistic for a given genotype  $X$  can be calculated across all loci  $M$  as follows:

$$\Delta \ell(p|X) := \frac{1}{2M} \sum_{m=1}^M X_m \log \left( \frac{\langle \hat{q}^K, p_{\cdot m} \rangle}{\langle \hat{q}^L, p_{\cdot m} \rangle} \right) + (2 - X_m) \log \left( \frac{1 - \langle \hat{q}^K, p_{\cdot m} \rangle}{1 - \langle \hat{q}^L, p_{\cdot m} \rangle} \right). \quad (1)$$

Here,  $\langle q, p_{\cdot m} \rangle := \sum_{k=1}^K q_k p_{km}$  is the probability of possessing the reference allele for an individual with admixture  $q$  at locus  $m$ . Note that we assume that all loci are biallelic. A thorough development of the calculation of the likelihood and methods to compute  $\hat{q}^K$  and  $\hat{q}^L$  is provided in Pfaffelhuber and Rohde (2022).

With high-density SNP genotypes, we expect linkage disequilibrium (LD) in which there is a nonrandom association of alleles among proximate loci. Model-based approaches that describe the association of genotypes with respective genetic clusters generally assume linkage equilibrium, necessitating the thinning of loci in close linkage (Alexander et al., 2009; Wringe et al., 2019). However, the description of LD within a dataset can be spuriously associated with correlations in allele frequencies attributable to genetic structure (Falush et al., 2003). Accordingly, we evaluated the influence of LD on the delta likelihood statistic by implementing a novel three-step procedure (Appendix S1: Figure S1). First, we identified the relative informativeness of individual loci for differentiating among the specified reference groups. Typically, LD pruning algorithms identify marker dyads that exceed the specified linkage threshold (e.g.,  $R^2 \geq 0.5$ ) and then retain the locus with the higher minor allele frequency; however, minor allele frequency may not have a bearing on the power of linked loci to discriminate among reference groups. Accordingly, for marker dyads that exceeded our LD threshold, we retained the locus that was more informative (higher  $F_{ST}$ ) for differentiating among the reference groups based on the rankings generated with assignPOP (Chen et al., 2018). Second, to address the influence of genetic structure in estimating LD, we restricted the inference of LD to a single focal reference group (e.g., mixed-commercial), pruning a single locus from linked marker dyads in accord with step one. We then pruned loci among the remaining four reference groups (e.g., Duroc, heritage, primitive, and European wild boar) to align with the loci retained from the single reference group and repeated this process

iteratively—treating each of the five reference groups as the focal group. Third, we evaluated the influence of the marker sets retained with focal group-based LD pruning on the distribution of delta likelihood statistic by seeking to maximize the difference between the delta likelihood statistics calculated for domestic pig and wild boar reference groups under the assumption that maximizing this difference would similarly provide the greatest statistical power to differentiate animals that possessed wild boar ancestry from domestic pigs.

To detail the implementation of the stepwise LD pruning procedure, we first ranked the informativeness of all loci in differentiating among the five reference groups using the R package assignPOP (function “assign.MC”; Chen et al., 2018; R Core Team, 2023; Appendix S1: Figure S1). Specifically, we conducted 1000 Monte Carlo bootstrap iterations, retaining 90% of reference genotypes as training data to rank loci by  $F_{ST}$  within each iteration and concatenating all iterations to rank the informativeness of loci. We then quantified linkage among genomically proximate loci using PLINK version 1.9 (Purcell et al., 2007; option “-r2” to generate linkage report for loci with  $R^2 \geq 0.5$  using a default window size of 10 loci and step size of 1). For marker dyads that exceeded our LD threshold of  $R^2 \geq 0.5$ , we retained the locus that was more informative for differentiating among the reference groups based on the rankings generated with assignPOP. With linked loci pruned from a given focal reference group (e.g., mixed-commercial), we then used PLINK (option “-snp-extract”) to prune the loci of the remaining reference groups (e.g., Duroc, heritage, primitive, and wild boar) to the same set of loci retained for the focal group. We then proceeded to calculate the delta likelihood statistic for all 1421 reference genotypes, plotting the distribution of the statistic by reference group. To identify the influence of the focal reference group used for LD pruning on the distribution of the respective delta likelihood statistics, we calculated the difference between the upper tail of the kernel density estimates (base R function “density” using the default parameters; R Core Team, 2023) observed among the domestic pig reference groups versus the lower tail of the wild boar reference group (Appendix S1: Figure S4). We then repeated the LD pruning routine and calculation of the delta likelihood statistic iteratively, treating each of the five reference groups as the focal population and retaining the focal group-based LD pruned marker set that maximized the difference between domestic pig and wild boar reference groups.

With the reference set drawn largely from previously published datasets with unbalanced sample sizes among breeds or populations, we implemented a bootstrapping routine executed with a custom R script to quantify the

variance in individual delta likelihood statistics attributable to the unique assembly of genotypes in the reference set and associated uncertainty in the differentiation of animals possessing wild boar ancestry from domestic pigs (Appendix S1: Figure S1). Specifically, holding sample sizes constant for each of the respective reference groups, we resampled genotypes from within each reference group with replacement while allowing the representation of included breeds or populations to vary with each iteration. We updated allele frequencies for each of the reference groups and then calculated individual delta likelihood statistics for each of the 1421 reference genotypes (regardless of whether the individual reference sample was included in the bootstrapped iteration), repeating this process 1000 times.

We then sought to fit a statistical distribution to the calculated delta likelihood statistics to describe the associated probability that a given genotype was derived strictly from allele frequencies represented among domestic pig reference groups as opposed to representing ancestry contributions from wild boar (Appendix S1: Figure S1). Specifically, we fit statistical distributions to the bootstrapped values of the test statistic for (1) an aggregate of all domestic pig genotypes and (2) each of the four domestic reference groups individually, selecting the bootstrapped test statistics over the strictly empirical data to overcome the limitations imposed by the opportunistic assembly of the reference set as described above. Calculation of the delta likelihood statistic can result in negative values when the likelihood that a genotype is drawn from a mix of both wild boar and domestic pig allele frequencies is lower than the likelihood if the same genotype was drawn strictly from domestic pig allele frequencies; genotypes producing negative delta likelihood values were interpreted as providing no evidence that the genotype in question possessed wild boar ancestry. Accordingly, we truncated negative delta likelihood values calculated for the reference set and fit statistical distributions only to positive values, as including negative values would limit some of the statistical distributions available for evaluation. Specifically, we fit normal, Weibull, lognormal, and gamma distributions to the bootstrapped values calculated for the domestic pig reference groups, using maximum likelihood estimation in the R package *fitdistrplus* version 1.1-8 (function “*fitdist*” with default parameters; Delignette-Muller & Dutang, 2015; R Core Team, 2023). Given that we truncated delta likelihood values at zero, we similarly needed to truncate the fitted statistical distributions to positive values using the R package *truncdist* version 1.0-2 (functions “*dtrunc*,” “*ptrunc*,” and “*qtrunc*”; Novomestky & Nadarajah, 2016). Finally, we evaluated the fit of these alternative statistical distributions with Kolmogorov–Smirnov statistic (KS), Anderson–Darling statistic (AD), Akaike

information criterion (AIC), and the Bayesian information criterion (BIC) calculated using the *fitdistrplus* R package (function “*gofstat*,” Delignette-Muller & Dutang, 2015).

