Population genomic, olfactory, dietary, and gut microbiota analyses demonstrate the unique evolutionary trajectory of feral pigs

Simona Petrelli1, Maria Buglione1, Valeria Maselli1, Claudia Troiano2, Greger Larson3, Laurent Frantz4, Aurelie Manin3, Ezio Ricca1,5, Loredana Baccigalupi6,6, Dominic Wright7, Christian Pietri8, Domenico Fulgione1,5

1Department of Biology, University of Naples Federico II, Naples, Italy
2Department of Humanities, University of Naples Federico II, Naples, Italy
3The Palaeogenomics & Bio-Archeology Research Network, Research Laboratory for Archaeology and History of Art, University of Oxford, Oxford, UK
4School of Biological and Chemical Sciences, Queen Mary University of London, London, UK
5Task Force on Microbiome Studies, University of Naples Federico II, Napoli, Italy
6Department of Molecular Medicine and Medical Biotechnology, University of Naples Federico II, Naples, Italy
7IFM Biology, AVIAN Behavioural Genomics and Physiology Group, Linköping University, Linköping, Sweden
8Fédération Départementale des Chasseurs de Haute-Corse (FDCHC), Résidence Nouvelle-Corniche, St Joseph, Bastia, France

Abstract
Domestication is an intriguing evolutionary process. Many domestic populations are subjected to strong human-mediated selection, and when some individuals return to the wild, they are again subjected to selective forces associated with new environments. Generally, these feral populations evolve into something different from their wild predecessors and their members typically possess a combination of both wild and human selected traits. Feralisation can manifest in different forms on a spectrum from a wild to a domestic phenotype. This depends on how the rewilded domesticated populations can readapt to natural environments based on how much potential and flexibility the ancestral genome retains after its domestication signature. Whether feralisation leads to the evolution of new traits that do not exist in the wild or to convergence with wild forms, however, remains unclear. To address this question, we performed population genomic, olfactory, dietary, and gut microbiota analyses on different populations of Sus scrofa (wild boar, hybrid, feral and several domestic pig breeds). Porcine single nucleotide polymorphisms (SNPs) analysis shows that the feral population represents a cluster distinctly separate from all others. Its members display signatures of past artificial selection, as demonstrated by values of $F_{ST}$ in specific regions of the genome and bottleneck signature, such as the number and length...
1 | INTRODUCTION

When domestic populations that have undergone millennia of human-induced and relaxed selection return to living in environments free of direct human interactions, they are once again faced with natural selective pressures. Adaptation to new environments through natural selection in populations that return to the wild, referred to as feralisation (Price, 1984), sometimes leads to the astatistic re-emergence of predomestication traits (Zhang, Wang, et al., 2020). Feralisation can be seen as a reversal of domestication, even if it does not result in the complete return to a wild “ancestral” form (Johnsson et al., 2016). This process has occurred independently on multiple occasions in populations of domesticated animals, such as chickens (Roberts, 1991), dogs (Zhang, Wang, et al., 2020), horses (Goodwin, 2007), pigs (Evin et al., 2015) and sheep (Van Vuren & Bakker, 2009). Thus, this phenomenon offers a unique opportunity to study how natural and sexual selection act on a domestic population whose survival is less immediately and less directly governed by humans (Der Sarkissian et al., 2015; Johnsson et al., 2016).

Occasionally, feralisation can restore predomestication features such as fangs and crests (as Mapston, 2004 reports for hogs). Importantly, the reappearance of these traits does not necessarily rely on the same genetic variation present in wild populations but can be the result of new genetic mechanisms (Gering, Incorvaia, Henriksen, Conner, et al., 2019). Indeed, recent studies of feral chickens and sheep found limited evidence for overlapping genomic regions involved in domestication and feralisation processes (Johnsson et al., 2016). For instance, the feral chickens of Kauai do not maintain the traits selected during domestication but instead possess signatures of selection on genomic regions involved in comb mass and fecundity, showing different phenotypes of these traits when contrasted with their domestic counterparts (Johnsson et al., 2016).

Swine (Sus scrofa) are an excellent model system to investigate feralisation since they possess one of the widest native distributions among all terrestrial mammals worldwide, are considered an invasive pest to agriculture and a threat to biodiversity due to their plasticity in feeding behaviour (Ballari & Barrios-Garcia, 2014) and high reproductive rates. The early domestication process of west Palaearctic pigs began in the Near East ~10,500 years B.P. (Frantz et al., 2019). They then dispersed alongside people into Europe where, as a result of gene flow with local wild boar, they lost virtually all nuclear and mitochondrial traces of their Near Eastern forebears (Frantz et al., 2019). Although gene flow with wild populations has played a significant role in their evolutionary history, domestic pigs have developed a variety of dramatically different phenotypes when compared to their wild progenitor (for example pigs possess a greater number of vertebrae) as a result of strong selection in some specific breeds.

In pigs, a clear signature of domestication is the reduction of the sense of smell and a decrease in the ratio of brain weight to body weight, due to both a stable supply of resources and variability of ancestry in domestic pig breeds (Lega et al., 2015; Lord et al., 2020; Maselli, Polese, et al., 2014; Zhu et al., 2017), especially in cerebral regions involved in responses to external stimuli (reviewed in Zeder, 2012). This is the case for olfaction, one of the keenest senses that pigs possess, which is involved in the acquisition of food and water (Cronely et al., 2003), sexual activity and predator avoidance through a complex pathway beginning in the olfactory epithelium (Nguyen et al., 2012).

In wild boar, food selection is largely affected by both innate physiological responses to olfactory information and learned behaviours. In fact, it has been demonstrated that flavour learning takes place during the prenatal stage (Fulgione et al., 2017; Oostindjer et al., 2011), when the foetus is exposed to the mother’s amniotic fluid and the ability to discern food and develop trophic preferences is acquired (Fulgione et al., 2017). Wild boars are gregarious in nature, with young piglets following their mothers for about a year, which enables them to learn olfactory traces and feeding areas (Allwin et al., 2016).

Human-induced selection and/or relaxation of natural selection in domestic pigs has also resulted in a shortening of the intestinal tract and of its digestive capabilities (Uhr, 1995). A reshaping of the gut microbial community has been observed and indicated as a consequence of feed ingredients, antibiotic administration, reduced interaction with wild environments and intensive farming (Ferrario et al., 2017; Mckenzie et al., 2017; Ushida et al., 2016). The composition of the intestinal microbiota begins during infancy (de Agüero et al., 2016) and, although strongly influenced by the diet (Arrieta et al., 2014; Laursen et al., 2017), can be modulated by exposure to environmental factors (Yatsunenko et al., 2012).

Feralisation in pigs takes place through a variety of mechanisms including animal abandonment, free grazing practices or farming practices such as pig transhumance (Albarella et al., 2011). Most derived traits found exclusively in domestic forms are thought to lead to reduced fitness in wild settings (Battocchio et al., 2017; Fang et al., 2009). However, other traits can be beneficial during feralisation, for
example increased fertility in Italian feral pigs (Fulgione et al., 2016) (Figure 1) and a regained relative brain size with smaller diencephalon and medulla oblongata, and a larger cerebellum in feral pigs from the Galapagos (Kruska & Röhrs, 1974). A similar process is taking place in Sardinian feral pigs, which regained both the brain size and the cell density of the olfactory mucosa found in wild boar, but not the olfactory markers (Maselli, Polese, et al., 2014).

In some cases, hybridisation arises among feral populations due to spontaneous admixture of its members with wild boar (ex- oferal sensu Gering, Incorvaia, Henriksen, Conner, et al., 2019); though some human practices can counteract this process. For example, there is evidence that a small amount of artificial selection still acts on Sardinian feral pig populations and that this limits gene flow with the wild form. Specifically, herders kill piglets which exhibit a hybrid phenotype (wild type striped coats) because of their reduced growth. As reported by Albarella et al. (2007), “all breeders agreed that interbreeding between wild boar and domestic pigs occurs, but the hybrids are invariably slaughtered immediately, as they do not grow sufficiently. Hybridisation is therefore regarded as inevitable but undesirable”. Herders can also maintain selection during the sows’ oestrous cycle by sequestering them in a pigsty and providing them with food and water during the gestation period.

Feral individuals can respond to environmental variation either within their lifetimes through plasticity (i.e., phenotypic modifications that are not related to an immediate heritable genetic change) or through evolutionary adaptation over several generations (Harrisson et al., 2014).

Here, we sought to characterise feral pig populations and clarify how they adapt to a wild context. Indeed, the myriad of different adaptive pathways and phenotypes associated with feralisation raises numerous questions including: does feralisation in pigs lead to the reversal to ancestral “predomestication” traits (e.g., through gene flow with local wild boar) or does it mostly result in novel adaptations that do not exist in wild populations? To address these questions, we generated and analysed genetic data obtained from several swine populations from Sardinia, Corsica and Southern Italy affected to different degrees by natural and human-induced selection. Specifically, we compared the feral pig genomic ancestry, olfactory system, diet and gut microbiota with that of multiple domestic and wild populations in different regions.

2 | MATERIALS AND METHODS

2.1 | Categories of sample assignment

Given the high phenotypic and social behavioural variability expressed by S. scrofa, mainly due to experienced history, sampled individuals were binned into different experimental categories based on phenotype traits such as coat colour, shapes of ears and tails, presence/absence of fangs, number of active nipples in farrowing

![FIGURE 1](wileyonlinelibrary.com) Some feral phenotypes of individuals included in this study [Colour figure can be viewed at wileyonlinelibrary.com]
boar. The incidence of hybrid individuals (Figure S3) co-occurring with wild (42°16.60’N/9°28.0’E) (Figure Southern Italy) (Figure wild boar and feral pig. The region of Campania (40°30’N/15°16’E, Sardinia, Italy) (Figure 2c–d) were selected due to admixture events between wild boar and domestic pigs escaped from farms; sampled in Southern Italy and Corsica.

Although hybrids and feral pigs both represent two forms of rewilding, we differentiated them based on genealogy and selection pressure (hybrids are only subjected to natural selection, while feral pigs are still subjected to limited human intervention only for avoiding admixture with wild boar).

2.2 | Study area

Three different geographical contexts (Figure 2a–d) were selected in order to sample all the categories considered. The upland of Golgo, in the Western part of the Ogliastra region (40°5.21’N/9°40.2’E, Sardinia, Italy) (Figure 2c; Figure S1), was chosen as a site to study feralisation process due to the presence of sympatric populations of wild boar and feral pig. The region of Campania (40°30’N/15°16’E, Southern Italy) (Figure 2d) and some territories of Upper Corsica (42°16.60’N/9°28.0’E) (Figure 2b; Figure S2) were selected due to the incidence of hybrid individuals (Figure S3) co-occurring with wild boar.

2.3 | Sampling and collection of biological samples

From February 2017 to February 2019, we conducted an extensive field survey to collect biological materials from wild boar, domestic pigs, hybrids and feral pigs. We collected ~5 cm sections of muscle tissue for genetic characterisation from all animals immediately after death. Furthermore, from the sympatric Sardinian specimens we also collected a ~7 cm section of olfactory epithelium for gene expression analysis as well as excrements, sampled immediately after observation of animal defecation or directly from the colon of culled individuals, to characterise their feeding habits and gut microbiota community through a DNA metabarcoding approach. Each record was georeferenced using a global positioning system (GPS; UTM-WSG 84) and loaded in a GIS environment using QGIS 3.4.2.

2.4 | Genomic characterisation of populations (Illumina 60K SNP)

DNA isolation from tissues was performed using the Qiagen DNeasy 96 Blood & Tissue Kit (QIAGEN GmbH Valencia, CA, USA) following manufacturer’s instructions. To discriminate hybrids from Southern Italy we applied three approaches: coat colour observation, MC1R sequence (useful to discriminate F1 hybrids), and single nucleotide polymorphisms (SNPs) analysis (using the Illumina 60 K SNP).

The latter was used both to discriminate hybrid individuals with greater accuracy compared with the other methods and to analyse the genomic relationships between the studied forms. We genotyped 96 pig samples (Table 1) using the Illumina 60 K SNP (Ramos et al., 2009), and pooled them to 97 external samples (Table 2; Yang et al., 2017). The quality of raw data was assessed using PLINK (Purcell et al., 2007). We removed genotypes with a minor allele frequency (MAF) lower than 0.05, resulting in a final data set of 39,605 SNP. These were used to construct neighbour joining (NJ) trees, obtained using the R package ape (Paradis et al., 2004) and visualised in figtree v1.4. To compute support for each node we bootstrapped our ped file 100 times and recomputed a NJ tree each time.