We used two disparate methods to evaluate the sensitivity and type I error rate associated with using the delta likelihood statistic to identify genotypes that possessed wild boar ancestry—testing with both empirical data and simulated genotypes. First, the breeds included in the reference set were selected to represent the breadth of diversity observed among domestic lineages that could have potentially contributed to invasive wild pig populations in the United States, yet only represent a subset of the full diversity observed among domestic pigs. Given the objective of restricting the probability of falsely classifying a domestic pig as an animal possessing wild boar ancestry (type I error), it was necessary to evaluate whether the exclusion of a given breed from the reference set would result in the breed being misclassified as possessing wild boar ancestry. Accordingly, we calculated delta likelihood values for an additional 29 breeds of domestic pig represented by 435 unique genotypes (Appendix S1: Table S1), which were excluded from the reference set to specifically evaluate whether delta likelihood values for domestic breeds excluded from the reference groups would fall within the distribution observed for other domestic pigs or whether test statistic values might be inflated due to the breed’s exclusion from the reference set.

As a second, independent test of the discriminatory power of the delta likelihood approach, we used the pedigree simulator available in R package *gscramble* (Anderson, 2023) to simulate genotypes that included known proportions of wild boar ancestry. Specifically, we populated the pedigree by randomly sampling without replacement a total of 16 individuals from the reference set, sequentially selecting, with each iteration, 16–1 genotypes from one of the four domestic pig reference groups and a complementary number (0–15) of wild boar genotypes (e.g., pedigrees with 16, 15, ..., 1 genotypes drawn from the mixed-commercial reference group with 0, 1, ..., 15 complementary genotypes drawn from the wild boar reference group). Within the specified pedigree, we then allowed for four generations of one-to-one pairing with no inbreeding while allowing for biologically realistic rates of recombination (1 cM/Mb) to reflect the linkage structure among proximate loci and tracking the proportion of each genotype attributable to wild boar ancestry versus domestic pig ancestry. We then randomly sampled a single simulated genotype from each pedigree with a known proportion of wild boar ancestry. We repeated this process 1000 times for all 16 combinations of founders within the specified pedigree (16–1 domestic pig reference genotypes with 0–15 wild boar reference genotypes) for each of the four domestic pig reference groups

(mixed-commercial, Duroc, heritage, and primitive; 16 pedigree combinations  $\times$  4 domestic pig reference groups  $\times$  1000 simulations = 64,000 simulated genotypes). We then calculated the delta likelihood statistic for each simulated genotype and used regression analysis to evaluate the relationship between the known proportion of wild boar ancestry as determined within the described pedigree and the associated test statistic.

## Evaluation of invasive wild pigs

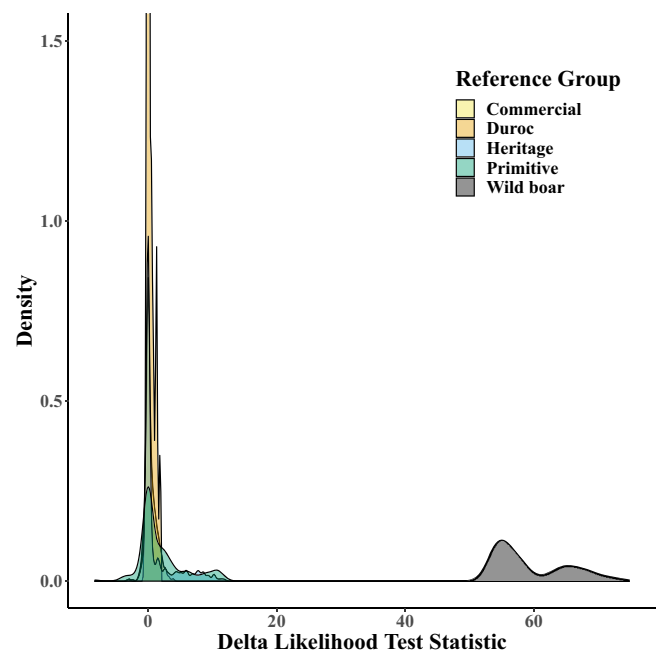
To evaluate the discriminatory power of the delta likelihood approach for *S. scrofa* genotypes of unknown origin, we applied this method to 6566 invasive wild pigs sampled across the full extent of the invaded range within the contiguous United States. These genotypes were published previously with a detailed description of data collection (Smyser et al., 2020). Briefly, genetic samples were collected from invasive wild pigs that overwhelmingly were lethally removed as a component of disease surveillance or damage mitigation efforts conducted by the US Department of Agriculture (USDA)—Animal Plant Health Inspection Services (APHIS)—Wildlife Services (WS). DNA was extracted from hair or tissue with commercially available magnetic bead recovery kits (MagMax DNA, Thermo Fisher Scientific, Waltham, Massachusetts) and genotyped with GeneSeek's Genomic Profiler for Porcine HD. Genotypes with call rates  $<95\%$  were excluded from the published dataset. Delta likelihood statistics for all wild pig genotypes were calculated as described above. As with the simulated genotypes, we were interested in comparing delta likelihood statistics to previous characterizations of ancestry for these animals (Smyser et al., 2020). Accordingly, we conducted a regression analysis to evaluate the association between the delta likelihood values for individual invasive wild pigs and associated European wild boar ancestry as calculated in Smyser et al. (2020; sum of Sardinian wild boar [K4] and European mainland wild boar [K17]). To further illustrate the discriminatory power of the delta likelihood approach versus ancestry methods, we similarly evaluated the statistical relationship between estimates of European wild boar ancestry and delta likelihood values for the domestic pig and wild boar reference samples. Finally, to characterize spatial patterns across the invaded range, we mapped delta likelihood values calculated for invasive wild pigs, summarized at the county level.

## RESULTS

With the stepwise LD pruning approach, treating the heritage reference group as the focal population maximized

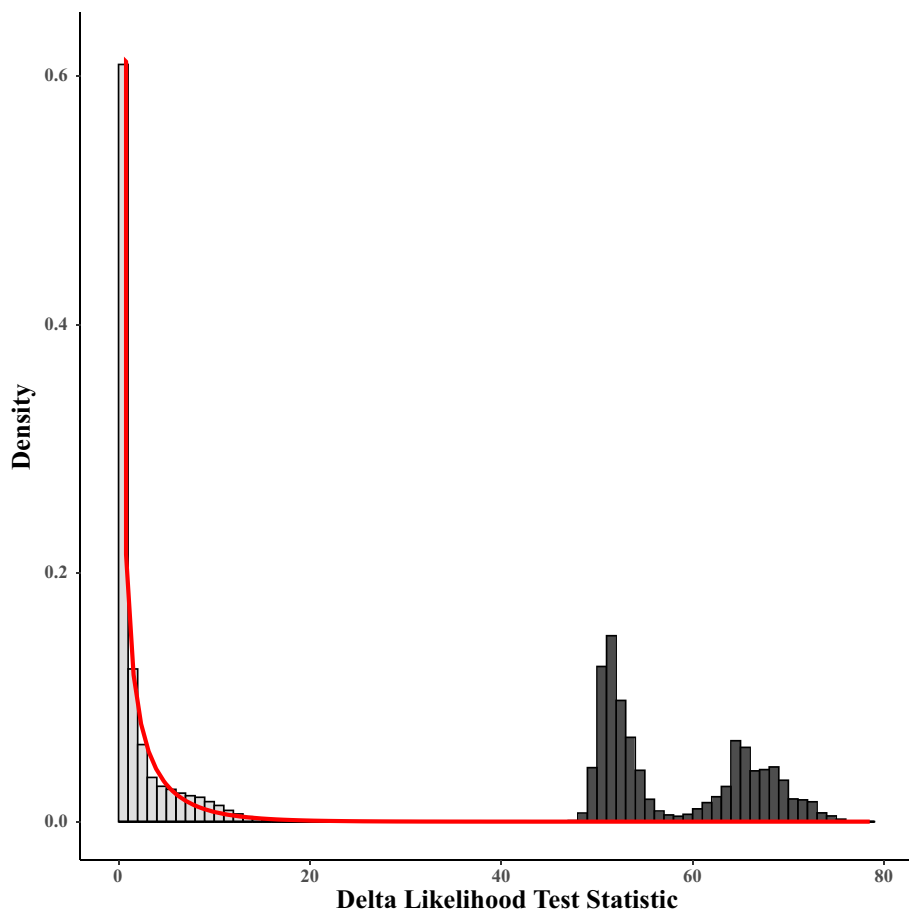
the difference between the kernel density estimates of the delta likelihood statistic observed for all domestic pig reference groups and wild boar (Figure 1; Appendix S1: Figure S4). Accordingly, as informed by LD pruning for the heritage reference group, 18,790 of the 28,545 candidate loci were retained for all subsequent analyses. Delta likelihood values for the 1039 domestic pigs evaluated among the four corresponding reference groups ranged from  $-8.25$  to  $12.08$ , whereas associated values for the 382 wild boar ranged from  $53.26$  to  $74.79$ . As anticipated, evaluation of delta likelihood values for the reference set with 1000 bootstrap iterations broadened the distribution, with values for domestic pig genotypes ranging from  $-11.13$  to  $19.66$  and wild boar genotypes ranging from  $39.79$  to  $78.51$ .