Furthermore, SNPs were useful for identifying signatures of selection (SOS) on genomes and for the assessment of runs of homozygosity (ROHs). Long ROHs usually indicate intense selection pressures (Metzger et al., 2015) and/or recent inbreeding (Al-Mamun et al., 2015), whereas short ROHs are often caused by more ancient bottleneck effects (Al-Mamun et al., 2015) and/or inbreeding events (Howrgan et al., 2011).

We then removed data in high linkage disequilibrium ($r^2 > 0.2$) resulting in 15,531 SNPs for principal component (PCA) and admixture analyses. PCA was carried out in PLINK and visualised by R programming language (R Core Team, 2013) using the ggplot2 package (Wickham, 2016) in order to position feral pigs within Suidae family.

To detect SOS in the Sardinian system, we calculated the fixation index ($F_{ST}$) (Weir & Cockerham, 1984) as measure of group differentiation per locus following the pipeline proposed by Porto-Neto et al. (2013) and comparing the value for each population with that of the meta-population. We divided SNP data into genic (cis-regulatory 5’ and 3’ UTR regions, nonsynonymous SNPs that would experience increased selective pressure) and nongenic (Barreiro et al., 2008). To statistically test this arrangement, we compared the average population values and their variance. SOS were identified by calculating the average and standard deviation of $F_{ST}$ values for each population and identifying the regions that have values higher than the average plus three standard deviations (mean $F_{ST} ± 3$ SD) (Porto-Neto et al., 2014). Gene annotation was performed using ENSEMBL comparative genomic resources and further filtering, was carried out in the Panther database based on gene biological functions (Mi et al., 2016).
**FIGURE 2**  Sampling. (a) Study areas. (b) Corsican area. (c) Sardinian area. (d) Southern Italy area. (e) Principal component analysis based on 15,531 SNPs. Each point represents a genotyped individual coloured according to geographic origin and category. The star-shaped symbols represent the different Italian pig breeds: cyan, Calabrese breed; pink, Cinta Senese breed; grey, Casertana breed; and yellow, Nero siciliano breed. Blue squares: Italian Wild Boar, with black border individuals from Northern-Central Italy and without black border individuals from Southern Italy [Colour figure can be viewed at wileyonlinelibrary.com]
TABLE 1 Swine forms genotyped

| Population                | Number of individuals |
|---------------------------|-----------------------|
| Southern Italy wild boars | 9                     |
| Sardinian wild boars      | 10                    |
| Corsican wild boars       | 5                     |
| Southern Italy hybrids    | 31                    |
| Corsican hybrids          | 5                     |
| Sardinian ferals          | 23                    |
| Corsican breed            | 5                     |
| Pigs                      | 8                     |

TABLE 2 External genotypes from Yang et al. (2017)

| Populations                      | Number of individuals |
|----------------------------------|-----------------------|
| Northern-Central Italy wild boars| 19                    |
| Sardinian wild boars             | 20                    |
| Calabrese breed                  | 15                    |
| Cinta Senese breed               | 13                    |
| Casertana breed                  | 15                    |
| Nero Siciliano breed             | 15                    |

ROHs were calculated for Sardinian wild boar and feral pigs, and for Southern Italy wild boar and pigs, with the ROH tool in PLINK setting the minimum length to 500 kb with at least 50 SNPs, allowing five missing calls and one heterozygous SNP. We estimated the number of ROHs per each group and the sum of all ROH lengths for each individual within the population.

2.5 | Differential gene expression of the odorant binding protein gene (OBP)

Total RNA was isolated from each olfactory mucosa using TRI Reagent (EuroClone) according to the manufacturer’s instructions. For cDNA synthesis with integrated removal of the genomic DNA contamination, QuantiTect Reverse Transcription Kit (QIAGEN) was used as described by the manufacturer. qPCRs were carried out with QuantiFast SYBR Green PCR Kit (Qiagen) on a Rotor-Gene Q cycler (Qiagen). The assays were performed in 25 µL final volume of reaction containing 1 µM of each primer (OBP-forward: 5’-CGGAACCAAACAAGAAGGC-3’; OBP-reverse: 5’-CCGGTCTCTTCTGTGACCT-3’) to amplify the target region of 214 bp. After an initial enzyme activation at 95°C for 5 min, 40 amplification cycles were performed using 95°C for 10 s for denaturation and 60°C for 30 s for annealing. Quantitative qPCR analysis was conducted by using the 2(-ΔΔC[T]) method. For each qPCR experiment, data were normalised to the expression of the β-actin housekeeping gene (Maselli, Polese, et al., 2014). In order to measure reaction efficiency, a standard curve was generated using the standards of 1000, 100, 10, 1 ng of total starting RNA. Statistical analyses were performed via a one-way ANOVA test.

2.6 | Faecal DNA extraction

All genetic analyses on excrements were conducted in a dedicated laboratory used exclusively for environmental DNA processing. DNA extractions, based on the hexadecyltrimethylammonium bromide (CTAB) method (Doyle & Doyle, 1987) and quality checks were performed as reported in Bugione et al. (2018).

2.7 | Marker selection for diet analysis

For the diet characterisation, we performed metabarcoding analysis on feral and wild boar stool samples, whereas we estimated the diet of domestic pigs from the composition of the commercial feed supplied to the animals by farmers. Considering the omnivorous food habits of wild boar and feral pigs, we selected g and h primers (Taberlet et al., 2007) for amplifying 200 bp of the P6 loop of the chloroplast trnL (UAA) intron in angiosperms and gymnosperms, and MICOlintF (Leray et al., 2013), associated with PolyShortCoIR (Carr et al., 2011), targeting a 350 bp fragment of mitochondrial cytochrome oxidase subunit I (COX1) to detect metazoan component of the diet. Moreover, a specific blocking oligonucleotide (Robeson et al., 2018) was used in order to reduce DNA amplification of its COX1, minimising interference of the host sequences with others. All primers, except Blocking S. scrofa, were modified by the addition of a specific Illumina adapter to their 5’ ends that acted as an identifier to recover data from each sample post-sequencing (Colissac, 2012; Valentini et al., 2009).

2.8 | Detection of diet component

To analyse the plant component of diet, a PCR was performed in a total volume of 25 µl with 100–150 ng of DNA as the template, 1 U of Taq Solis polymerase (BioDyne), 2.5 µl of 10x Buffer B (0.8 M Tris-HCl, 0.2 µM (NH₄)₂ SO₄, 0.2% w/v Tween-20), 2.5 µl of 25 mM MgCl₂, 4 µl of 2.5 mM dNTP mix and 2.5 µl of 25 µM of g and h primers. Three independent PCR replicates were performed for each sample. Positive and negative controls were included in each amplification to monitor the performance of the PCR process. The PCR conditions began with an initial denaturation step at 95°C for 3 min, followed by 35 cycles of 95°C for 3 min, 55°C for 1 min, and 72°C for 30 s and a final extension step at 72°C for 5 min. The same protocol was used to characterise animal composition of the diet, except for 0.125 µl of 100 µM of MICOlintF, PolyShortCoIR and Blocking S. scrofa primers, 1 µl BSA 2 mg/ml (Sigma) and extension at 72°C for 40 s.

2.9 | Post PCR processing

Prior to sequencing, PCR products were purified using Illustra GFX PCR DNA and Gel Band Purification Kit (GE Healthcare) and
were omitted.

in equimolar concentrations (Harris et al., 2010) in all amplicons. We pooled samples according to S. scrofa form and diet content:

Pool W1: plant component of wild boar diet (N = 7);  
Pool F1: plant component of feral diet (N = 11);  
Pool W2: animal component of wild boar diet (N = 7);  
Pool F2: animal component of feral diet (N = 11).

The diet data was analysed considering all individuals in a pool as a single group in order to obtain the diet of wild boar or feral pig as a whole.

2.10 | Sequencing and bioinformatic analysis

Large-scale sequencing was performed on an Illumina MiSeq platform (Illumina, Inc.) with a 2 × 150 bp paired-end run (for tnrL ampli- cons) and 2 × 300 bp paired-end run, (for COX1 amplicons) following the Nextera DNA sample preparation protocol at the Genomix4Life Srl (http://www.genomix4life.com/it/). Sequencing data analysis was conducted as previously described in Buglione et al. (2020) and the Blast results were filtered in relation to the regional list of plant (Fenu et al., 2010; Maxia et al., 2003; www.actaplanturn.org) and animal (personal database) species to increase the accuracy of the automatic taxonomic assignation.

2.11 | Microbiota characterisation

A total of 25 individuals (seven wild boar, 10 feral pigs and eight do- mestic pigs) were used for gut microbiota analyses. Partial 16S rDNA gene sequences were amplified using the primer pair Probio_Uni and Probio_Rev, which target the V3 region as previously described in Milani et al. (2013). Gene sequencing was performed on an Illumina MiSeq platform at the DNA sequencing facility of GenProbio srl (www.genprobio.com) according to the protocol reported in Milani et al. (2013). Following sequencing and demultiplexing, the obtained reads of each sample were filtered to remove low quality and poly- clonal sequences. All quality-approved, trimmed and filtered data were exported as fastq files and processed using a script based on the QIIME software suite (Caporaso et al., 2010). Paired-end reads were assembled to reconstruct the complete Probio_Uni/Probio_Rev amplicons. We retained sequences that after quality control were 140–400 bp in length and with mean sequence quality score >20. Sequences with homopolymers >7 bp and mismatched primers were omitted.

2.12 | Statistical analysis

For microbiota data, the calculation of β diversity by performing a principal coordinates analysis (PCoA) using Bray–Curtis dissimilarity (Bray & Curtis, 1957) was performed in PAST v 3.2. (Hammer et al., 2001), which includes the relative abundance information of bacte- ria genera. Statistically significant differences between groups were tested with one-way PERMANOVA in PAST v 3.2. In order to describe the core of the intestinal microbiota across the three experimental categories, a symmetrical Venn diagram of the shared and exclusive microbial genera of the categories under study (count > 0) was generated. The calculation was obtained from the presence/absence data of the bacterial genera in the three categories using the free Venn Diagram Tool (http://bioinformatics.psb.ugent.be/webtools/Venn/).

3 | RESULTS

3.1 | Genome-wide SNPs for population assignment and signatures of selection

The variation accounted for by the first two principal components (27.5%) evidenced that the genotyped individuals clustered mostly according to their geographic location and experimental catego- ries (Figure 2a–d). More specifically, Sardinian feral pigs clustered separately from all other populations (Figure 2e; Figures S4 and S5) except for four individuals close the different breeds of domestic pigs, and only one individual near Sardinian wild boar. Finally, we also identified a single Sardinian wild boar that completely overlaps with the Sardinian feral pigs (Figure 2e; Figures S4 and S5). Corsican hybrids fell in an intermediate position between pigs and wild boar, while the Southern Italian hybrid populations clustered together with their local wild counterparts. Domestic pig breeds together dia- metrically opposite to the wild boar (Figure 2e).

Both natural and human-induced selection can lead to genomic changes in terms of long ROHs. Moreover, ROHs arise from back- ground relatedness promoted by demographic processes that in- crease homozygosity and reduce population size (for example, cultural and/or social factors that favour consangunity and natural selection) (Szpiech et al., 2013). We identified 568 ROHs across four different groups: feral pigs, domestic pigs and the wild boar popu- lations from Southern Italy and Sardinia. The relationship between per-individual ROH number and total genomic length covered by ROHs per individual varied considerably both among and within considered swine forms (Figure 3). Feral populations had signifi- cantly fewer and shorter ROHs (146 in total, 43 Mb long) relative to both Southern Italian and Sardinian wild boar, which possessed more and longer ROHs (174 and 98 Mb; 167 and 69 Mb, respectively). Signaturees of selection were identified by quantifying the differ- ence in allelic frequency using FST in the SNPs of feral pigs relative to wild boar populations and domestic pigs (Figure 4). The smoothed value of median FST (Figure 4a) showed, as expected, a clear signal that is higher for domestic pigs when compared with the two wild living forms, principally due to the SNP array, where most of the polymorphisms were discovered in commercial European and US breeds (Ramos et al., 2009). However, recent studies have shown
that this Illumina array provides a solid assessment of genomic diversity in comparative studies between European populations despite the ascertainment bias (Herrero-Medrano et al., 2014).