The gamma distribution best fits the bootstrapped delta likelihood values for the aggregate of all domestic pig genotypes as characterized by KS, AD, AIC, and BIC (Figure 2, Table 2). Among the individual domestic pig reference groups, the gamma distribution was the best fit for the mixed-commercial, heritage, and primitive groups; however, the lognormal distribution was the best fit for the Duroc group (Appendix S1: Table S2). Accordingly, we used the gamma distribution to describe



**FIGURE 1** Kernel density estimates of delta likelihood statistics calculated from high-density single nucleotide polymorphism genotypes for *Sus scrofa* reference genotypes, organized into reference groups representing mixed-commercial breeds ( $n = 310$ ), Durocs ( $n = 159$ ), heritage breeds ( $n = 381$ ), primitive breeds ( $n = 189$ ), and European wild boar ( $n = 382$ ) with 18,790 loci retained for analysis reflecting linkage disequilibrium pruning based on the heritage breed reference group.





**FIGURE 2** A gamma distribution (red line) fitted to bootstrapped delta likelihood statistics for 1039 domestic pigs reference genotypes representing 33 described domestic breeds, with comparable values for 382 European wild boar (dark gray bars) illustrated for reference.

**TABLE 2** Kolmogorov–Smirnov statistic, Anderson–Darling statistic, Akaike information criterion (AIC; expressed as  $\Delta$ AIC relative to the best model), and the Bayesian information criterion ( $\Delta$ BIC) were used to evaluate the fit of the normal, Weibull, lognormal, and gamma statistical distributions to delta likelihood statistics calculated for 1039 domestic pigs (*Sus scrofa*) reference genotypes (representing 33 described domestic breeds and assumed to be free of wild boar introgression) as a means of characterizing the null distribution of the delta likelihood statistic as a statistical test of wild boar ancestry among *S. scrofa* of unknown origin.

| Goodness-of-fit criteria | Normal    | Weibull | Lognormal | Gamma  |
|--------------------------|-----------|---------|-----------|--------|
| Kolmogorov–Smirnov       | 0.3987    | 0.0637  | 0.0932    | 0.0585 |
| Anderson–Darling         | 492,648   | 5780    | 12,524    | 4777   |
| $\Delta$ AIC             | 1,421,538 | 12,423  | 95,170    | 0      |
| $\Delta$ BIC             | 1,421,538 | 12,423  | 95,170    | 0      |

the probability ( $p$  value) that a given genotype was derived strictly from allele frequencies observed among the domestic pig reference groups.

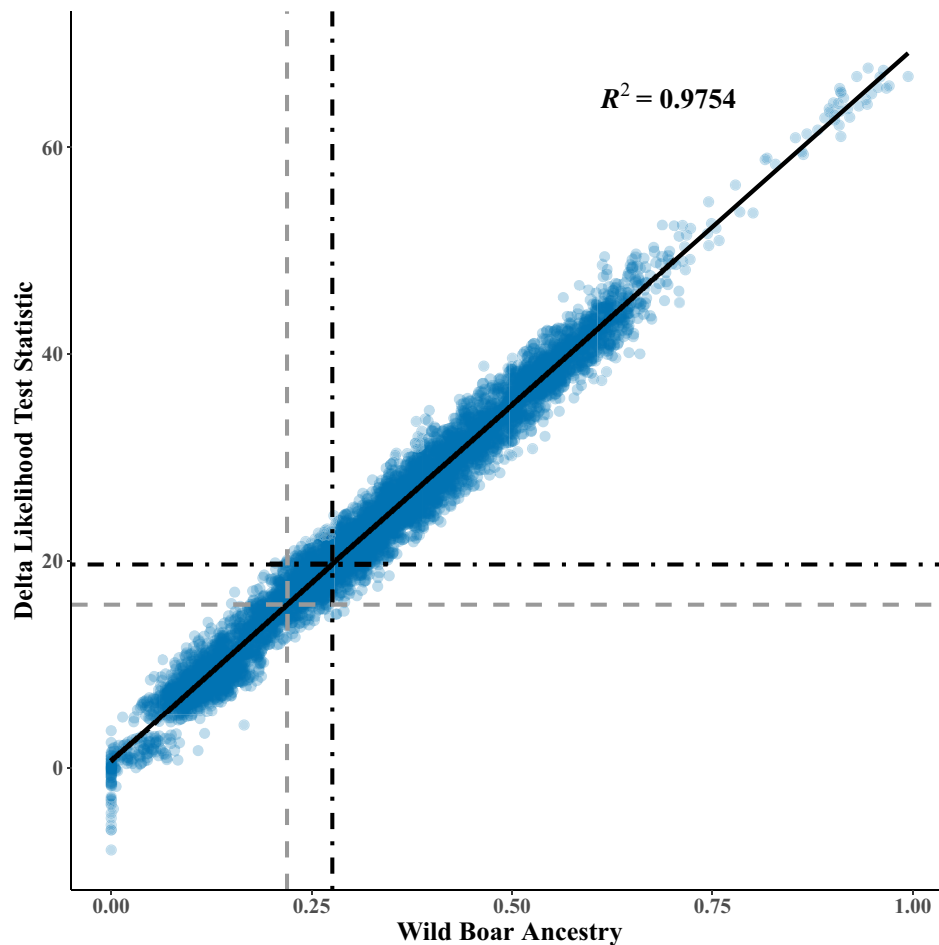
In our empirical test of the delta likelihood approach, evaluating whether breeds excluded from the domestic pig reference set would be misidentified as possessing wild boar ancestry, all 435 genotypes representing an additional 29 breeds fell within the distribution characterized from the domestic pig reference genotypes (range =  $-1.43$  to  $7.51$ ; associated  $p = 1.00$ – $0.06$ ; Appendix S1: Figure S5).