It is interesting to note that feral forms are closer to wild ones rather than to pigs, suggesting they show genomic signatures of natural selection, which usually acts as a balancing selection and, in the case of ferals, also as a purifying selection. The peaks above the smoothed line indicate regions exceeding the significance threshold of the $F_{ST}$ value and, hence, allegedly affected by positive selection. The peaks above the smoothed line indicate regions exceeding the significance threshold of the $F_{ST}$ value and, hence, allegedly affected by positive selection (artificial or natural, presumably the former in domestic pigs). In domestic pigs, these regions contain genes for obesity and intramuscular fat deposition (NDUFS4 and AGBL4: Chen et al., 2013; Locke et al., 2015), carbohydrate (WDTC1: Liu et al., 2015) and lipidic (PSAP: Jégou et al., 2016) metabolic processes and adaptive immune response (CD3E: Meddens et al., 2018).

Interestingly, feral pigs also possess some regions with $F_{ST}$ values exceeding the threshold and hosting genes involved in meat quality (MYOT and FRMD8: Kim et al., 2020; Velez-Irizarry et al., 2019), feed efficiency (VPR1 and PFKFB4: Hao et al., 2016; Hou et al., 2018; Yun et al., 2012), uptake and carbohydrate metabolism (SLC16A1 and LNPEP: Bosse et al., 2015; Ghareeb et al., 2015; Liu et al., 2015), obesity and fat deposition (AGBL4 and NR2C2: Locke et al., 2015; Ramayo-Caldas et al., 2014) spermatogenesis and reproduction (WT1, HSPA4, HSPA9 OSBP2: Dun et al., 2012; Hashemitabar et al., 2015; Lahbib-Mansais et al., 1997; Zhang, Yang, et al., 2020), brain development (DAB1: Long et al., 2011) and immunity (SH2D2A, ANKRD34B, ERAP2, BHLHE40, MASP1: Al-Shaibi & Ghosh, 2004; André et al., 2010; Berge et al., 2012; Boldt et al., 2016; Lin et al., 2014). Some regions under selection have been identified in the wild form as well, although they only slightly exceed the threshold value. These regions include genes regulating carbohydrate metabolism and backfat (LNPEP and GLIPR1: Bosse et al., 2015; Liu et al., 2015, 2019), immunity (ERAP2: André et al., 2010) and spermatogenesis (HSPA4 and HSPA9: Dun et al., 2012; Hashemitabar et al., 2015).

Previous studies have revealed regions that were differentially selected in populations of the same species. In fact, selection
preferentially targets genic over nongenic regions (Barreiro et al., 2008). It has also been observed that variants leading to amino-acid changes (nonsynonymous mutations) or located in cis-regulatory regions (5' UTR and 3' UTR) (Figure 4b) would be under stronger selective pressure than "silent" genic mutations (nongenic and intronic variants) (Barreiro et al., 2008).

**FIGURE 4**  
(a) Genome wide plots of smoothed $F_{ST}$ (red line) for three swine forms (wild boar, domestic pig and feral pig). Notice how the smoothed $F_{ST}$ highlights at least five regions under potential selection for pig populations. The horizontal line shows the significance threshold.  
(b) Manhattan plots of $F_{ST}$ for genic and nongenic regions. $F_{ST}$ values above the threshold line (black horizontal line) show SNPs under potential selection. The different SNP classes (y-axis) - nonsynonymous, cis-regulatory, nongenic and intronic – are examined across the 18 autosomal chromosomes and X chromosome, shown in different colours on the x-axis (the X chromosome represented in green on the far right) [Colour figure can be viewed at wileyonlinelibrary.com]
Thus, intronic and nongenic SNPs have higher overall $F_{ST}$ values in pigs. Conversely, nonsynonymous and cis-regulatory regions presented a significant excess of low $F_{ST}$ values in feral pigs (with values similar to those in wild boar), which can be explained by a balancing/purifying selection signature on genic regions (Figure 4b). Interestingly, none of the SNPs that fall in the genic regions of loci involved in sense of smell in feral pigs (based on the Panther database) showed significantly high $F_{ST}$ values, suggesting that these regions evolved under balancing selection.

3.2 | Adaptive strategies: Sense of smell, trophic ecology, and microbial community

Based on our results, we cannot exclude the possibility that some genes involved in olfaction are regulated in order to respond with specific adaptations to natural selection. This is the case of the OBP gene, which showed a significant (2.4-fold) increase in mRNA in feral pigs compared with domestic pigs (Figure 5a).

3.3 | Metabarcoding diet analysis

A possible recovery of the sense of smell in environments beyond the human niche could affect food choice ability. We found that feral pigs, similar to wild boar, had predominantly plant-based diets. In particular, wild boar mostly fed on 33 plant taxa assigned to 23 families (Figure 5b). Fabaceae (70.90%) and Rosaceae (14.97%) accounted for about 87% of the total plant items identified in the diet. The former family was represented mainly by acorns, Quercus sp. (70.78%) (Table S1), which is typical of woody habitats with mature beeches and oaks favoured by the European wild boar. Furthermore, the Fabaceae family showed the greatest richness in the total number of plant taxa ($N = 5$). Prunus spinosa (blackthorn) (10.92%) that produces edible fruits and roses was the most abundant species among Rosaceae. We also detected Oleaceae (8.91%), Geraniaceae (3.22%) and Asteraceae, (asters, daisies, sunflowers) (1.23%). All other families were present at lower frequencies (<1%).

On the other hand, the diet of the feral pigs included 34 plant taxa assigned to 21 families (Figure 5c). Fabaceae (spontaneous and cultivated herbaceous plants) showed the greatest diversity in number of taxa ($N = 5$) and represented the most frequently occurring family (86.32%) in the diet (Figure 5c). In particular, Trifolium sp. accounted for 73.46% of the total diet. Fabaceae was followed by Geraniaceae (5.58%), Rosaceae (1.80%), Poaceae (1.62%) and Orchideaceae (1.04%). All other families showed a frequency of occurrence <1% (Table S1). The diet of domestic pigs provided by humans was extremely simplified and based on Fabaceae (~15%), Poaceae (~80%) and Rosaceae (~5%).

The animal component of wild boar diet included seven taxa (six families in total, Figure 5d), of which 92.85% were invertebrates (i.e., Lumbriidae and Capitellidae) and 7.14% vertebrates (i.e., mammals) (Figure 5d and Table S2). The dominant invertebrate item was earthworms (Lumbricus terrestris), that accounted for 35.70% of the total animal component of the diet (Table S2).

The animal component of feral pig diet included 14 taxa (with 13 families in total, Figure 5e) and consisted of 98.31% of invertebrates (i.e., Erebidae and Capitellidae) and 1.69% of vertebrates (i.e., mammals and reptiles). This is most probably the result of the necrophagous feeding behaviour of feral pigs as supported by the observation, during field activities, off feral animals feeding on the stomach of goat carcasses.

Compared to wild boar, feral animals appeared to consume a greater diversity of animal-based food, with sequences assignable to phytophagous insects and parasites such as flies of plants and fungi (i.e., Tychius pusillus, Anthomyiidae and Halimococcidae) (Table S2). These components could easily be derived from indirect causes, such as via the ingestion of the substrates (plant, roots, fruits, leaves) on which these Arthropoda perform their life cycle. For example, Lymantria dispar accounted for 86.07% of the total animal diet of feral pigs. However, this finding is strongly influenced by a large abundance of larvae (caterpillars) in the study area during the sampling period. Further studies should be carried out to understand if this is an intentional food selection or the result of indirect ingestion. It is worth pointing out that earthworms were absent in the feral diet.

3.4 | Metabarcoding gut microbiota analysis

The total number of bacterial genera identified in intestinal samples of pigs, feral pigs and wild boar was 101, 108 and 126, respectively (Table S3). As shown by the Venn diagram (Figure 6a), a core of 97 bacterial genera were shared by all three groups. Pigs shared only one genus with feral pigs and only two with wild boar, while feral and wild animals shared nine genera (Figure 6a and Table S5). Wild animals showed the highest variability with 18 bacterial genera exclusively present in their intestine (Table S4 and Figure 6a).

By plotting individual gut communities on a multivariate space PCoA, we observed a clear discrimination in the first and second coordinates (Figure 6b), with the gut microbiota of each group clustering apart (p = .0001, permutational multivariate analysis of variance PERMANOVA). This result indicates a significant difference in
their composition. The microbial communities of wild boar (blue in Figure 6b) showed the highest variability (as shown by the multivariate space they occupy), while that of feral pigs (green) the lowest. Both of them showed a proximity with the cluster formed by the microbiota of domestic pig (orange).

4 | DISCUSSION

Although feralisation is a powerful model to explore rapid evolutionary changes in response to natural selection, it has received relatively little attention (Neaux et al., 2020). In this study, feral pigs appear genetically distinct from domestic ones and wild boar as well as from wild/domestic hybrids. Feral pigs also appear to have maintained a unique genetic signature that is distinct from that of sympatric wild boar, potentially as a result of their genealogy and of the culling of newborn wild/feral hybrids imposed by farmers.

The genomes of Southern Italian hybrids are virtually identical to those found in the sympatric wild boar population probably as a result of deep introgression into the latter. Our multiple-marker approach (coat colour, MC1R, SNPs) revealed this complex pattern that can be further explored through the characterisation of complete genomes obtained from a wide survey of Southern Italian wild boar showcasing all possible phenotypes. This would be the only way to define the level of introgression from domestic pigs present in Italian wild populations, with important implications for conservation biology.

In contrast, the Corsican populations revealed a pattern, expected in cases of hybridisation, in which pigs and wild boar clustered in extreme positions, and the hybrids positioned in between members of the two categories. Our results indicate that Sardinian feral pigs represent a lineage that is distinct from wild and domestic ones. This suggests that the morphological similarity observed between wild and feral animals is not due to the reversal through gene flow but is instead the result of unique adaptations by feralisation. This highlights the considerable flexibility of the pig genome, which is able to produce similar phenotypes in wild and feral pigs despite different genetic backgrounds.

On the basis of homozygous genomic regions, feral pigs are distinguished by few and short ROHs. This suggests that feral pigs could have derived from a small founding population and do not seem to have been affected by inbreeding in recent times (Brito et al., 2017). Thus, we cannot exclude the effect of genetic drift in Sardinian feral pig differentiation.

Feralisation, and to a lesser degree domestication, can be considered an unstable equilibrium (Gering, Incorvaia, Henriksen, Wright, et al., 2019). If direct and/or indirect human selection was to cease, would all these pig populations be routed along evolutionary trajectories that converge to the same path of wild boar? In other words, what is their main evolutionary stable attractor (ESA sensu Rand et al., 1994)? Possible answers can be provided by the pressures that feral populations are undergoing and deduced by the analysis of the variation of genomic regions (Barreiro et al., 2008), namely quantifying the difference in allelic frequency using $F_{ST}$.

A generalised genome-based approach using the median of the $F_{ST}$ values for each locus shows similarly low levels in the two wild living forms, suggesting that natural selection might be affecting feral populations on top of eventual genetic drift. In other words, there are multiple alleles preserved in the gene pool at frequencies larger than expected from genetic drift alone. Another possible explanation could be the case of highly ascertained variants at medium frequency in domestic pigs and low in both wild and feral ones, lowering the $F_{ST}$ values. Interestingly, some chromosomal regions which show $F_{ST}$ values exceeding the threshold line seem to be negatively impacted by human-induced selection for characters possibly related to feed efficiency, fat deposition, spermatogenesis and reproduction, brain development and immunity. The extent to
which newborn (which exhibit striped coats and liveliness) culling affects these signals of selection remains uncertain. The regions of the smoothed \( F_{ST} \) line that slightly exceed the threshold line in the wild boar may be linked to functional adaptations typical of an R-strategy species, including immunity and spermatogenesis.