Evaluating the delta likelihood approach with simulated genotypes, in which genotypes were drawn from pedigrees that combined 16–1 genotypes sampled from one of the four domestic pig reference groups with 0–15 corresponding genotypes sampled from the wild boar reference group, allowed us to evaluate both the type I error rate and sensitivity to detect wild boar ancestry. The distribution of delta likelihood statistics for the genotypes simulated without wild boar hybridization generally aligned with distributions of the empirical

data from the associated reference group, providing validation for the inference drawn from simulated genotypes. Of the 4000 genotypes simulated from pedigrees populated strictly with genotypes from the four domestic pig reference groups (16 domestic pig genotypes/0 wild boar genotypes; implicitly simulating 1000 pure domestic pig genotypes from each of the domestic pig reference groups), the maximum delta likelihood test statistic value was 5.72 (drawn from a pedigree representing primitive breeds), which is within the distribution expected for domestic pigs ( $p$  value = 0.10). Delta likelihood values calculated for simulated domestic pig-wild boar hybrid genotypes were strongly correlated with the proportion of genotypes attributable to wild boar ancestry ( $R^2 = 0.9834$ ,  $p < 2.2e-16$ ; Appendix S1: Figure S6). In regard to the sensitivity in detecting wild boar ancestry, characterizing the distribution of delta likelihood values for domestic pigs with a

fitted gamma distribution enables the identification of a threshold value from which to statistically classify simulated genotypes as possessing wild boar ancestry. For example, the threshold value drawn from the gamma distribution of  $p = 0.01$  (delta likelihood = 15.77) was associated with a wild boar ancestry of 20.06% as inferred from the regression of wild boar ancestry, known for simulated genotypes, versus corresponding delta likelihood values.

Among invasive wild pigs, delta likelihood values calculated for 6566 wild pigs sampled throughout the contiguous United States ranged from  $-7.95$  to  $67.67$ , with the breadth in values reflective of the genetic heterogeneity observed both among and within populations. Delta likelihood values calculated for invasive wild pig genotypes were strongly correlated with European wild boar ancestry as estimated in Smyser et al. (2020;  $R^2 = 0.9754$ ,  $p < 2.2e-16$ ; Figure 3), similar to associations observed



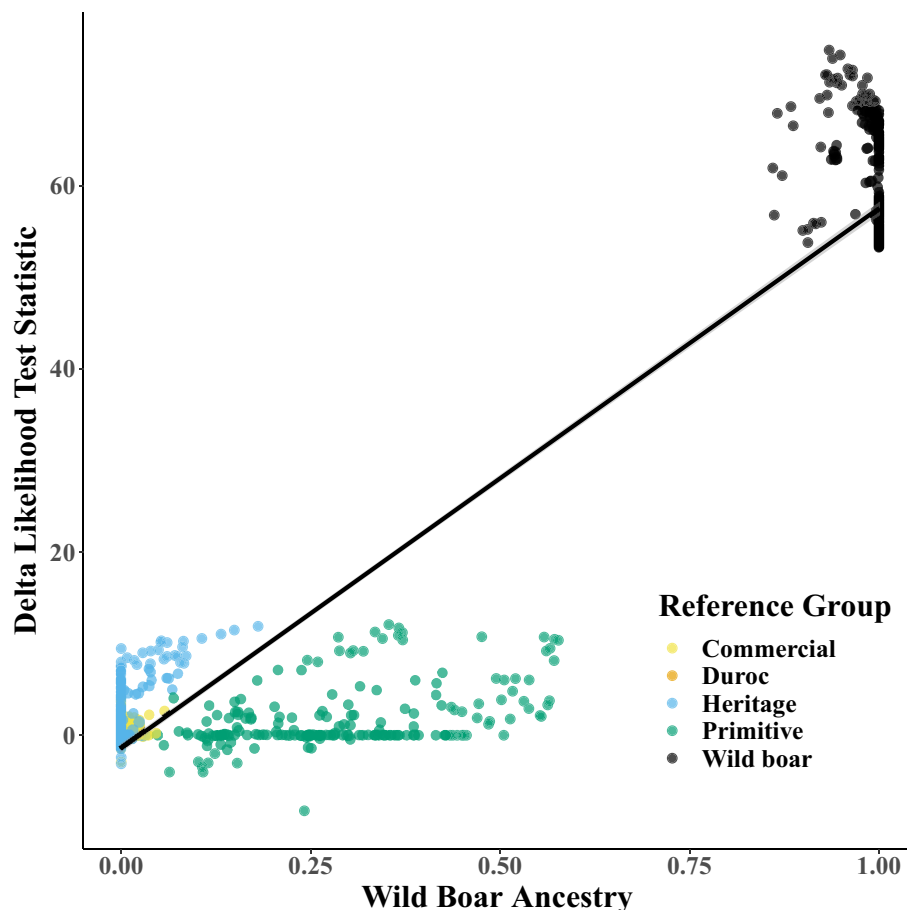
**FIGURE 3** Delta likelihood statistics calculated from 18,790 biallelic single nucleotide polymorphic loci versus European wild boar ancestry as estimated in Smyser et al. (2020) for 6566 wild pigs sampled throughout the invaded range within the contiguous United States, with the maximum bootstrapped value from domestic pig reference genotypes (delta likelihood = 19.66; corresponding to 27.58% European wild boar ancestry) presented as dash-dotted black lines and a value drawn from the fitted gamma distribution associated with  $p = 0.01$  (delta likelihood = 15.77; corresponding to 21.93% European wild boar ancestry) presented as dashed grey lines.

with simulated genotypes as described above. Of the wild pig genotypes evaluated, 80.11% (5260/6566) had delta likelihood values greater than the maximum observed values for all reference domestic pig genotypes (delta likelihood >12.08). Using a threshold value of  $p = 0.01$  drawn from the gamma distribution fitted to the bootstrapped values from reference domestic pig genotypes (corresponding delta likelihood statistic of 15.77) would enable 73.77% (4844/6566) of wild pig genotypes to be classified as possessing wild boar ancestry. Further, this delta likelihood threshold value of 15.77 corresponded to an estimated European wild boar ancestry of 21.93% as inferred from the association between wild boar ancestry and corresponding delta likelihood values calculated for invasive wild pigs. Thus, we can infer the relative sensitivity of the delta likelihood approach with the capacity to statistically differentiate a hybrid from a domestic pig at a statistical standard of  $p = 0.01$  if a given animal possesses approximately 22% wild boar ancestry. Providing a greater level of statistical

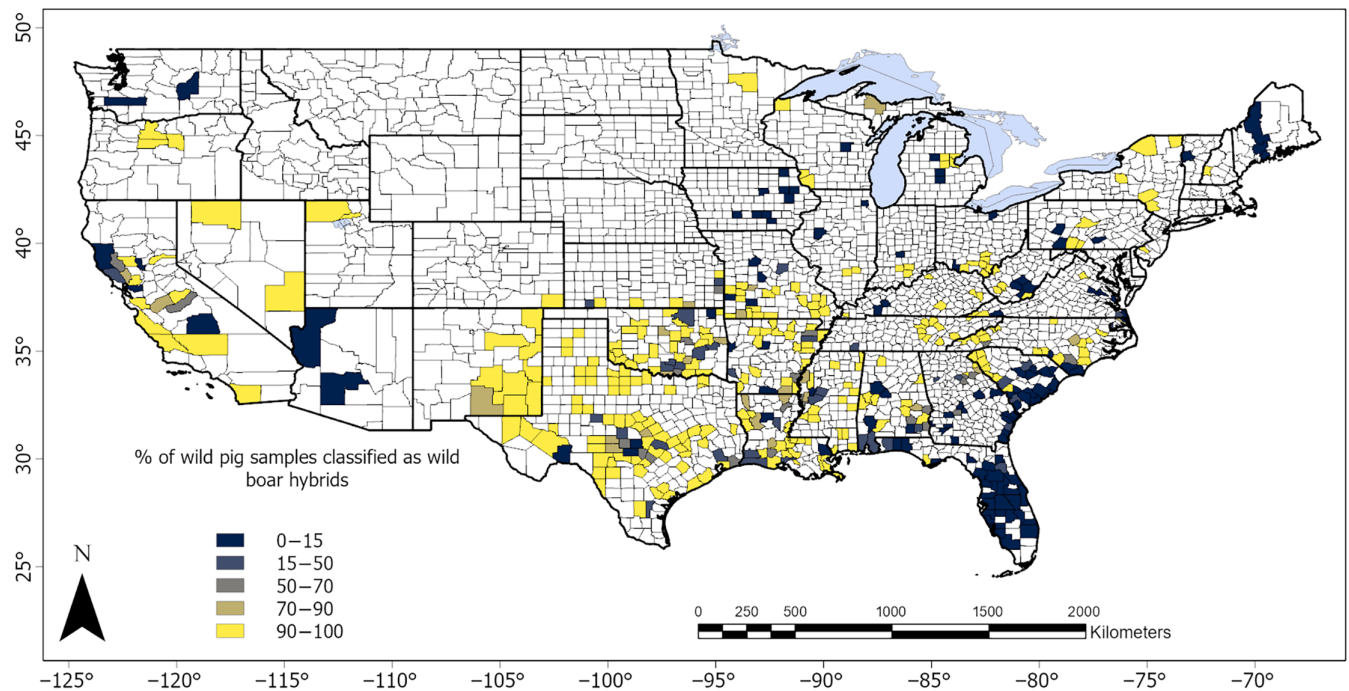
certainty for classification as possessing wild boar ancestry, 63.43% (4165/6566) of wild pig genotypes had delta likelihood values that exceeded the maximum bootstrapped value calculated for all domestic pig reference genotypes (delta likelihood = 19.66,  $p = 0.004$ ), corresponding to an estimated European wild boar ancestry of 27.58%.