The polymorphisms typed in this study (Illumina 60 K array) include both synonymous and nonsynonymous substitutions. The analysis of the effects of selection on the genome, considering functionally different regions, can provide a further contribution to the understanding of the pressures affecting feral populations. As expected, the highest number of polymorphisms with lower \( F_{ST} \) were found in neutral gene regions. High \( F_{ST} \) values for intronic and non-genic SNPs, which are neutral markers, shown by the three forms suggest an effect of demographic factors including reduced gene flow. The lower number of polymorphisms associated with non-synonymous functional gene regions and cis-regulatory SNPs is in agreement with the effect of selection. The values of \( F_{ST} \) under the threshold line for these polymorphisms suggest a lower degree of population differentiation for some functional genes as well.

Interestingly, the genic regions related to olfactory functions were found below the significance threshold, suggesting no signature of selection, in accordance with low levels of expression of genes involved in neurotransmission (Maselli, Polese, et al., 2014), despite the fact that feral pigs showed greater proliferation of these cells together with high expression levels of mucosal carrier proteins (Figure 5a). This lack of signal in our marker, however, potentially stems from a technical artefact. Indeed, different pig lineages often possess very different numbers of olfactory genes (Paudel et al., 2013, 2015). Such variability in copy number often results in inflated levels of observed polymorphism (see Paudel et al., 2013, 2015) probably affecting the resolution of selection scans based on allelic frequency such as \( F_{ST} \). Analyses of copy number variation based on whole genome sequencing may provide a viable alternative to better characterise the genetic differentiation at olfactory genes among wild, feral and domestic populations.

The sense of smell directly affects the diversity and richness of the feral animal diet. The broad spectrum of food ingested by feral pigs identified in this study suggests that this population regained olfactory abilities that went lost during domestication. As our results demonstrate, the richness of plants and animals consumed by feral pigs is comparable to that observed in wild boar, with some key differences. The absence of earthworms in feral pig diet could be linked to a better olfactory detection of these prey animals by wild boar and should be further investigated. Assuming that wild boar and feral pigs, living freely in the same region, have access to identical resources, the discrepancy in their dietary components could be attributed to a different knowledge of specific properties of trophic resources in terms of energetic content and healthiness. This hypothesis stems from the presence of Salicaceae, Vitaceae and Equisetaceae families found, which have antioxidant, anti-inflammatory, diuretic and mineralising properties (Asgarpanah & Roohi, 2012; Katalinić et al., 2010; Kedage et al., 2007) in the diet of wild boar but not in feral pigs. The diet of this group, however, was enriched with at least 10 plant taxa that are neglected by wild boar. It would be interesting to investigate this finding, which probably relates to differences between the intestinal microbiota of the two populations. Indeed, we cannot exclude that dietary dissimilarities are both the cause and the consequence of the intestinal microbial system.

The ability of feral pigs to achieve a greater complexity of their intestinal microbiota community is surprising. Our results suggest that microbiota simplification, a consequence of domestication (Ferrario et al., 2017; Mckenzie et al., 2017), is a modifiable condition once pigs live outside the direct influence of the human niche. This result reinforces the idea that domestic forms are, at least in some regard, capable of readapting to wild environments.

Feral pigs are a living example of how some characteristics can be gained in different ways and can be useful in developing adaptive conditions (Johnsson et al., 2016). However, it is important to note that the persistence of feral pigs depends on the close proximity to people, and their populations could disappear as traditional pastoralist practices vanish.

Overall, our study underscores the complexity of feral pig population history that has been shaped by both past human-induced and present natural selection. These results indicate that feralisation is not a mere reversal of domestication but instead involves novel adaptations that are unique to feral lineages.

Our story-tell is easily exportable to other domesticated species. In a rapidly changing world, mainly driven by economic strategies, the relationship between human and domestic animal can open up new scenarios. The abandonment of agricultural areas and traditional pastoralism leaves a trail of domestic populations without management resulting often in feral animals (Herrera, 1995; Velamazán et al., 2018). This brings up many intriguing topics for evolutionary biology. Apart from wild boar, feralisation arises in many other animal populations: horse (DeSilvey & Bartolini, 2019; Donlan et al., 2006; Naundrup & Svenning, 2015), cattle (McTavish et al., 2013) as well as goat and sheep (Bullock, 1991; Doro et al., 2016; Hess et al., 2017; Pareja et al., 2020).

Sometimes, uncontrolled feral populations can have negative impacts both on indigenous wild and domestic fauna and plants (Abe, 2021; Eldridge et al., 2020; Mihailou & Massaro, 2021; Scandurra et al., 2016). However, abandoned populations of horses, cattle and goats, or wildly grazed practices could have beneficial effects on the landscape, biological diversity as well as containment of invasive species (Pareja et al., 2020; Ruiz-Mirazo et al., 2011; Troiano et al., 2021). So, like for our study on Sus scrofa, a multidisciplinary approach will be crucial in defining the environmental contexts in which these feralisation processes take place and how the human component exerts its influence. Studying complex biological process, such as local adaption, of these feral populations can help to understand how to manage them in order to have benefits on ecosystem integrity.

Finally, feralisation in large mammals also has intriguing conservation implications, representing the restoring of megafauna that
was decimated by humans during the Quaternary Age (Ellis et al., 2021; King, 2009; Stuart, 2021).

ACKNOWLEDGEMENTS

We thank specialised hunters and herders Giorgio and Mauro Pusole for assistance in the sampling activities. LF, AM, and GL were supported by a Natural Environment Research Council grant (NE/S00078X/1). L.F. and G.L. were also supported either by European Research Council grants (ERC-2013-STG-337574-UNDEAD and ERC-2019-STG-853272-PALAEOFARM) and Natural Environmental Research Council grants (NE/K005243/1, NE/K003259/1, NE/S007067/1) or both. Open Access Funding provided by Universita degli Studi di Napoli Federico II within the CRUI-CARE Agreement.

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

AUTHOR CONTRIBUTIONS

Simona Petrelli, Maria Buglione, Valeria Maselli, Domenico Fulgione and Christian Pietri performed sample collection. Simona Petrelli, Maria Buglione, Valeria Maselli, Claudia Troiano and Domenico Fulgione performed the research. Simona Petrelli, Maria Buglione, Valeria Maselli, Claudia Troiano, Domenico Fulgione, Loredana Baccigalupi and Ezio Ricca analysed the data. Simona Petrelli, Maria Buglione, Valeria Maselli, Claudia Troiano, Domenico Fulgione, Greger Larson, Aurelie Manin, Laurent Frantz, Dominic Wright, Loredana Baccigalupi and Ezio Ricca wrote the manuscript.