Extending similar regression analyses to reference samples demonstrated that the delta likelihood approach more effectively differentiates domestic pigs from wild boar (Figure 4). Specifically, some domestic pig reference genotypes, particularly from the primitive reference group, were characterized as having substantial European wild boar ancestry (as estimated by Smyser et al., 2020) but low delta likelihood values.

Investigating the spatial associations of delta likelihood values across the invaded range reveals regional differences among wild pig populations (Figure 5). Wild pig populations with delta likelihood values that fell within the distribution of domestic pig reference samples (delta likelihood <19.66) were largely restricted to the



**FIGURE 4** Delta likelihood statistics calculated from 18,790 biallelic single nucleotide polymorphic loci versus European wild boar ancestry as estimated in Smyser et al. (2020) for 1039 domestic pig reference genotypes (representing 33 described breeds) and 382 European wild boar (representing 16 populations) demonstrating greater discriminatory power between domestic pig and European wild boar with the delta likelihood approach as compared with ancestry efforts described in Smyser et al. (2020).



**FIGURE 5** Frequency (summarized by county) in which delta likelihood values calculated for 6566 invasive wild pigs, sampled throughout the invaded range within the contiguous United States, exceeded the bootstrapped distribution for 1039 domestic pigs (representing 33 breeds). Invasive wild pigs with test statistic values greater than those drawn from the distribution of domestic pigs can be interpreted as possessing European wild boar ancestry.

Florida peninsula and adjacent areas in the southeastern portions of the invaded range. Other areas characterized by low delta likelihood values were associated with localities that are generally not considered to support self-sustaining invasive wild pig populations (e.g., Iowa, Minnesota, North Dakota, Washington), with free-ranging animals putatively attributable to incidental escape from pork production operations or Vietnamese potbellied pigs (Smyser et al., 2020). Conversely, the identification of wild boar hybrids among other states generally recognized as not supporting self-sustaining invasive populations (e.g., Wisconsin, New York) demonstrates the risks of ongoing introductions, with animals most likely translocated from established invasive populations or representing the release of captive-reared wild boar hybrids.

## DISCUSSION

As anthropogenic processes have driven the rapid expansion of invasive wild pigs throughout much of the United States since the 1980s (Hernández et al., 2018; Smyser et al., 2020; Tabak et al., 2017), additional analytical resources are needed to enforce prohibitions on the possession and transport of these destructive animals.

With the application of the delta likelihood approach, we have demonstrated the capacity to statistically differentiate genetically typical invasive wild pigs—those animals that possess wild boar ancestry—from domestic pigs. The scientifically grounded capacity to differentiate invasive wild pigs from domestic pigs will serve to support the enforcement of prohibitions on the possession or transport of wild pigs. The increasing availability and application of high-resolution molecular tools have provided a foundation to begin addressing the complex and computationally intensive challenges of differentiating genetically similar groups and testing for hybridization (Ramos et al., 2009). By presenting hybrid classification within a context of statistical uncertainty (associated  $p$  value), the delta likelihood approach provides transparency as to whether the weight of the genetic evidence falls in violation of local statutes as interpreted by the criminal justice system. Relative to traditional clustering approaches, the delta likelihood approach improved the resolution and statistical certainty in which wild boar hybridization could be identified (Figures 3 and 4; Goedbloed et al., 2013a, 2013b; Smyser et al., 2020). Specifically, the genetic similarity between primitive breeds and wild boar limited the capacity in which these groups could be resolved with traditional applications of genetic clustering methods (Smyser et al., 2020). By explicitly considering the allele

frequencies that characterize primitive breeds, the application of the delta likelihood approach proved to be more effective for differentiating this group from genetically similar wild boar. Further, by evaluating an individual genotype within the context of a testable hypothesis, we were able to derive an associated probability ( $p$  value) that a given genotype was attributable strictly to the allele frequencies observed among domestic pigs.

In applying the delta likelihood approach to a comprehensive sample of invasive wild pigs, 63% (4165/6566) of sampled genotypes could be confidently classified as possessing wild boar ancestry with test statistic values more extreme than any of the 1039 domestic pig reference genotypes evaluated with 1000 bootstrap iterations (delta likelihood  $>19.66$ ,  $p < 0.004$ ). Invasive wild pigs in the contiguous United States are genetically heterogeneous, with ongoing introduction pressure associated with accidental or deliberate releases from both livestock production and the pet trade (i.e., potbellied pigs; Smyser et al., 2020). Thus, we would not expect any genetic approach to have the capacity to differentiate all invasive wild pigs from domestic sources. Similarly, in developing an approach for differentiating invasive wild pigs from domestic pigs, it was imperative that domestic pigs were not falsely identified as possessing wild boar ancestry (type I error). Evaluating this approach with a robust sample of domestic breeds withheld from the reference set provided validation of the delta likelihood approach as none of the 435 representative genotypes were misidentified as wild boar hybrids. Similar results were found for 4000 simulated pure domestic pig genotypes (pedigrees populated with 16 domestic pig genotypes/0 wild boar genotypes), with all simulated genotypes falling within the delta likelihood distribution characterized by domestic pig reference genotypes ( $p \geq 0.10$ ). Additionally, we reduced the risk of type I error by deriving associated  $p$  values for the delta likelihood statistic from the more conservative bootstrapped distribution from the domestic pig reference genotypes. A consequence of minimizing type I error is that ancestry estimates suggest a subset of sampled invasive wild pigs possessed wild boar ancestry, but these genotypes could not be confidently differentiated from domestic pigs with the delta likelihood approach (type II error; Figure 3). By expressing classification within a probability ( $p$  value) context, a benefit of the delta likelihood is that it enables weighing type I (i.e., the risk of falsely identifying a domestic pig as a wild boar hybrid) versus type II (failing to correctly identify a hybrid animal as such) within a given context. For example, enforcement of legal prohibitions on the possession of invasive wild pigs may choose to implement a more stringent statistical threshold, to minimize type I error, than managers monitoring wild populations

for evidence of gene flow or domestic introgression, in which type II error may be of greater concern (Goedbloed et al., 2013a; Iacolina et al., 2018; Wells et al., 2019).