DATA AVAILABILITY STATEMENT

Subset of SNP data have been made available from the Dryad Digital Repository: doi: https://doi.org/10.5061/dryad.05qftfow. Faecal metagenomic raw reads of the V3 region of the 16S rDNA gene are deposited into the SRA under accession PRJNA669575. Diet sequencing raw reads are available at the European Nucleotide Archive (ENA) under accession number PRJEB48041. External SNP data from Yang et al. (2017). Data from: Genome-wide SNP data unveils the globalization of domesticated pigs, Dryad, https://doi.org/10.5061/dryad.30tk6.

ORCID

Simona Petrelli https://orcid.org/0000-0003-2862-1397
Maria Buglione https://orcid.org/0000-0001-6022-707X
Valeria Maselli https://orcid.org/0000-0002-1452-8063
Claudia Troiano https://orcid.org/0000-0003-2448-2721
Greger Larson https://orcid.org/0000-0002-4092-0392
Laurent Frantz https://orcid.org/0000-0001-8030-3885
Aurelie Manin https://orcid.org/0000-0002-1061-3754
Ezio Ricca https://orcid.org/0000-0002-7355-0332
Loredana Baccigalupi https://orcid.org/0000-0003-3070-161X
Dominic Wright https://orcid.org/0000-0003-2329-2635
Domenico Fulgione https://orcid.org/0000-0002-5121-3298

REFERENCES

Abe, T. (2021). Impacts by feral goats on critically endangered Crepidiostomum granicolium (Compositae) endemic to the Ogasawara Islands. Plant Species Biology, 36(2), 361–367. https://doi.org/10.1111/1442-1984.12314
Albarella, U., Manconi, F., & Trentacoste, A. (2011). A week on the plateau: Pig husbandry, mobility and resource exploitation in central Sardinia. In U. Albarella, & A. Trentacoste (Eds.), Ethnozoarcheology. The present and past of human-animal relationships (pp. 143–159). Oxbow Books.
Albarella, U., Manconi, F., Vigne, J. D., & Rowley-Conwy, P. (2007). Ethnoarchaeology of pig husbandry in Sardinia and Corsica. Pigs and Humans, 10, 285–307.
Allwin, B., Swaminathan, R., Mohanraj, A., Suhas, G. N., Vedaminckam, S., Gopal, S., & Kumar, M. S. (2016). The Wild Pig (Sus scrofa) Behavior – A Retrospective Study. Journal of Veterinary Science & Technology, 7, 333. https://doi.org/10.4172/2157-7579.1000333
Al-Mamun, H. A., Clark, S. A., Kwan, P., & Gondro, C. (2015). Genome-wide linkage disequilibrium and genetic diversity in five populations of Australian domestic sheep. Genetics Selection Evolution, 47, 1–14. https://doi.org/10.1186/s12711-015-0169-6
Al-Shaibi, N., & Ghosh, S. K. (2004). A novel phosphoprotein is induced during bone marrow commitment to dendritic cells. Biochemical and Biophysical Research Communications, 321, 26–30. https://doi.org/10.1016/j.bbrc.2004.06.106
Andrés, A. M., Dennis, M. Y., Kretzschmar, W. W., Cannons, J. L., Lee-Lin, S.-Q., Hurle, B., Schwartzberg, P. L., Williamson, S. H., Bustamante, C. D., Nielsen, R., Clark, A. G., & Green, E. D. (2010). Balancing selection maintains a form of ERAP2 that undergoes nonsense-mediated decay and affects antigen presentation. PLoS Genetics, 6, e1001157. https://doi.org/10.1371/journal.pgen.1001157
Arrieta, M. C., Stiensma, L. T., Amenyooge, N., Brown, E. M., & Finlay, B. (2014). The intestinal microbiome in early life: health and disease. Frontiers in Immunology, 5, 427. https://doi.org/10.3389/fimmu.2014.00427
Asgarpanah, J., & Roohi, E. (2012). Phytochemistry and pharmacological properties of Equisetum arvense L. Journal of Medicinal Plants Research, 6, 3689–3693. https://doi.org/10.5897/JMPR12.234
Ballari, S. A., & Barrios-Garcia, M. N. (2014). A review of wild boar Sus scrofa diet and factors affecting food selection in native and introduced ranges. Mammal Review, 44, 124–134. https://doi.org/10.1111/mam.12015
Barreiro, L. B., Laval, G., Quach, H., Patin, E., & Quintana-Murci, L. (2008). Natural selection has driven population differentiation in modern humans. Nature Genetics, 40, 340. https://doi.org/10.1038/ng.78
Battocchio, D., Iacolina, L., Canu, A., & Mori, E. (2017). How much does it cost to look like a pig in a wild boar group? Behavioural Processes, 138, 123–126. https://doi.org/10.1016/j.beproc.2017.03.001
Berge, T., Grønningsæter, I. H. B., Lorvik, K. B., Abrahamsen, G., Granum, S., Sundvold-Gjerstad, V., Corthay, A., Bogen, B., & Spurkland, A. (2012). SH2D2A modulates T cell mediated protection to a B cell derived tumor in transgenic mice. PLoS One, 7, e48239. https://doi.org/10.1371/journal.pone.0048239
Boldt, A. B. W., Boschmann, S. E., Catarino, S. J., Andrade, F. A., & de Mesillas-Reason, I. J. (2016). MASP1 and MASP2. In S. Choi (Ed.), Encyclopedia of signaling molecules (pp. 1–18). Springer.
Bosse, M., Lopes, M. S., Madsen, O., Megens, H. J., Crooijmans, R. P., Frantz, L. A., Groenen, M. A. (2015). Artificial selection on in-. Mammal Review, 44, 1–18. https://doi.org/10.1111/mam.12015
Bray, J. R., & Curtis, J. T. (1957). An ordination of the upland forest communities of Southern Wisconsin. Ecological Monograph, 27, 325–349. https://doi.org/10.2307/1942268