By exploiting wild boar ancestry as a common characteristic—but not necessarily a defining attribute—of genetically heterogeneous invasive wild pigs in the United States (and likely Canada; Koen et al., 2021), the delta likelihood approach was useful across large parts of the United States. However, this approach was inadequate for differentiating wild pigs in specific regions (i.e., populations in Florida and southern Georgia that are not characterized by a history of wild boar hybridization; Figure 5; Smyser et al., 2020). Disparate selective pressures incurred by domestic pigs and wild pigs may lead to increased divergence of specific genomic regions (Barmantlo, 2021; Petrelli et al., 2022). Thus, as an alternative approach, high-resolution molecular tools targeting a limited number of genomic regions under intense but divergent selective pressure could provide greater sensitivity in differentiating wild pigs from domestic pigs across the full extent of the invaded range. However, more work is needed to evaluate whether a divergent evolutionary signal is shared among heterogeneous wild pigs, which vary in ancestral origins and have successfully established self-sustaining populations in (and thus may be differentially adapting to) a diversity of ecosystems.

Beyond the specific application to invasive wild pigs, the delta likelihood approach represents a computationally efficient and generalizable method for the probabilistic classification of individuals as either founders or hybrids with broad ecological and economic applications. For example, by reoptimizing the reference set, this approach could be applied to the pork industry to verify breed labeling, with application to markets in which products from specific breeds command a premium price (Lin et al., 2014; Szántó-Egész et al., 2016; Szemethy et al., 2021; Wilkinson et al., 2012). Similarly, additional development of the analyses presented here could be applied to address conservation concerns of domestic pig introgression among wild boar populations in their native European range (Fulgione et al., 2016; Goedbloed et al., 2013a, 2013b; Iacolina et al., 2018; Petrelli et al., 2022; Schleimer et al., 2022). Beyond applications to pigs or their wild progenitor, domestic introgression into wild populations poses a conservation concern for other species. For example, the delta likelihood approach could be used to inform bison (*Bison bison*) herd management decisions reflecting a goal of reducing the genetic influence of past cattle introgression (Hartway et al., 2020) or classify wild Atlantic salmon (*Salmo salar*) with a history of introgression from cultured populations

(Wringe et al., 2019). More broadly, the delta likelihood approach can be generalized to serve as a method to quantify directional gene flow by statistically identifying individuals of mixed ancestry when genetic populations of interest can be delineated. As a generalizable test, the delta likelihood approach could be implemented across study systems with any number of genetic markers, recognizing that the statistical power of the test will be influenced by the true genetic differences between reference groups and the ability to resolve those differences with the available marker set (i.e., could be applied to low-density data as with nonmodel species or genotypes produced from trace evidence as in forensic applications). Thus, the delta likelihood approach could be applied to numerous other ecological systems regardless of whether gene flow may be cause for conservation concern (e.g., koloa maoli or Hawaiian duck [*Anas wyvilliana*], Wells et al., 2019) or whether the detection of natural gene flow might obviate the need to conduct translocations for the purposes of genetic rescue (arctic fox [*Vulpes lagopus*], Hasselgren et al., 2018).

In sum, the development of the delta likelihood approach and application to the challenges posed by invasive wild pigs provides a needed tool for natural resources managers, delivering statistical support for the enforcement of prohibitions on the possession and movement of this destructive species. As a component of a comprehensive management plan, the increased capacity to enforce such prohibitions should serve as a deterrent for the human-facilitated spread of invasive wild pigs. Recognizing that the rapid expansion of this invasive species observed since the 1980s has been driven by the deliberate introduction of animals into uninvaded habitats as opposed to incremental biological processes of natal dispersal from established populations, the capacity to curtail illicit translocation is an essential component of reducing ecological destruction and economic damage caused by invasive wild pigs.

## ACKNOWLEDGMENTS

The research was supported by the USDA-APHIS-WS-National Feral Swine Damage Management Program and the USDA-APHIS-WS-National Wildlife Disease Program. We are grateful for the numerous Wildlife Services employees and partners from other agencies that made sampling wild pigs throughout the invaded range within the contiguous US possible. We thank Eric C. Anderson for guidance in implementing simulations in gscramble. Contributions to this work from Valentin Adrian Bâlțeanu were supported by research project TE 113/2022—Research projects for stimulating young independent teams, PNIII Subprogram 1.1—Human Resources, which is currently

funded by the Ministry of Research, Innovation and Digitization of Romania. We appreciate assistance from Zachary Moen with the development of figures and comments on the manuscript provided by Daniel M. McDonald.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

## DATA AVAILABILITY STATEMENT

Genotypic data and code (Smyser et al., 2024) are available from Dryad: <https://doi.org/10.5061/dryad.tjq2bw48>.

## ORCID

Amy J. Davis  <https://orcid.org/0000-0002-4962-9753>

## REFERENCES

- Alexander, D. H., J. Novembre, and K. Lange. 2009. "Fast Model-Based Estimation of Ancestry in Unrelated Individuals." *Genome Research* 19: 1655–64.
- Anderson, A., C. Sloatmaker, E. Harper, J. Holderieath, and S. A. Shwiff. 2016. "Economic Estimates of Feral Swine Damage and Control in 11 US States." *Crop Protection* 89: 89–94.
- Anderson, A., C. Sloatmaker, E. Harper, R. S. Miller, and S. A. Shwiff. 2019. "Predation and Disease-Related Economic Impacts of Wild Pigs on Livestock Producers in 13 States." *Crop Protection* 121: 121–26.
- Anderson, E. C. 2023. "gscramble." eriqande. <https://github.com/eriqande/gscramble/>.
- Ballari, S. A., and M. N. Barrios-Garcia. 2014. "A Review of Wild Boar *Sus scrofa* Diet and Factors Affecting Food Selection in Native and Introduced Ranges." *Mammal Review* 44: 124–134.
- Barmentlo, N. 2021. "Examining the Spread of Invasive Swine in the US in Context of the Invasive Bridgehead Effect." MSc thesis, Biological Sciences, University of Amsterdam.
- Barrios-Garcia, M. N., and D. Simberloff. 2013. "Linking the Pattern to the Mechanism: How an Introduced Mammal Facilitates Plant Invasions." *Austral Ecology* 38: 884–890.
- Bellard, C., P. Cassey, and T. M. Blackburn. 2016. "Alien Species as a Driver of Recent Extinctions." *Biology Letters* 12: 20150623.
- Bertelsmeier, C., and L. Keller. 2018. "Bridgehead Effects and Role of Adaptive Evolution in Invasive Populations." *Trends in Ecology & Evolution* 33: 527–534.
- Boughton, E. H., and R. K. Boughton. 2014. "Modification by an Invasive Ecosystem Engineer Shifts a Wet Prairie to a Monotypic Stand." *Biological Invasions* 16: 2105–14.
- Bratton, S. P. 1975. "The Effect of the European Wild Boar, *Sus scrofa*, on Gray Beech Forest in the Great Smoky Mountains." *Ecology* 56: 1356–66.
- Brown, V. R., M. C. Marlow, T. Gidlewski, R. Bowen, and A. Bosco-Lauth. 2020. "Perspectives on the Past, Present, and Future of Feral Swine Disease Surveillance in the United States." *Journal of Animal Science* 98: 1–3.
- Brown, V. R., R. S. Miller, S. C. McKee, K. H. Ernst, N. M. Didero, R. M. Maison, M. J. Grady, and S. A. Shwiff. 2021. "Risks of

- Introduction and Economic Consequences Associated with African Swine Fever, Classical Swine Fever and Foot-and-Mouth Disease: A Review of the Literature.” *Transboundary and Emerging Diseases* 68: 1910–65.
- Carlisle, K. M., N. Didero, S. McKee, J. Elser, and S. A. Shwiff. 2021. “Towards a More Comprehensive Understanding of Wild Pig (*Sus scrofa*) Impacts on Agricultural Producers: Insights from a Texas Case Study.” *Crop Protection* 150: 105793.
- Chen, K.-Y., E. A. Marschall, M. G. Sovic, A. C. Fries, H. L. Gibbs, and S. A. Ludsins. 2018. “assignPOP: An R Package for Population Assignment Using Genetic, Non-Genetic, or Integrated Data in a Machine-Learning Framework.” *Methods in Ecology and Evolution* 9: 439–446.
- Chinn, S. M., J. C. Kilgo, M. A. Vukovich, and J. C. Beasley. 2021. “Influence of Intrinsic and Extrinsic Attributes on Neonate Survival in an Invasive Large Mammal.” *Scientific Reports* 11: 11033.
- Comer, C. E., and J. J. Mayer. 2009. “Wild Pig Reproductive Biology.” In *Wild Pigs: Biology, Damage, Control Techniques, and Management*, edited by J. J. Mayer and I. L. Brisbin, Jr., 51–75. Aiken, SC: Savannah River National Laboratory. SRNL-RP-2009-00869.
- Cushman, J. H., T. A. Tierney, and J. M. Hinds. 2004. “Variable Effects of Feral Pig Disturbances on Native and Exotic Plants in a California Grassland.” *Ecological Applications* 14: 1746–56.
- Delignette-Muller, M. L., and C. Dutang. 2015. “fitdistrplus: An R Package for Fitting Distributions.” *Journal of Statistical Software* 64: 1–34.
- Falush, D., M. Stephens, and J. K. Pritchard. 2003. “Inference of Population Structure Using Multilocus Genotype Data: Linked Loci and Correlated Allele Frequencies.” *Genetics* 164: 1567–87.
- Fern, M., R. Barlow, C. Sloomaker, J. Kush, S. Shwiff, L. Teeter, and J. Armstrong. 2021. “Economic Estimates of Wild Hog (*Sus scrofa*) Damage and Control among Young Forest Plantations in Alabama.” *Small-Scale Forestry* 20: 503–516.
- Fulgione, D., D. Ripa, M. Buglione, M. Trapanese, S. Petrelli, and V. Maselli. 2016. “Unexpected but Welcome. Artificially Selected Traits May Increase Fitness in Wild Boar.” *Evolutionary Applications* 9: 769–776.
- Gaillard, J. M., M. Festa-Bianchet, and N. G. Yoccoz. 1998. “Population Dynamics of Large Herbivores: Variable Recruitment with Constant Adult Survival.” *Trends in Ecology & Evolution* 13: 58–63.
- Goedbloed, D. J., H.-J. Megens, P. van Hooft, J. M. Herrero-Medrano, W. Lutz, P. Alexandri, R. P. M. A. Crooijmans, et al. 2013. “Genome-Wide Single Nucleotide Polymorphism Analysis Reveals Recent Genetic Introgression from Domestic Pigs into Northwest European Wild Boar Populations.” *Molecular Ecology* 22: 856–866.
- Goedbloed, D. J., P. van Hooft, H.-J. Megens, K. Langenbeck, W. Lutz, R. P. M. A. Crooijmans, S. E. van Wieren, R. C. Ydenberg, and H. H. T. Prins. 2013. “Reintroductions and Genetic Introgression from Domestic Pigs Have Shaped the Genetic Population Structure of Northwest European Wild Boar.” *BMC Genetics* 14: 43.
- Hardin, G. 1968. “The Tragedy of the Commons.” *Science* 162: 1243–48.
- Hartway, C., A. Hardy, L. Jones, B. Moynahan, K. Traylor-Holzer, B. McCann, K. Aune, and G. Plumb. 2020. *Long-Term Viability of Department of the Interior Bison under Current Management and Potential Metapopulation Management Strategies*. Fort Collins, CO: National Park Service. Natural Resource Report NPS/NRSS/BRD—2020/2097.
- Hasselgren, M., A. Angerbjorn, N. E. Eide, R. Erlandsson, O. Flagstad, A. Landa, J. Wallen, and K. Noren. 2018. “Genetic Rescue in an Inbred Arctic Fox (*Vulpes lagopus*) Population.” *Proceedings of the Royal Society B-Biological Sciences* 285: 20172814.
- Hernández, F. A., B. M. Parker, C. L. Pylant, T. J. Smyser, A. J. Piaggio, S. L. Lance, M. P. Milleson, J. D. Austin, and S. M. Wisely. 2018. “Invasion Ecology of Wild Pigs (*Sus scrofa*) in Florida, USA: The Role of Humans in the Expansion and Colonization of an Invasive Wild Ungulate.” *Biological Invasions* 20: 1865–80.
- Hulme, P. E. 2009. “Trade, Transport and Trouble: Managing Invasive Species Pathways in an Era of Globalization.” *Journal of Applied Ecology* 46: 10–18.
- Iacolina, L., C. Pertoldi, M. Amills, S. Kusza, H.-J. Megens, V. A. Balteanu, J. Bakan, et al. 2018. “Hotspots of Recent Hybridization between Pigs and Wild Boars in Europe.” *Scientific Reports* 8: 17372.
- Jombart, T. 2008. “Adegenet: A R Package for the Multivariate Analysis of Genetic Markers.” *Bioinformatics* 24: 1403–5.
- Koen, E. L., E. J. Newton, and E. H. Ellington. 2021. “Evaluating Potential Sources of Invasive Wild Pigs in Ontario.” *Ecology and Evolution* 11: 14744–57.
- Koen, E. L., E. Vander Wal, R. Kost, and R. K. Brook. 2018. “Reproductive Ecology of Recently Established Wild Pigs in Canada.” *American Midland Naturalist* 179: 275–286.
- Lewis, J. S., M. L. Farnsworth, C. L. Burdett, D. M. Theobald, M. Gray, and R. S. Miller. 2017. “Biotic and Abiotic Factors Predicting the Global Distribution and Population Density of an Invasive Large Mammal.” *Scientific Reports* 7: 44152.
- Lin, Y. C., H. M. Hsieh, J. C. I. Lee, C. T. Hsiao, D. Y. Lin, A. Linacre, and L. C. Tsai. 2014. “Establishing a DNA Identification System for Pigs (*Sus scrofa*) Using a Multiplex STR Amplification.” *Forensic Science International-Genetics* 9: 12–19.
- Livestock Conservancy. 2023. *What Are Heritage Breeds?* Pittsboro, NC: Livestock Conservancy. [livestockconservancy.org/heritage-breeds/](https://livestockconservancy.org/heritage-breeds/).
- Lorenzini, R., R. Fanelli, F. Tancredi, A. Siclari, and L. Garofalo. 2020. “Matching STR and SNP Genotyping to Discriminate between Wild Boar, Domestic Pigs and Their Recent Hybrids for Forensic Purposes.” *Scientific Reports* 10: 3188.
- Lowe, S., M. Browne, S. Boudjelas, and M. De Poorter. 2000. *100 of the World's Worst Invasive Alien Species: A Selection from the Global Invasive Species Database*. Auckland: The Invasive Species Specialist Group.
- Mack, R. N., D. Simberloff, W. M. Lonsdale, H. Evans, M. Clout, and F. A. Bazzaz. 2000. “Biotic Invasions: Causes, Epidemiology, Global Consequences, and Control.” *Ecological Applications* 10: 689–710.

- Mayer, J. J., and I. L. Brisbin, Jr. 1991. *Wild Pigs in the United States: Their History, Morphology and Current Status*. Athens, GA: University of Georgia Press.
- McClure, M. L., C. L. Burdett, M. L. Farnsworth, M. W. Lutman, D. M. Theobald, P. D. Riggs, D. A. Grear, and R. S. Miller. 2015. "Modeling and Mapping the Probability of Occurrence of Invasive Wild Pigs across the Contiguous United States." *PLoS One* 10: e0133771.
- McKee, S., A. Anderson, K. Carlisle, and S. A. Shwiff. 2020. "Economic Estimates of Invasive Wild Pig Damage to Crops in 12 US States." *Crop Protection* 132: 105105.
- Miller, R. S., S. J. Sweeney, C. Sloomaker, D. A. Grear, P. A. Di Salvo, D. Kiser, and S. A. Shwiff. 2017. "Cross-Species Transmission Potential between Wild Pigs, Livestock, Poultry, Wildlife, and Humans: Implications for Disease Risk Management in North America." *Scientific Reports* 7: 7821.
- Nolte, D. L., and W. S. Anderson. 2015. "Collaborating to Halt Feral Swine Damage." *The Wildlife Professional* 9: 56–59.
- Novomestky, F., and S. Nadarajah. 2016. *truncdist: Truncated Random Variables (R Package Version 1.0-2)*. Vienna: R: A Language and Environment for Statistical Computing. The Comprehensive R Archive Network.
- Parker, I. M., D. Simberloff, W. M. Lonsdale, K. Goodell, M. Wonham, P. M. Kareiva, M. H. Williamson, et al. 1999. "Impact: Toward a Framework for Understanding the Ecological Effects of Invaders." *Biological Invasions* 1: 3–19.
- Petrelli, S., M. Buglione, V. Maselli, C. Troiano, G. Larson, L. Frantz, A. Manin, et al. 2022. "Population Genomic, Olfactory, Dietary, and Gut Microbiota Analyses Demonstrate the Unique Evolutionary Trajectory of Feral Pigs." *Molecular Ecology* 31: 220–237.
- Pfaffelhuber, P., and A. Rohde. 2022. "A Central Limit Theorem Concerning Uncertainty in Estimates of Individual Admixture." *Theoretical Population Biology* 148: 28–39.
- Poudyal, N. C., C. Caplenor, O. Joshi, C. Maldonado, L. I. Muller, and C. Yoest. 2017. "Characterizing the Economic Value and Impacts of Wild Pig Damage on a Rural Economy." *Human Dimensions of Wildlife* 22: 538–549.
- Purcell, S., B. Neale, K. Todd-Brown, L. Thomas, M. A. R. Ferreira, D. Bender, J. Maller, et al. 2007. "PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses." *American Journal of Human Genetics* 81: 559–575.
- R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. <https://www.R-project.org/>.
- Ramos, A. M., R. P. M. A. Crooijmans, N. A. Affara, A. J. Amaral, A. L. Archibald, J. E. Beever, C. Bendixen, et al. 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: e6524.
- Schleimer, A., L. Richart, F. Drygala, F. Casabianca, O. Maestrini, H. Weigand, C. Schwartz, M. Mittelbronn, and A. C. Frantz. 2022. "Introgressive Hybridisation between Domestic Pigs (*Sus scrofa domestica*) and Endemic Corsican Wild Boars (*S. s. meridionalis*): Effects of Human-Mediated Interventions." *Heredity* 128: 279–290.
- Smyser, T., P. Pfaffelhuber, R. Giglio, M. DeSaix, A. Davis, C. Bowden, M. Tabak, et al. 2024. "Probabilistic Genetic Identification of Wild Boar Hybridization to Support Control of Invasive Wild Pigs (*Sus scrofa*)." Dataset. Dryad. <https://doi.org/10.5061/dryad.tjq2b4w48>.
- Smyser, T. J., M. A. Tabak, C. Sloomaker, M. S. Robeson, R. S. Miller, M. Bosse, H.-J. Megens, et al. 2020. "Mixed Ancestry from Wild and Domestic Lineages Contributes to the Rapid Expansion of Invasive Feral Swine." *Molecular Ecology* 29: 1103–19.
- Snow, N. P., M. A. Jarzyna, and K. C. VerCauteren. 2017. "Interpreting and Predicting the Spread of Invasive Wild Pigs." *Journal of Applied Ecology* 54: 2022–32.
- Stegeman, L. C. 1938. "The European Wild Boar in the Cherokee National Forest, Tennessee." *Journal of Mammalogy* 19: 279–290.
- Strickland, B. K., M. D. Smith, and A. L. Smith. 2020. "Wild Pig Damage to Resources." In *Invasive Wild Pigs in North America*, edited by K. C. VerCauteren, J. C. Beasley, S. S. Ditchkoff, J. J. Mayer, G. J. Roloff, and B. K. Stickland, 143–174. Boca Raton, FL: CRC Press.
- Szántó-Egész, R., A. Jánosi, A. Mohr, G. Szalai, E. Koppányné Szabó, A. Micsinai, R. Sipos, J. Rátky, I. Anton, and A. Zsolnai. 2016. "Breed-Specific Detection of Mangalica Meat in Food Products." *Food Analytical Methods* 9: 889–894.
- Szemethy, D., B. Mihalik, K. Frank, T. Nagy, D. Ujvary, S. Kusza, L. Szemethy, E. Barta, and V. Steger. 2021. "Development of Wild Boar Species-Specific DNA Markers for a Potential Quality Control and Traceability Method in Meat Products." *Food Analytical Methods* 14: 18–27.
- Tabak, M. A., A. J. Piaggio, R. S. Miller, R. A. Sweitzer, and H. B. Ernest. 2017. "Anthropogenic Factors Predict Movement of an Invasive Species." *Ecosphere* 8: e01844.
- Waithman, J. D., R. A. Sweitzer, D. van Vuren, J. D. Drew, A. J. Brinkhaus, and I. A. Gardner. 1999. "Range Expansion, Population Sizes, and Management of Wild Pigs in California." *Journal of Wildlife Management* 63: 298–308.
- Warr, A., N. Affara, B. Aken, H. Beiki, D. M. Bickhart, K. Billis, W. Chow, et al. 2020. "An Improved Pig Reference Genome Sequence to Enable Pig Genetics and Genomics Research." *GigaScience* 9: g1aa051.
- Wells, C. P., P. Lavretsky, M. D. Sorenson, J. L. Peters, J. M. DaCosta, S. Turnbull, K. J. Uyehara, et al. 2019. "Persistence of an Endangered Native Duck, Feral Mallards, and Multiple Hybrid Swarms across the Main Hawaiian Islands." *Molecular Ecology* 28: 5203–16.
- White, S. 2011. "From Globalized Pig Breeds to Capitalist Pigs: A Study in Animal Cultures and Evolutionary History." *Environmental History* 16: 94–120.
- Wilcove, D. S., D. Rothstein, J. Dubow, A. Phillips, and E. Losos. 1998. "Quantifying Threats to Imperiled Species in the United States." *BioScience* 48: 607–615.
- Wilkinson, S., A. L. Archibald, C. S. Haley, H.-J. Megens, R. P. M. A. Crooijmans, M. A. M. Groenen, P. Wiener, and R. Ogden. 2012. "Development of a Genetic Tool for Product Regulation in the Diverse British Pig Breed Market." *BMC Genomics* 13: 580.
- Wringe, B. F., E. C. Anderson, N. W. Jeffery, R. R. E. Stanley, and I. R. Bradbury. 2019. "Development and Evaluation of SNP Panels for the Detection of Hybridization between Wild and Escaped Atlantic Salmon (*Salmo salar*) in the Western



Atlantic.” *Canadian Journal of Fisheries and Aquatic Sciences* 76: 695–704.

Zadik, B. J. 2005. “The Iberian Pig in Spain and the Americas at the Time of Columbus.” Master of Arts, Latin American Studies, University of California.

### 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:** Smyser, Timothy J., Peter Pfaffelhuber, Rachael M. Giglio, Matthew G. DeSaix, Amy J. Davis, Courtney F. Bowden, Michael A. Tabak, et al. 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>